UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENVIRONMENTAL AND LIFE SCIENCES

Biological Sciences

Volume 1 of 1

Plasticity and adaptation in the multigenerational plant response to rising atmospheric CO_2 concentrations

by

Jasmine Mariette Saban

Thesis for the degree of Doctor of Philosophy

November 2018

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENVIRONMENTAL AND LIFE SCIENCES

Biological Sciences

Thesis for the degree of Doctor of Philosophy

Plasticity and adaptation in the multigenerational plant response to rising atmospheric CO_2 concentrations

Jasmine Mariette Saban

Rising atmospheric carbon dioxide concentrations ($[CO_2]$) will expose extant plant species to novel $[CO_2]$ on a global scale. In studies of plant acclimation and adaptation to elevated $[CO_2]$ diverse acclimatory phenotypes have been observed. Further investigation of gene expression profiles has begun to elucidate changes in expression that drive observed phenotypic changes. Although multigenerational studies of plant responses to elevated $[CO_2]$ suggest that plants do adapt to these conditions, evidence of causative changes to the underlying genetic sequence is limited and controversial.

Naturally occurring CO_2 springs provide a resource to study the contribution of plasticity, indirect and direct genetic effects to the multigenerational responses to elevated $[CO_2]$. Previously popular to study plant physiology in elevated $[CO_2]$, these sites can be revisited using new genetic and epigenetic technologies to answer fundamental questions about the multigenerational response of plants to elevated $[CO_2]$. Chapter two of this work synthesises available phenotypic data from plants at natural CO_2 springs to provide an overview of broad-scale trends in the multigenerational elevated $[CO_2]$ response. Experimental Chapters three and four utilise *Plantago lanceolata* L. at a natural CO_2 spring to examine mechanisms facilitating the multigenerational plant response to elevated $[CO_2]$. In Chapter three the contribution of parental and grandparental effects is examined through statistical analysis of a multigenerational experiment. Chapter four provides novel insight into the role of methylation and genetic sequence change in the multigenerational response of plants to elevated $[CO_2]$ using High Throughput Sequencing technologies. These analyses are valuable in their contribution to understanding how plants respond to elevated $[CO_2]$ over multiple generations, with critical impact to the prediction of plant response to climate change.

Table of Contents

List of Tables

List of Figures

Academic Thesis: Declaration Of Authorship

I, Jasmine Mariette Saban declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Plasticity and adaptation in the multigenerational plant response to rising CO_2 concentration under climate change

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

   Chapter 2: Saban, J. M., Chapman, M. A. & Taylor G. (2018) FACE facts hold for multiple generations: Evidence from natural CO_2 springs. *Global Change Biology*, *In Press.*

Signed: .....

Date: .....

Acknowledgements

A thesis is often viewed (and sometimes feels like) a solo enterprise, but science is a collective effort, and this thesis has been informed by the words of thousands of people through their published papers or through spoken word. I feel immensely privileged to have been a part of the academic community for the last four years. Here I extend my sincere gratitude to all of the people that have contributed to this thesis, recognising that I only have the space here to acknowledge a handful by name.

Firstly I must thank my funders, The National Environmental Research Council (NERC) through the SPITFIRE DTP and the University of Southampton. The SPITFIRE DTP funds innovative and multidisciplinary research into the environment, and has provided me with amazing training opportunities including five training courses and a three month academic placement to Australia. From this experience I not only learned new skills but the exposure to new people and new ideas gave me new energy, ideas and focus to do better research.

Next I want to thank Professor Gail Taylor for providing me with all the wonderful opportunities (and the wonderful data) that I've been so lucky to have on this journey. I have been fortunate to have her very experienced guidance on this thesis and I sincerely value her encouragement and support in my personal and professional development. I also thank my second supervisor, Dr Mark Chapman, who has provided almost daily support and advice on genomics and bioinformatics throughout my PhD, and who continues to challenge and nurture my work in these areas in equal measure. I am grateful to Dr Alex Watson-Lazowski and Dr Yunan Lin for all of the hard work that preceded my project, and particularly to Alex for his advice when I first started working with *Plantago*. Additional thanks to Dr Tom Ezard for asking me a difficult question about maternal effects in my first ever (knee-shaking) environmental biosciences seminar. Chapter three of this thesis stems from that question, and I am also extremely grateful to have had his guidance on the statistical analysis in that chapter. For organising our first visit to the CO_2 spring, for help with preliminary experiments and for early discussions about the project I also thank Kevin Staniland. Here I also acknowledge the use of the

University of Southampton supercomputer; Iridis, and thank the support staff there for the many times they've really helped me out of a bind.

This PhD has been an amazing life-changing experience both on a personal and professional level. All of Taylor Lab and Chapman Lab have been instrumental in that experience, providing considerable support throughout my four years here, and all of my most treasured memories of this research. My work has benefited from their practical help with lab and field work, and their intellectual input on analysis and ideas. Particular thanks of course to Suzie Milner who has always gone the extra mile, and pried me away from my desk when a break was very much needed, and to Anna, Rachael, Annabelle, Ollie, both Libbys, Zoe, Nikol, Asya, Hazel and all of the others.

As part of my placement in Australia I thank Ben Schwessinger, Steve Eichten, Miriam Schalamun, John Rathjen and Justin Borevitz for the considerable time and resources they spent on teaching me new skills, both in methylation analysis and in Oxford Nanopore sequencing, I hope to have my own MinION one day soon! The warm welcome I received there and their highly collaborative approach to science rejuvenated my energy and enthusiasm for this project.

I suspect that part of the reason that I was chosen for this PhD in the first place was because of a reference written for me by my MSc dissertation supervisor Laurence Hurst. I am sincerely thankful to him not just for this, but for first introducing me (with great patience) to the rewards of coding, and for a somewhat infectious passion for evolution. I have had a number of other wonderful teachers, for example Marianne Sullivan, that I also want to thank here for inspiring me, and teaching me that learning should never stop.

On a very personal note I owe this academic achievement to my Mum and Dad for their unwavering support, for teaching me the value of hard work and for always finding things to be curious about. My whole family (including the Ready's as my now extended family) have contributed to this work, in the construction of 80 tent frames two years in a row, the making of a chamber door with gloves to operate a porometer in a closed system, for proof reading, and for all the weekends of recuperation that meant I could get back to my work when my motivation had ground to a halt.

Finally, to my long-suffering partner Andrew Ready, words can never express my gratitude for all of your support and patience through this process. For excel wizardry, constructing plant tents, manually checking all my references, lending me your expensive waterproof jacket to get muddy on field work, listening to me endlessly talk about the most minute details of this project, recognising when I've had a hard day and getting extra dessert, and for all of the rest. I couldn't have done this without you and I think in the end we've faired slightly better than Sam and Frodo. Although at times this thesis has been an all-consuming obsession, I don't think I'll have too much trouble casting it into Mount Doom (handing it in to the graduate office) and heading home for a holiday!

Abbreviations

Abbreviation	Definition
A	Photosynthetic rate at growth [CO_2]
ABA	Abscissic acid
AFLP	Amplified fragment length polymorphism
ANOVA	Analysis of variance
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BUSCO	Benchmarking Universal Single-Copy Orthologs
C	Cytosine
CAM	Crassulacean acid metabolism
C:N	Carbon to Nitrogen ratio
CO_2	Carbon dioxide
[CO_2]	Carbon dioxide concentration
CG	A region of DNA where a cytosine nucleotide is followed by a guanine nucleotide in the 5'->3' direction (also often abbreviated to CpG)
CHG	A region of DNA where a cytosine nucleotide is followed by a cytosine, thymine or adenine, and then a guanine nucleotide in the 5'->3' direction i.e. CCG
CHH	A region of DNA where a cytosine nucleotide is followed by any combination of two of cytosine, thymine or adenine in the 5'->3' direction i.e. CAT
CI	Credible intervals

COBRA	Combined bisulfite restriction analysis
CTAB	Cetyltrimethyl ammonium bromide
DE	Differentially expressed
df	Degrees of freedom
DNA	Deoxyribonucleic acid
DMR	Differentially methylated region
DMS	Differentially methylated site
D_{xy}	Absolute divergence
ECA	Epidermal cell area
FACE	Free Air CO_2 Enrichment
FDR	False discovery rate
FPKM	Fragments per kilobase of transcript per million mapped reads
F_{st}	Fixation index
GATK	Genome Analysis ToolKit
Gb	Gigabase
GCL	Guard cell length
GLMM	Generalised linear mixed model
GO	Gene ontology
g_s	Stomatal conductance
GtC	Gigatonnes of carbon
$GtCO_2$	Gigatonnes of carbon dioxide
GVCF	Genomic variant call format

HTS	High throughput sequencing
IGA	Instituto di Genomica Apilicata
IPCC	Intergovernmental Panel on Climate Change
J_{max}	Maximum rate of electron transport
kb	Kilobase
LAI	Leaf Area Index
Lat.	Latitude
LD	Linkage disequilibrium
Leaf N content	Leaf Nitrogen content
lnR	Log response ratio
logFC	Log Fold Change
Long.	Longitude
MCMC	Monte Carlo Markov Chain
MeDip	Methylated DNA immunoprecipitation
mRNA	messenger RNA
MS	Methylation sensitive
MSAP	Methylation sensitive amplified polymorphism
N50	The minimum contig length needed to cover 50 % of the genome
NERC	National Environmental Research Council
NGS	Next generation sequencing
PAR	Photosynthetically active radiation
PBAT	Post bisulfite adaptor tagging

PC	Principal component/coordinate
PCA	Principal component analysis
PCoA	Principle coordinate analysis
PCR	Polymerase Chain Reaction
pMCMC	p-value for parameter estimate in MCMC glmm
ppm	Parts per million
QTL	Quantitative trait loci
RCP	Representative concentration pathway
RNA	Ribonucleic acid
RNA-Seq	RNA Sequencing
rRNA	Ribosomal RNA
RuBisCO	Ribulose-1,5-bisphosphate carboxylase/oxygenase
SD	Stomatal density
SI	Stomatal index
SLA	Single leaf area
SMRT	Single molecule real time
SNP	Single nucleotide polymorphism
T	Thymine
TAIR10	The Arabidopsis Information Resource version 10
TES	Transcription end site
TMM	Trimmed mean of M values
TNSC	Total non-structural carbohydrates

TSS	Transcription start site
VCF	Variant call format
V_{cmax}	Maximum carboxylation rate
WGBS	Whole genome bisulfite sequencing
WGS	Whole genome sequencing

Chapter 1 General Introduction

Prior to the mid-18th Century atmospheric CO_2 concentration [CO_2] had remained stable at between 180 and 300 ppm for at least 800,000 years (Luthi *et al*., 2008). Evidence from boron/calcium ratios in foraminifera further suggests that atmospheric [CO_2] has been less than 400 ppm for 14-23 million years of plant evolution (Tripati *et al*., 2009, Pearson and Palmer, 2000). Anthropogenic activities since the Industrial Revolution (notably the burning of fossil fuels and deforestation) has led to the accumulation of atmospheric [CO_2] at an accelerating speed, reaching 400 ppm in April 2014 (Figure 1.1). The IPCC now predicts that [CO_2] concentrations are likely to reach 720-1000 ppm by the end of the century in the absence of further mitigation (IPCC, 2014). Unlike precipitation patterns and local temperature, the increase in [CO_2] occurs with consistent spatial distribution globally, exposing plants in every part of the world to elevated [CO_2]. This increase in carbon availability for photosynthesis will undoubtedly impact the growth and development of plants. The utilisation of additional carbon by plant species may affect community composition and ecosystem dynamics (Barnaby and Ziska, 2012) as well as global food security (Long *et al*., 2006; Wheeler and von Braun, 2013; Myers *et al*., 2014). The ability of the world's vegetation to act as a sink for carbon may also be ameliorating the effects of increased atmospheric [CO_2]. With global temperatures increasing by 3.7 - 4.8 ˚C by 2100 under the worst case scenarios (IPCC, 2014), and increasingly unpredictable rainfall patterns, it is imperative to understand the impact of elevated [CO_2] at both the ecosystem level and to understand plant responses at the molecular level over many generations.

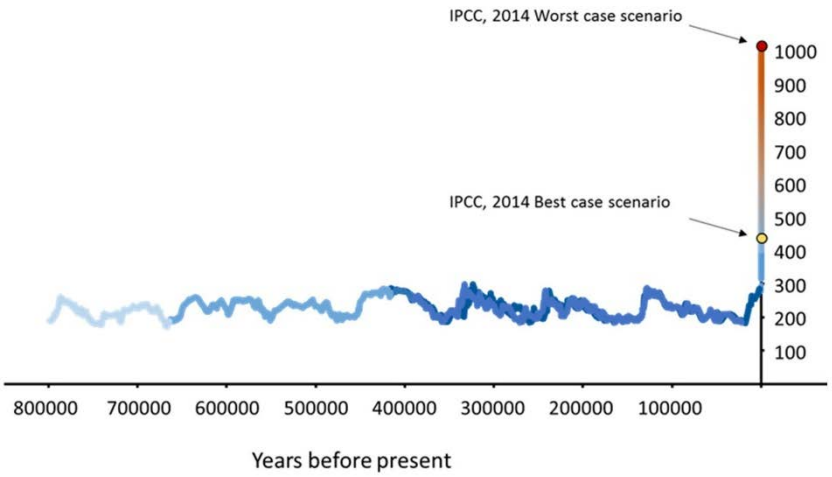


Figure 1.1 Increasing atmospheric CO_2 concentrations from the last 800,000 years.

Nine sets of data including ice core records (Neftel et al., 1994, Petit et al., 2001, Monnin et al., 2001, Barnola et al., 2003, Siegenthaler et al., 2005, MacFarling Meure et al., 2006 and Luthi et al., 2008) and historic sampling records (Tans and Keeling, 1958-2014) have been used to reconstruct the [CO_2] of the last 800,000 years. Predictions for the end of the century are highlighted with yellow (best case scenario) and red (worst case scenario) circles (IPCC, 2014).

1.1.1 The 'closing door' of emissions targets and the status of [CO_2] research

CO_2 emissions targets are set according to their degree of resultant warming, since global warming is a significant threat to the biome (IPCC, 2014). To reach these temperature targets, cumulative global CO_2 emissions must remain below a quota. The emissions quota consistent with remaining below the 2 °C temperature target (the 'best case scenario') will be exhausted within 30 years at the rate of emission of the year 2014 (Friedlingstein *et al*., 2014). However, rates of emission have been increasing by around 2.5% over the past ten years, and as Friedlingstein *et al*. (2014) argue, these figures and predicted economic growth are more consistent with the worst-case scenarios of 720-1000 ppm by the end of the century (Figure 1.2). Therefore, until a new and more stringent climate agreement is reached (next meeting 2020) this range of [CO_2] will be used to inform research into plant responses to elevated [CO_2] throughout the following experiments.

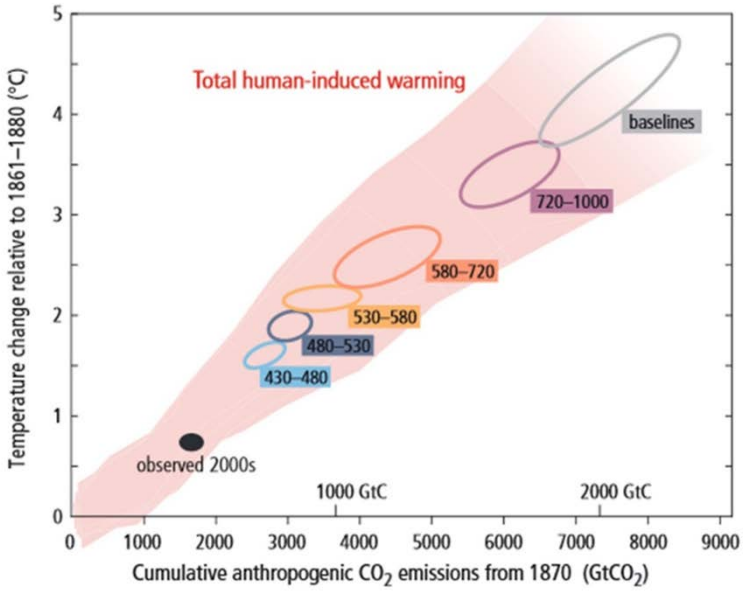


Figure 1.2: From IPCC (2014); Summary for Policy Makers. Emissions scenarios plotted by given net CO_2 cumulative emissions against the resulting global mean temperature caused by that emission scenario.

Coloured plume represents past and present projections from a number of different models, fading with decreasing availability of models. Filled black ellipsis shows observed cumulative emissions of 2005 and average temperature between 2000 and 2009. Other ellipses demonstrate emissions scenarios with associated atmospheric [CO_2] ranges (ppm) in boxes.

Since plants are static and atmospheric [CO_2] is increasing globally, plants are limited in their potential to respond to elevated [CO_2]. The closing door on emissions targets makes it more important than ever to understand long-term and multigenerational plant responses to elevated [CO_2]. A particular concern is the threat to human nutrition, since crop yields under elevated [CO_2] have been shown to be overestimated by early studies (Long *et al*., 2006) and some crops have been shown to decline in nutritional value (Myers *et al*., 2014).

There have been widely observed differences between how different species and plant types respond to elevated [CO_2] (Ainsworth and Long, 2004; Ainsworth and Rogers, 2007) and this is an important consideration when modelling general trends in plant responses to elevated [CO_2], and a constraint on generalising those responses. A simple example is the effect of life history traits; a long-lived tree may experience a doubling of [CO_2] within its lifetime, while a biennial may experience only very small increases in its lifetime – but will have a much shorter generation time for selection to act upon. This makes it

imperative to expand research into plant response to elevated [CO_2] to previously unstudied species, particularly endangered and keystone species.

Currently used methods for studying plant responses to elevated [CO_2] include both closed and open topped chamber experiments, Free Air Carbon dioxide Enrichment (FACE) facilities and natural CO_2 springs. In accordance with the speed and increasing awareness of anthropogenic impact on the global climate, research into the response of plants to IPCC predicted [CO_2] has increased in the last two decades, reaching a peak number of publications in 2001 (Figure 1.3). Since then funding has been decreasing while long-term studies into ecosystem responses remain lacking and the mechanisms by which plants sense and respond to elevated [CO_2] are yet to be fully elucidated (Becklin *et al.*, 2017). This leaves a significant gap in our understanding and ability to predict and prepare for future climate change scenarios.

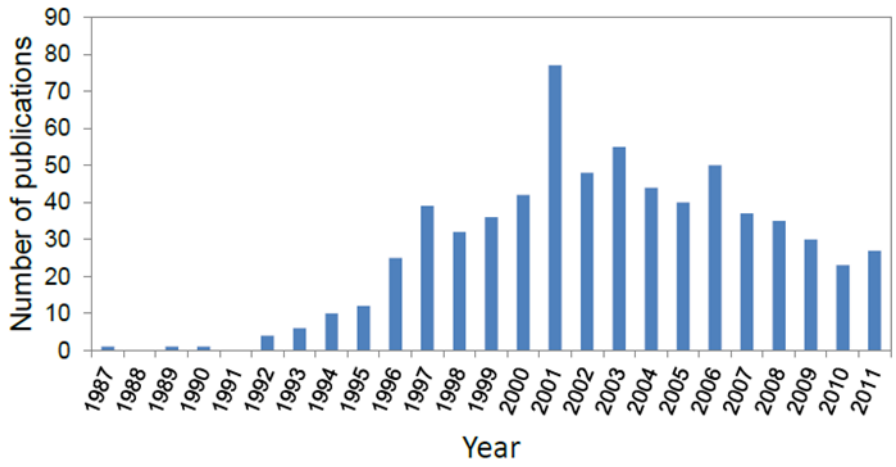


Figure 1.3: The number of publications in the field of plant response to elevated [CO_2] in the last 20 years, Jones *et al*. 2014

1.2 Plant response to environmental change

Broadly, plant responses to environmental change can be split into three categories; plastic responses, indirect genetic responses and direct genetic responses (Figure 1.4). Although not mutually exclusive the key difference between these types of response are the timescale at which they act. Plastic responses may be defined as the ability of a specific genotype to express altered phenotypes (biochemical or physiological) in different environments (Gratani, 2014). These responses occur within a plants lifetime, and are generally reversible, though they may be irreversible particularly if occurring during plant early development (Forsman, 2015). Direct genetic responses result from fixed changes to the genetics of individuals, and the differential success of these altered genotypes within a population over many generations, where the cost is demographic - some plants reproduce while others will not. Finally, indirect genetic effects are the causal effect of an individual's genotype or phenotype on the phenotype of an individual in the same species (Metz *et al*., 2015; Wolf and Wade, 2009). This includes within-generational and transgenerational indirect genetic effects, an example of within generational effects include a plant that produces a shadow over another plant and thereby affects shade avoidance in the plant it shadows. This influences the plastic expression of traits as part of the plant environment. Transgenerational indirect genetic effects include effects such as maternal effects, where the maternal genotype has a causal effect on offspring phenotype beyond the genotype inherited by the offspring (Wolf and Wade, 2009). Transgenerational indirect genetic effects can therefore both plastically respond to the environment and affect the phenotype of the next generation.

Plastic, indirect genetic effects and evolved direct genetic responses are not mutually exclusive of each other (Nicotra *et al*., 2010) and all may or may not be adaptive (Ghalambour *et al*., 2007). The term adaptation is reserved to describe adaptive direct and indirect genetic effects on an organism in a given environment (Table 1.1). All three may contribute to the same phenotype (i.e. increased size) but the mechanism by which this occurs is different, for direct genetic effects this phenotype may be mediated by a mutation in a gene promotor, for transgenerational indirect genetic effects this could be increased maternal provisioning to offspring, and for plastic responses this could purely be a consequence of increased resource availability. However over multiple generations these mechanisms interact extensively via complex molecular networks and through

evolutionary processes to coordinate the multigenerational plant response to the environment.

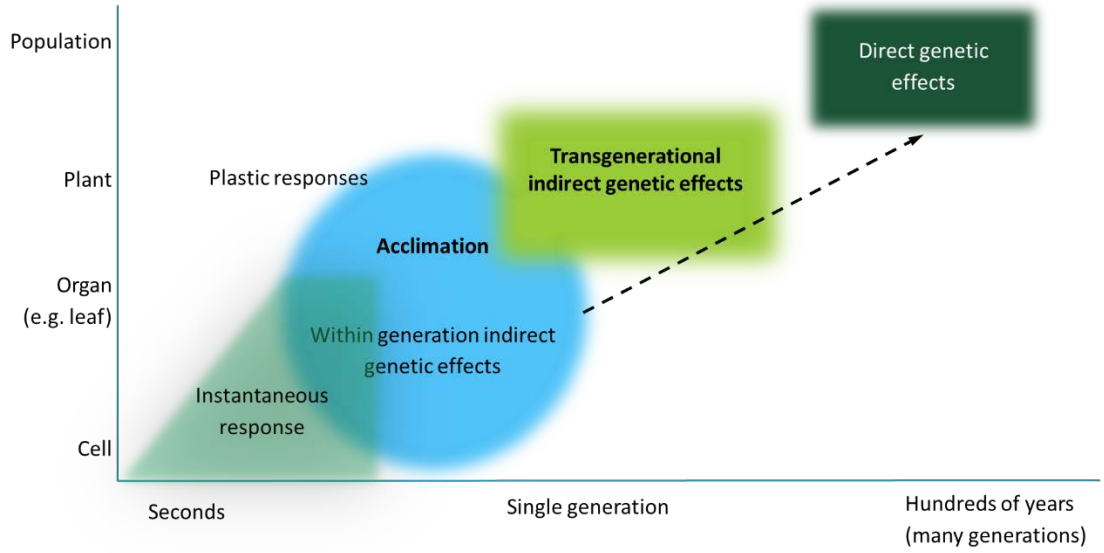


Figure 1.4: Plant responses to environmental change.

The y-axis indicates the level at which the response takes place, from individual cell up to population. The x-axis shows the time taken for the response to materialise. The arrow linking acclimation and evolved responses should be interpreted as 'may lead to', -for example by genetic assimilation or accommodation.

Phenotypic plasticity – The ability of a specific genotype to express different phenotypes in different environments (Gratani, 2014).

Adaptation –A derived heritable characteristic that enhances survival or reproduction of an organism in a given environment (Ghalambor *et al*., 2007, Futuyma, 2009).

Evolution – The change in frequency of alleles within a population over time (Ghalambor *et al*., 2007).

Direct genetic effects – The causal effect of an individual's inherited genotype on its own phenotype (Falconer *et al*., 1996)

Indirect genetic effects - The causal effect of an individual's genotype or phenotype on the phenotype of an individual in the same species (Wolf *et al*., 1998).

Within-generation indirect genetic effects – Indirect genetic effects that effect the phenotype within the same generation as the effector.

Transgenerational indirect genetic effects (also 'transgenerational effects', 'parental/grandparental effects' or 'transgenerational plasticity') - The causal effect of ancestral phenotype or genotype on offspring phenotype, beyond the directly inherited genotype (Metz *et al*., 2015, Wolf and Wade, 2009).

Epigenetic- the study of phenomena and mechanisms that cause chromosome-bound, heritable changes to gene expression that are not dependent on changes to DNA sequence (Deans and Maggert, 2015, Eichten *et al*., 2014).

Epigenetic memory - The mitotic and meiotic transmission of chromatin modifications within a plant lifetime in response to environmental change (Crisp *et al*., 2016).

Epigenetic inheritance – The stable transgenerational transmission of chromatin modifications.

Table 1.1: Definition of terms used to describe processes contributing to the multigenerational plant response to the environment.

1.3 Plasticity in plant response to elevated [CO_2]

In recent years plasticity has been recognised as a crucial factor in the response of specific plant traits to environmental conditions (Gratani, 2014; Hendry, 2015). The ability of a plant to respond plastically depends on its ability to sense changes to environmental conditions and initiate signal cascades that result in altered phenotype (Nicotra *et al*., 2010). Thus plasticity has a functional genomic component and is mediated by molecular interactions (Pigliucci, 2003). Many questions remain about how plants directly sense elevated [CO_2] at the molecular level, and the genetic architecture of phenotypic traits responsive to elevated [CO_2] has yet to be fully elucidated (Engineer *et al*., 2016), though much progress has been made in recent years through the combination of next generation sequencing (NGS) and functional genomic studies (Gamage *et al*., 2018).

Plastic responses can often be temporally separated into instantaneous and acclimatory responses, and this is a particularly important consideration in examining plastic plant responses to elevated [CO_2]. The instantaneous response of plants to elevated [CO_2] occurs within seconds to minutes, a sudden influx of CO_2 increases photosynthetic rate by providing increased substrate availability. However, over hours to days a decline in photosynthetic rate stimulation has been evidenced in a large number of studies in C_3 plants, a phenomenon termed photosynthetic acclimation (see Gamage *et al*., 2018). Whilst early studies indicated that this acclimation could completely negate the stimulation of photosynthetic rate (Sage, 1994), more recent experiments have concluded that photosynthetic rate largely remains enhanced despite acclimation (see Leakey *et al*., 2009). Mechanisms that have been proposed to coordinate photosynthetic acclimation, include CO_2 enrichment inhibition of shoot nitrate assimilation (Bloom *et al*., 2012), non-specific reduction of nitrogen partitioning in leaf blades (Seneweera *et al*., 2011), reduced sink capacity (Ainsworth *et al*., 2004) and carbohydrate suppression of protein synthesis (Cheng *et al*., 1998). In reality acclimation may well be achieved by a combination of these mechanisms. Since plant response to a sudden increase in experimental [CO_2] does not reflect the gradual increase in atmospheric [CO_2] expected under climate change, here we focus on acclimatory rather than instantaneous responses to elevated [CO_2]. Further, around 15 % of plant species have evolved a carbon concentrating mechanism to optimise photosynthesis (C_4 and CAM photosynthesis) and these plants generally do not see the same stimulation in photosynthesis (Ehleringer *et al*., 1997) but here we focus on

the responses of the 85 % of plants without a carbon concentrating mechanism (C_3) except where expressly stated.

Nearly all plastic plant responses to elevated [CO_2] in traits such as water use efficiency, nitrogen use efficiency and biomass accumulation occur as downstream consequences of the enhancement of photosynthetic rate (Leakey *et al*., 2009). Plant acclimatory phenotypes have been relatively well studied in a variety of settings, both natural and artificial and a number of trends have been identified that affect almost every aspect of plant morphology and development (Figure 1.5).

The expression of plasticity may play a key role in plant adaptation to elevated [CO_2] since at the population level, plasticity plays two vital roles in the survival and fitness of plant populations. In the short term it may facilitate population persistence by buffering them against selection, while in the long term it may affect adaptation potential, as a source of novel phenotypes for evolution or being subject to selection (Lande, 2009). For example, higher baseline gene expression allows for greater 'fine-tuning' of gene expression to variable conditions (Murren *et al*., 2015). On the other hand, adaptive trait plasticity in a novel environment could reduce or negate selection, reducing the evolutionary potential of the species and leading to population decline (Fusco and Minelli, 2010)

Whether plants will continue to respond plastically to elevated [CO_2] is unclear. Reduced responsiveness of plant traits to [CO_2] of the future relative to past [CO_2] (Dippery *et al*., 1995; Ward and Strain, 1997; Temme *et al*., 2015) supports the hypothesis that plants will exhaust their potential to plastically respond to elevated [CO_2] under climate change, as a result of the saturation of C_3 photosynthesis at around 1000 ppm (Korner, 2006; Ward and Strain, 1997). However these observations are highly species dependent with some species showing proportional increases in biomass up to and beyond predicted concentrations for the end of the century (Rasse *et al*., 2005).

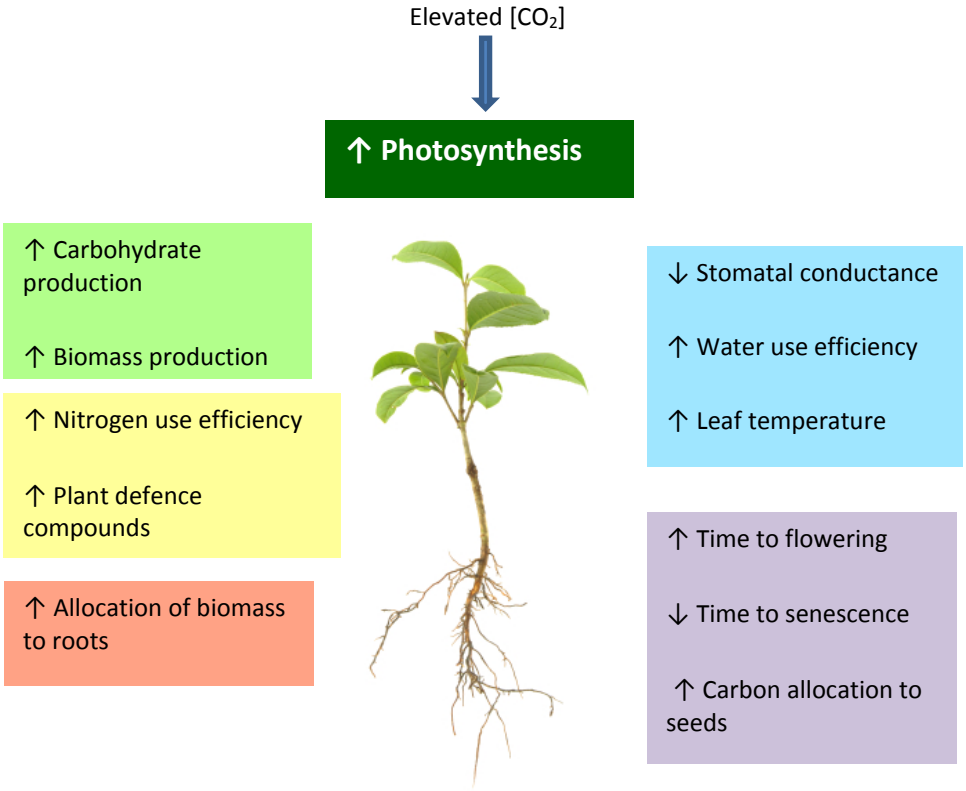


Figure 1.5: Generally observed phenotypic responses to elevated [CO_2].

All plant response begins with the increase in CO_2 availability to photosynthesis. This results in enhanced carbon assimilation, despite photosynthetic acclimation. Increased carbon assimilation provides the building blocks for increased growth and can alter secondary metabolism, including increased nitrogen use efficiency and increased production of defence compounds. Nutrients and water availability may become rate limiting to growth. Since there is increased substrate availability for photosynthesis, stomata may be increasingly closed to minimise water loss and stomatal density may subsequently be downregulated in developing leaves. Increased growth and increased availability of carbohydrates can lead to altered phenology.

1.3.1 High Throughput Sequencing (HTS) methods to study plant acclimation to elevated [CO_2]

Experiments characterising the plastic response of plants to elevated [CO_2] have focussed on plant physiological and morphological responses, but the increasing availability and affordability of HTS tools is facilitating further dissection of the mechanisms coordinating plastic responses (De Souza *et al.*, 2016). Comparative transcriptomics provides insight into plasticity (and adaptation in appropriate experimental designs) by quantifying gene expression changes associated with a treatment. Transcriptomics is the study of all of the RNA transcripts produced in a given tissue or individual at a given developmental stage

and in a given environment (DeBiasse and Kelly, 2016). Of specific interest is the abundance and diversity of messenger RNA (mRNA) which provides quantification of gene expression, and acts as a proxy for protein production and therefore cell functioning and response. Studying gene expression changes of plants growing under elevated [CO_2] therefore facilitates the identification of modular changes and key genes associated with specific traits and pathways responsive to elevated [CO_2] (Gupta *et al*., 2005; Ainsworth *et al*., 2006; Taylor *et al*., 2005). Increasingly RNA Sequencing (RNA-Seq) has been utilised to identify genes with differential expression in plants exposed to elevated [CO_2] (Tallis *et al*., 2010; Watson-Lazowski *et al*., 2016; Lin *et al*., 2016; Kumar *et al*., 2017). Broadly, plants typically show vastly different gene expression profiles when grown under elevated relative to ambient [CO_2], that are characterised by many small changes in gene expression, which can make detection of these changes difficult (Watson-Lasowski *et al*., 2016).

RNA-Seq has now largely superseded the earlier hybridisation based methods to analyse differential gene expression for reasons outlined in Table 1.2 (Wang *et al*., 2009). It is important to note that most studies in ecology and evolution that perform RNA-Seq analysis only analyse mRNA and not other types of RNA that may play regulatory roles in gene expression, such as microRNA.

	Microarray	**RNA-Seq**
Price	Low	High
RNA analysed	Specified by the user through the design of hybridisation probes	Any RNA that is transcribed
Identification of genes expressed	Known through binding of transcripts to probes	Identified through BLAST search
Detection of sequence differences	Sequence differences affect hybridisation	Sequences differences can be detected and analysed with efficient assemblers
Quantification	Relative intensities	Relative counts
***De novo* analysis**	Not possible	Possible
Multiplexing samples	Not possible	Possible
Amount of RNA required	High	Low

Table 1.2: A comparison of microarray and RNA-Seq approaches to analysing gene expression.

Technical considerations in RNA-Seq

Technical considerations of sources of error in RNA-Seq are key to robust interpretation of gene expression changes. Analysis of plasticity and adaptation in this thesis makes use of this technology, so these technical considerations are outlined briefly here. There are three key sources of error associated with transcriptome sequencing: biological variation, Poisson counting error and non-Poisson technical variance (Busby *et al*., 2013; Todd *et al*., 2016). All three of these sources can be addressed to some extent through experimental design and good processing practice.

Errors associated with biological variation can be addressed through appropriate biological replication, with a minimum of three replicates per treatment (with individual rather than pooled library preparation), and six replicates recommended by extensive analysis (Schurch *et al*., 2016). Since increased biological variation can arise from increased biological replication (particularly with the heterogeneity of field conditions), reliable and reproducible data requires that samples for RNA-Seq come from individuals in experiments that control for as much variation as possible, including through randomisation and blocking (Fisher, 1935).

Poisson counting error is the increased uncertainty that expression differences seen in low count transcripts reflect true and replicable expression differences between treatment groups. Increased sequencing depth reduces Poisson noise, but with diminishing returns (Liu *et al*., 2013; Wu *et al*., 2014). At 5-20 Million reads per sample and at least three samples per treatment, there should be sufficient depth to quantify differential transcripts with accuracy (Todd *et al*., 2016). Beyond this plateau in returns, the trade-off between sequencing depth and number of replicates makes it more efficient to invest in increased sample size (Lin *et al*., 2013).

Non-Poisson technical variation is the variation in expression differences that can be seen if the same sample is prepared twice. This variation is difficult to address as it can arise from differential processing of samples or from inherent properties of the RNA sample such as expression landscape, transcript length and depth of coverage (Todd *et al*., 2016). Only a very small fraction of the total RNA sample prepared for sequencing is actually sequenced and since biased degradation can occur, fast collection and freezing of tissue, uniform storage and processing of all samples is imperative (Romero *et al*., 2014).

Additionally samples from different treatment groups need to be distributed across flow cells and lanes to minimise the effects of sequencing bias.

The optimal assembly pipeline for mRNA sequencing reads depends on the availability of genomic resources. If there is a reference genome, reads can be mapped to the genome with a gapped mapper and stitched together into transcripts (Conesa *et al*., 2016). Newer genome guided approaches such as the genome-guided function in Trinity (Haas *et al*., 2013) will map reads to the genome, partition these reads into clusters according to locus, and then *de novo* assemble at each locus. If there is no reference genome available then the transcriptome can be assembled *de novo* through the construction of De Bruijn graphs in assemblers such as SOAPdenovo-Trans (Xie *et al*., 2014), Trans-ABySS (Grabherr *et al*., 2011) and Trinity (Haas *et al*., 2013).

Future analyses of plant responses to elevated [CO_2] are expected to take advantage of the decreasing cost of RNA-Seq and increasingly well-established pipelines for analysis of RNA-Seq data in non-model organisms (Ge *et al*., 2018; Dhami *et al*., 2018; Gamage *et al*., 2018). However, in recognising the limitations of this approach, it must be noted that gene expression as quantified by RNA-Seq does not necessarily translate into quantity of protein production, where post-transcription regulation of gene expression plays a role (Wang and Zhang, 2014). Therefore the use of proteomics to quantify protein abundance and diversity is needed (Voelckel *et al*., 2017), and functional analysis of genes of interest identified through transcriptomics are an important step in fully elucidating complex traits responsive to elevated [CO_2] (Gray *et al*., 2000).

1.4 Transgenerational indirect genetic effects in plant response to elevated [CO_2]

The evolution of transgenerational indirect genetic effects is adaptive if it can increase fitness of offspring in the environment in which they grow, providing a rapid and flexible mechanism for transgenerational adjustment of phenotype to environmental conditions (Uller, 2008). Increasingly the role of transgenerational indirect genetic effects in the expression of offspring phenotype have been explored within theoretical models (Hoyle and Ezard, 2012; Ezard *et al*., 2014; English, 2015; Leimar and McNamara, 2015) and with empirical data (Galloway and Etterson, 2007; Herman and Sultan 2011; Herman et al, 2012; Herman and Sultan, 2016). These studies have identified components that predict the evolution of transgenerational indirect genetic effects (see Auge *et al*., 2017), including developmental time lags between an environmental cue and selection on progeny (Ezard *et al*., 2014), the nature of environmental change (especially predictability) (Ezard *et al*., 2014; Kuijper and Hoyle, 2015) and accuracy with which parental environment can predict that of the offspring (Leimar and McNamara, 2015). Further, the contribution of transgenerational indirect genetic effects in shaping evolutionary trajectories under climate change is of key interest to predicting multigenerational plant responses (Herman and Sultan, 2011). The sessile nature of plants predicts that transgenerational indirect genetic effects are likely to evolve to a greater extent in plants than animals (Galloway, 2005).

1.4.1 Mechanisms of transgenerational indirect genetic effects

Initially the study of transgenerational indirect genetic effects focussed on the direct effect of altered resource provisioning, but increasing identification of specialised mechanisms has expanded this view (Herman and Sultan, 2016). Mechanisms that constitute transgenerational indirect genetic effects in the context of plant response to elevated [CO_2] can broadly be split into three categories, those relating to altered phenology of the maternal plant (which in turn alters the environment experienced by the seedling), seed resource provisioning and modifications, and epigenetic inheritance.

Some transgenerational indirect genetic effects of elevated [CO_2] on plant offspring have been widely recorded, and occur as a downstream consequence of enhanced carbon

assimilation. For example, elevated [CO_2] can alter flowering phenology, seed size (Jablonski *et al*., 2002), weight (Hussain *et al*., 2001; Darbah *et al*., 2007), quality (Cotrufo *et al*., 1998; Myers *et al*., 2014), viability (Farnsworth and Bazzaz, 1995) and germination properties (Thurig *et al*., 2003; Andalo *et al*., 1996) (Table 1.3). Meta-analyses indicate that these transgenerational indirect genetic effects are highly variable between species in response to elevated [CO_2] but also within species (Nakamura *et al*., 2011; Springer and Ward 2007). Thus a greater mechanistic understanding of the basis of these traits is necessary to fully understand the response to elevated [CO_2] over multiple generations.

Quantification of transgenerational indirect genetic effects in plant response to environmental change has traditionally focussed on maternal effects because maternal environment influences offspring phenotype through both pre- and postzygotic effects, as well as through seed dispersal, and therefore are considered to be more influential (Mousseau and Fox, 1998). Paternal effects have generally been assumed to be negligible (Roach and Wulff, 1987) and typically experimental design does not allow for the discrimination of paternal from maternal effects (Lacey, 1996). The significance of paternal effects in influencing offspring traits is highlighted by Lacey (1996). In this study controlled crosses *of Plantago lanceolata* were exposed to different pre- and postzygotic temperatures, and although postzygotic temperature effects were stronger than prezygotic parental effects, prezygotic effects of temperature on offspring traits were more likely to be transmitted paternally than maternally. Thus experimental design that facilitates the discrimination of paternal and maternal effects is key to furthering our understanding of the molecular mechanisms by which transgenerational indirect genetic effects are conferred. This is particularly important in understanding plant response to elevated [CO_2] where the contribution of paternal effects have not been assessed.

Grandparental effects can also influence offspring phenotypes, but are expected to be weaker than parental effects where they have evolved, because of the increased temporal lag between the response to the environmental cue and the production of the offspring phenotype (Groot *et al*., 2017). Grandparental effects and to a lesser extent paternal effects are not able to directly influence the environment of a seed in the same way that mothers can. Therefore mechanisms of indirect influence, particularly epigenetic inheritance are expected to facilitate these effects (Baker *et al*., 2018).

Transgenerational effect	Average response	Example reference	Species
Flowering time	Not summarised, but acceleration, deceleration and no change observed	(Springer and Ward, 2007)	90 crop and wild
Number of seeds	Increased	(Jablonski *et al*., 2002)	79 crop and wild
Seed size, mass or weight	Increased	(Jablonski *et al*., 2002)	79 crop and wild
Seed carbohydrate content	No change	(Grünzweig and Dumbur, 2012)	17 species
Seed protein content	Decrease	(Myers *et al*., 2014)	6 crops
Seed lipid content	No change	(Thomas *et al*., 2009) (Frick *et al*., 1994) (Ekman *et al*., 2007)	Kidney bean, *Brassica napus, A. thaliana*,
	Increased	(Uprety *et al*., 2010) (Hussain *et al*., 2001)	*Brassica juncea*, *Pinus taeda*,
Seed mineral nutrient content	Decreased	(Myers *et al*., 2014)	6 crop species
Germination rate	Increased	(Marty and BassiriRad, 2014)	30 crops and wild species
Viability	No change	(Marty and BassiriRad, 2014)	47 crops and wild species
Seed dispersal	Increased	(McPeek and Wang, 2007)	Dandelion
Seed coat thickness/permeability	No change	(Chen *et al*., 2015)	Rice
Gene expression	Varied	(Thomas *et al*., 2003)	Soybean

Table 1.3: A summary of transgenerational effects studied in plant response to elevated [CO_2].

Epigenetic inheritance has never been studied in plant response to elevated [CO_2]. Epigenetics is a broad scientific concept and a widely used term but the use of different definitions in different situations has confounded interpretations and methodologies (Deans and Maggert, 2015). Here we define epigenetics in accordance with Dean and Maggert, (2015) and Eichten *et al*., (2014) as 'patterns of chromosome-bound phenotypic variation that are heritable and independent of changes to DNA sequence'. The caveat that epigenetic modifications are both heritable and independent of genetic variation requires some *a priori* knowledge about the nature of a chromatin modification before it can be defined as epigenetic. As a result, here we will largely refer to chromatin modifications, which may occur by the same mechanisms as those in epigenetics but make no assumptions about the heritability of the modification, and epigenetic inheritance, which here is used to describe chromatin modifications that are heritable but not necessarily independent of genetic variation (Table 1.1).

1.4.1 Statistical methods to study transgenerational indirect genetic effects in plant response to elevated [CO_2]

Despite their importance, transgenerational indirect genetic effects are often neglected in the study of plant response to elevated [CO_2]. For many decades, transgenerational indirect genetic effects were seen as a nuisance source of statistical noise in experiments tracking the heritability of traits (Falconer *et al*., 1996). As such, studies looking to disentangle phenotypic plasticity and genetic adaptation usually attempt to minimise the contribution of transgenerational indirect genetic effects, typically by growth in a common [CO_2] for at least one generation prior to crossed factored analyses (Watson-Lazowski *et al*., 2016). This reduces the potential for biased interpretation of direct genetic effects that is actually an artefact of plasticity induced transgenerational indirect genetic effects (Latzel, 2015). Many studies have not however taken advantage of this experimental design to also study the role of transgenerational indirect genetic effects in their study system by monitoring traits throughout the generation in the same environment. This would provide more information about the reversibility of traits, and thus the potential contribution of transgenrational indirect genetic effects.

Since phenotypic variation has a genetic component that maybe shaped by natural selection, classical quantitative genetics methodology provides a powerful means with

which to explore the genetic architecture of phenotypic traits and how they are influenced by the environment and transgenerational indirect genetic effects (Kruuk *et al*., 2008). Under quantitative genetics the genetics of complex traits are modelled as being influenced by many genes and factors (Hill, 2010). Statistical models partition components of phenotypic variation attributed to genetic and non-genetic factors by describing the resemblance between relatives and between those in similar environments.

1.4.1.1 The animal model

The use of the 'animal model' to partition phenotypic variance into that attributed to heritability, transgenerational indirect genetic effects and plastic response to environmental change has largely superseded other approaches such as parent-offspring regression in recent years (Visscher *et al*., 2008). This is because of its increased flexibility, increased use of all available information on relationships through the construction of a pedigree, ability to model environmental and transgenerational indirect genetic effects with generalised linear mixed models (GLMM) and the increased tractability of the approach facilitated by newly developed software.

An animal model is a particular form of model in quantitative genetics where the additive effect of an individual genotype on phenotype expression ('breeding value') relative to the population average is included as an explanatory variable for the variation in that phenotype (Wilson *et al*., 2010). The animal model expresses the phenotype/trait (y) in every individual as depending on the mean value (μ) the breeding value (a) and residual deviation (r) as:

$$y = \mu + a + r$$

And this can be expanded to include other effects of interest, for example:

$$y_i = \mu + a_i + p_z + mat_d + pat_f + c_i + r_i$$

In this case, y_i is the record or observation of the trait of interest (for example; number of days to germination); μ is the overall mean of observations of the trait; a_i is the 'breeding value' which accounts for the additive genetic effect on the trait estimated from the pedigree; p_z is the permanent environmental effects which are included only if traits are

measured repeatedly in the same individual z; mat_d is the effect of the mother d; pat_f is the effect of the father f; c_i is the effect of a fixed environmental effect (such as CO_2 treatment) and r_i is the residual variance (Wilson *et al*., 2010; Noble *et al*., 2014).

Since we are interested in the components of variance contributing to variance in the trait, the variance components are estimated i.e.

$$Var(a_i) = \sigma_a^2$$

Estimates of heritability and parental effects are calculated by dividing the variance component by the sum of the other variance components in the model, (note that c_k and p_z are fixed effects and not a (random effect) variance component).

For example heritability is calculated as:

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{mat}^2 + \sigma_{pat}^2 + \sigma_r^2}$$

Assumptions of the animal model include that the population is large (infinite), association of alleles from parents is random, there is no influence of factors other than those modelled and that only additive genetic effects exist. In the animal model dominance and epistasis are generally assumed to make negligible contributions to the genetic component of phenotypic variation, further, quantifying them requires comparison of full and half siblings which is infrequently done with wild populations (Wolak and Keller, 2014). Heritability is therefore calculated as narrow sense heritability with the acknowledgement that it may be overestimated because of dominance or epistasis.

1.4.1.2 The animal model: Frequentist vs Bayesian statistics

Within the animal model, the statistical framework for modelling can be either Frequentist or Bayesian. A key difference between these two methods is that while Frequentist inference estimates the probability of data given a hypothesis, Bayesian inference estimates the probability that the hypothesis is true given the observed data (see Ellison (2004) for a summary) (Figure 1.6). Comparison of the two approaches has found that the Frequentist methods are quicker, more user friendly and perform better for data with Gaussian error structure, while Bayesian inference implemented as Monte Carlo Markov Chains (MCMC) generalised linear mixed models (GLMM) perform better for non-Gaussian traits (Ellison, 2004; Wilson *et al*., 2010). The major benefit of Bayesian inference over Frequentist is that uncertainty can be quantified in multivariate problems through analysis of a joint probability distribution, since both model parameters and data are modelled as random variables (Ovaskainen *et al*., 2008). The aim of Bayesian inference is to calculate the posterior distribution; the conditional probabilistic distribution of the parameters (e.g. maternal effects) to a degree of belief (Beaumont and Rannala, 2004). In this thesis we make use of the Bayesian inference approach because of the complex models of environmental, genetic and non-genetic effects of interest and the likelihood of missing observations (which requires complex marginalisations in frequentist methods) (Little and Rubin, 2002; Dorazio, 2016).

Figure 1.6: Bayesian inference.

Bayesian inference estimates the probability of a hypothesis being true given observed data and based on Bayes Rule

Bayes Rule:

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

In this equation θ denotes the parameter and y denotes the observed data (adapted from O'Hara *et al*., 2008). The equation states that the probability p() distribution that summarises the values the parameter of interest will likely take (p(θ|y) or posterior distribution) is proportional to the product of the likelihood of observing the data given the parameter (p(y|θ) or likelihood) and the distribution of the parameter before the data is observed based on prior belief (p(θ) or prior).

1.4.1.3 The animal model: MCMC algorithms

Many Bayesian inference models have posterior distributions that cannot be evaluated analytically and so must be approximated numerically (O'Hara *et al*., 2008). The development of MCMC algorithms to fit Bayesian models overcame the initial computational barrier to the widespread application of Bayesian methods to quantitative genetics (Dorazio, 2016). MCMC algorithms can simulate the posterior distribution of a model parameter (e.g. heritability or maternal effects) given the observed phenotypic data by randomly sampling the posterior probability of the distribution using Monte Carlo integration with Markov chains. Monte Carlo integration facilitates the sampling as a random subset rather than all combinations of mean and variance in parameter space. Markov chains are a sequence of variables, whereby the probability of the variable assuming a specified value is dependent on the values of specified variables that precede it in the chain (Gilks *et al*., 1995). The proposal to extend the Markov chain to a new position in parameter space is rejected or accepted probabilistically using the prior distribution and the observed data. Over sufficient number of iterations, the constructed Markov chain will converge to a stationary distribution equivalent to the posterior distribution.

The inclusion of priors is probably the most controversial component of Bayesian inference because of its seemingly subjective nature. Since priors inform the process for simulating the probability distribution, careful consideration of the most appropriate is required (Gill, 2014). For example, parameter expanded priors are useful for random effects to avoid poor mixing if parameter estimates are close to zero. The strength of the prior on the posterior distribution diminishes with increased sample size but calculated estimates should be checked that they are robust to prior specification. The use of MCMC algorithms necessitates diagnostics of the MCMC algorithm behaviour to check the convergence of the chain and the degree of autocorrelation in the chain for every parameter estimated.

Software that can implement MCMC GLMMs suitable for estimating quantitative genetic parameters include WinBUGS (Spiegelhalter *et al*., 2003), WOMBAT (Meyer, 2007) and the R package MCMCglmm (Hadfield, 2010) of which MCMCglmm may be the most

flexible because of the range of error distributions supported. This package also provides useful visual diagnostics to check convergence and autocorrelation.

1.4.2 HTS methods to study transgenerational indirect genetic effects: Methylation and epigenetic inheritance

Statistical models are useful to partition the component of variation attributed to transgenerational indirect effects. However they cannot provide information on the mechanisms coordinating these indirect genetic effects, and for this HTS methods can be employed. DNA methylation is one potential mechanism that could facilitate both transgenerational indirect genetic effects (through epigenetic inheritance), and the plastic response of plants to elevated [CO_2]. DNA methylation is a chromatin modification found extensively across plant and animal genomes, where it has a number of biological functions, including in transposon silencing, chromatin structural organisation and to some extent in gene regulation (He *et al*., 2011; Niederhuth and Schmitz, 2017). Methylation here means the addition of a methyl group (CH_3) to a cytosine residue within DNA. In plants unlike in animals it can be established in almost any cytosine sequence context. These are divided into three categories, the symmetrical contexts; CG and CHG, and non-symmetrical CHH (Law and Jacobsen, 2010). Methylation in different sequence contexts are established and maintained through independent mechanisms and have different functions (Niederhuth and Schmitz, 2017). Changes in methylation patterns have been extensively linked to changes in plant phenotypes, including in response to environmental cues (Das and Messing, 1994; Baulcombe and Dean, 2014; Bilichak and Kovalchuk., 2016; Kumar *et al*., 2013). Methylation in symmetrical contexts are maintained through DNA replication, with the methylated template strand initially resulting in hemimethylated double stranded DNA, which directs methylation of the nascent strand (Law and Jacobsen, 2010). Since CHH methylation is not symmetrical it has to be re-established *de novo* after replication (Gehring, 2016). Methylation marks can persist for multiple generations because there is no overt demethylation of CG and CHG methylation, and because CHH methylation is actively maintained through fertilisation (Heard and Martienssen, 2014).

A role for epigenetic inheritance of methylation patterns that are induced by the environment is an attractive hypothesis for a variety of reasons (Baulcombe and Dean,

2014) but remains controversial (particularly as a genome-wide mechanism for coordinating plant response) because of the lack of direct evidence (Boyko and Kovalchuk, 2008; Paszkowski and Grossniklaus, 2011; Pecinka and Scheid, 2012). Rather, methylation variation is increasingly associated with underlying genetic variation and the primary functions appear to be silencing transposons and increasing genome stability rather than in regulating gene expression over generations (Eichten *et al*., 2014).

Methylation patterns have never been studied in either plastic or adaptive plant response to elevated [CO_2] but a potential role for epigenetic inheritance can be inferred by analysing the decay of adaptive phenotypic traits through multiple generations at ambient [CO_2]. Physiological traits that are adaptive in elevated [CO_2] and are maintained in subsequent generations at ambient [CO_2] even when no longer adaptive may be indicative of either genetic adaptation or epigenetic inheritance. Where epigenetic inheritance is implicated it may be reversible in few generations (Becker *et al*., 2011) while the reversal of a change to the genetic sequence would be expected to be much rarer, happening less frequently over more generations (Furrow, 2014).

A number of different methods have been developed to study methylation (Table 1.4). For genome wide scale, single nucleotide resolution analysis, the current 'gold standard' technique is whole genome bisulfite sequencing (WGBS). Bisulfite conversion is a technique used in multiple other methylation analysis approaches (i.e. reduced representation bisulfite sequencing) that relies on the differential chemical interaction of sodium bisulfite with methylated versus unmethylated cytosines (Figure 1.7). In bisulfite conversion non-methylated cytosines (C) are converted to a uracil residue which become thymine (T) residues upon synthesis, while methylated cytosines are not converted. This technique effectively 'fixes' methylation at the point of conversion but relies on the effectiveness of conversion. Some studies have utilised whole genome bisulfite sequencing without a whole genome sequence for each individual, and then inferred methylation with an estimated margin of error (Platt *et al*., 2015) since C->T or T->C single nucleotide polymorphisms (SNPs) between individuals confound the identification of bisulfite converted sites. Combining whole genome and whole genome bisulfite sequencing for each individual is therefore considerably more accurate when there is genetic variation between sequenced individuals, although it effectively doubles the cost of sequencing.

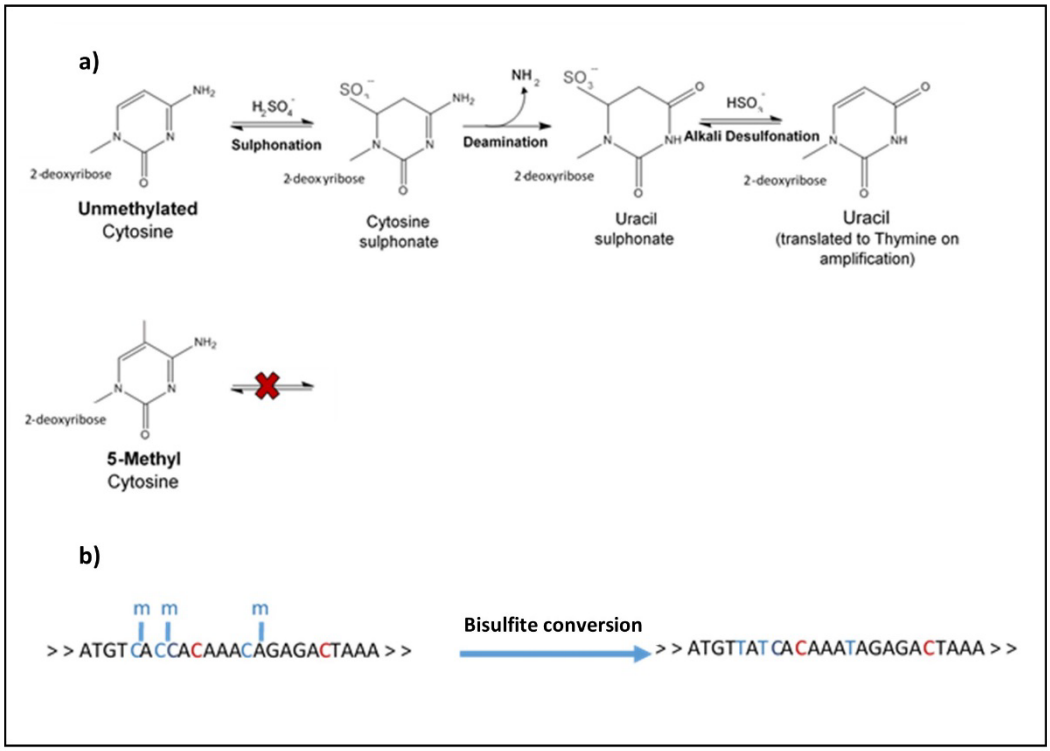


Figure 1.7: Bisulfite conversion of DNA.

The bisulfite anion will selectively convert unmethylated cytosine (C) to uracil upon alcohol desulfonation, methylation at the 5' carbon prevents the initial reaction step from occuring. A) The reaction mechanism for conversion of unmethylated C to uracil, later replaced by thymine (T) residues during amplification. B) DNA is sequenced both pre- and post- bisulfite conversion. The two sequences are compared. Where a C in the original sequence has remained a C in the converted sequence, this C was likely methylated. Where a T is present in the converted sequence but was a C in the original sequence, the C is presumed to be unmethylated.

Three approaches exist to create sequence libraries for bisulfite sequencing: MethylC-Seq, BS-Seq and Post Bisulphite Adaptor Tagging (PBAT). In this thesis we make use of PBAT because it is less costly and requires less input DNA than MethylC-Seq and because it is less bioinformatically challenging and computationally resource intensive to analyse than BS-Seq. Although the original PBAT method utilised costly biotinylated primers (Miura *et al.*, 2012), subsequent development has reduced the cost by replacing these with random hexamers. The sequencing libraries are made by first bisulfite converting the DNA which also fragments it to the required size distribution, and then adding the primers (in contrast to MethylC-Seq which adds primers and then bisulfite converts). The random hexamer primers bind to DNA fragments and these are amplified over two stages of strand synthesis to enable sequencing of small input samples. These libraries can then

be sequenced as for other types of sequencing, with the exception that sequencing requires a ~20% PhiX (single stranded bacteriophage DNA) spike-in to balance the nucleotide composition of the sample which facilitates greater sequencing throughput (Nair *et al*., 2018).

The conversion of unmethylated cytosine residues to thymine residues during synthesis results in a loss of information for assembly. In order to assemble WGBS reads, mapping occurs against two versions of the reference genome that are created *in silico*, one that represents a fully methylated genome following bisulfite conversion, and one that represents the fully unmethylated genome. The unmethylated version will have all C's replaced by T's and the methylated will retain all C's. As a result of the loss of information, mapping efficiency is 50-70 % lower than may otherwise be expected, which has to be anticipated in experimental design (Tran *et al*., 2016). Additionally bisulfite conversion is a harsh chemical process that results in degradation of the DNA, which means that it must be sequenced to greater coverage in order to meet a coverage target.

We utilise WGBS because it provides the most comprehensive overview of methylation currently available. However, as single molecule sequencing techniques become more popular and algorithms to detect the ionic signal of methylated DNA through nanopores are refined, bisulfite conversion will likely be phased out of use and replaced by this technology.

Technique	Methodology	Advantages	Limitations
Methylated-DNA immunoprecipitation (MeDIP)	DNA is incubated with monoclonal 5-methyl-C antibodies and Immunoprecipitated. Microarrays (MeDIP-chip) or deep sequencing (MeDIP-seq) is used to analyse the enriched methylated DNA.	• Different antibodies can be used to target methylation at different site contexts • Can be applied to low DNA concentrations (160ng)	• High costs associated with generating specific antibodies and microarray or sequencing. • Less powerful than bisulfite sequencing
Methylation sensitive restriction endonuclease	Digestion of sequence with one methylation sensitive (MS) and one non MS restriction endonuclease isoschizomer. Digestion can be followed by quantitative PCR or Southern blotting.	• Good for methylation patterns in organisms with no reference genome. • Simple and inexpensive	• Only methylation in sites recognised by the endonucleases will be targeted. • The surrounding sequence of MS restriction sites are unknown
Methylation sensitive amplified polymorphisms (MSAP)	DNA fingerprinting technique utilised for marker-trait association analysis. Uses MS restriction enzymes to identify methylation sensitive amplified fragment length polymorphisms	• Highly reproducible and reliable. • Requires similar expertise and equipment as AFLP.	• Different stabilities of loci not accounted for. • There may be no obvious connection between methylation of a site and the expression of a specific gene • Failure of MS enzymes to cut at a restriction site could be due to an epigenetic OR a genetic change
Combined bisulfite restriction analysis (COBRA)	After Bisulfite conversion, the DNA is PCR amplified with restriction enzymes that are differentially sensitive to TpG conversion	• Simple and inexpensive	• Coverage limited to sites targeted by the restriction enzymes. • Difficult to design primers for T enriched bisulphite converted sequences
Bisulfite sequencing	Application of sodium bisulphite selectively converts unmethylated cytosine into uracil. The Bisulfite treated DNA is PCR amplified and sequenced, then compared to the untreated DNA sequence.	• High sensitivity, single nucleotide resolution	• Expensive, time consuming and labour intensive. • Possibility of incomplete conversion - detected as falsely methylated residues • Possibility of DNA degradation

Table 1.4: A comparison of approaches to methylation analysis

1.5 Direct genetic effects in plant response to elevated [CO_2]

If the expression of a plastic response to elevated [CO_2] is limited beyond a specific [CO_2] under climate change, only adaptation to elevated [CO_2] will then be able to optimise the fitness of a population to that environment (Jump and Penuelas 2005). Experimental designs to study adaptation (as indirect and direct genetic effects) to environmental change include resurrection experiments, 'space for time' studies such as crossed factored, reciprocal transplant and common garden experiments (Franks *et al*., 2013). Understanding the basis of genetic adaptation to elevated [CO_2] is critical both for wild populations where genetic variation is unknown, and for crops since domesticated plants have undergone severe reductions in genetic variation (Dempewolf *et al*., 2014).

Studies that combine plants with multigenerational elevated [CO_2] exposure in FACE or natural CO_2 spring sites with reciprocal transplant and crossed factored experiments have largely concluded that elevated [CO_2] can drive genetic adaptation in natural systems because plants grown for generations in elevated [CO_2] show trait differences to those grown for generations in ambient [CO_2], that are not reversed in offspring grown in ambient [CO_2] (Barnes *et al*., 1997; Polle *et al*., 2001; Onoda *et al*., 2009; Nakamura *et al*., 2011; Grossman and Rice, 2014; Watson-Lazowski *et al*., 2016). Though this conclusion is not universal (Van Loon *et al*., 2016).

Genetic variation in traits responsive to elevated [CO_2] has been evidenced in a wide range of plant taxa (Nakamura *et al*., 2011; De Costa *et al*., 2007; Lindroth *et al*., 2001; Wieneke *et al*., 2004; Ziska and Bunce, 2000), and genetic variation in these traits has proven to be heritable in some cases (Schmid, 1996; Case *et al*., 1996) but direct evidence that elevated [CO_2] is a strong enough selection pressure to induce population level genotypic changes under natural selection is limited and controversial (Leakey and Lau, 2012). Even in selection experiments, evidence that elevated [CO_2] is a driver of contemporary evolutionary adaptation is variable and limited (Potvin & Tousignant 1996; Ward *et al*. 2000; Wieneke *et al*., 2004; Teng *et al*., 2009; Frenck *et al*., 2013; Grossman and Rice, 2014). While overall these experiments seem to support a role for elevated [CO_2] as a driver of contemporary evolution, this evolutionary response depends on the strength of natural selection on reproductive and fitness traits imposed by the environment, including interactions with competitors. Indeed some studies argue that

there are only small and frequently insignificant if any, direct effects of elevated [CO_2] on evolutionary trajectories in the absence of competitors (Bazzaz *et al*., 1995; Lau *et al*., 2007; Steinger *et al*., 2007).

Climate change is also expected to alter the frequency and intensity of interactions between species (Tylianakis *et al*., 2008) and response to elevated [CO_2] varies widely both within (Ainsworth, 2016; Sanchez-Gomez *et al*., 2016; Resco de Dios *et al*., 2016) and between species (Blackman *et al*., 2016; Arndal *et al*., 2018). Competition between individuals plays a critical role in evolution by increasing selective pressures on plants in a community to utilise resources optimally. Differences between how species respond to elevated [CO_2] is expected to facilitate changes in competitive interactions that have consequences for community composition. A well-studied example is the interaction between C_3 and C_4 species under elevated [CO_2] (Wand *et al*., 1999; Ainsworth and Long 2004). In Savannah grassland ecosystems elevated [CO_2] is expected to favour the growth of C_3 tree seedlings over C_4 grasses, resulting in a conversion of grassland to forest (Bond and Midgley, 2012). Further, elevated [CO_2] is expected to favour invasive C_3 weeds over commercially important C_4 crops (Ziska, 2003a; Ziska, 2003b). However even these well studied and supported predictions for future climates have found contradictions, highlighting the need for long-term experiments (Reich *et al*., 2018).

Within species elevated [CO_2] may be expected to alter genotypic fitness rank order (Ward and Strain, 1997; Bernacchi *et al*., 2003; Davey *et al*., 2003; Sekhar *et al*., 2015; Blackman *et al*., 2016), and in addition to consequences to individual fitness this differential response of genotypes could impact population and species level responses to climate change (Moran *et al*., 2016). In dioecious plant species there may be sex specific differences in response to elevated [CO_2], where previous studies indicate that male plants may react more flexibly to elevated [CO_2] while females are less responsive (Li *et al*., 2013; Zhao *et al*., 2011; Wang and Curtis, 2001) with implications for sex ratios, distribution and reproductive capacity of the species (Li *et al*., 2013).

1.5.1 HTS methods to study direct genetic effects in plant response to elevated [CO_2]

There are a large number of methods available to study genetic sequence differentiation between populations in order to identify the genetic basis of adaptation to environmental change (see Davey *et al*., 2011; Matz, 2017). For species with a well annotated genome,

approaches such as Quantitative Trait Loci (QTL) mapping can be employed to elucidate the genetic basis of complex traits responsive to elevated [CO_2] in natural environments. For example this has been applied to study the genetic architecture of leaf growth, development, quality and senescence (Rae *et al*., 2006) and stomatal and epidermal cell traits in poplar (Ferris *et al*., 2002) and a number of yield traits in rice (Fan *et al*., 2005). However for non-model species the development of genomic resources is often the first required step for detailed genetic analyses. Rapid advances in sequencing technology and bioinformatics analysis tools have increased the accessibility and affordability of acquiring these resources (particularly *de novo* whole genome sequences) for non-model organisms (Ellegren, 2014).

De novo whole genome assemblies are useful tools for a multitude of other analyses, by providing a reference to which to align RNA and DNA sequences that have less coverage (Matz, 2017). Multiple individuals can then be sequenced at low coverage and mapped back to the reference in order to identify regions of selection (Unamba *et al*., 2015). A combination of short reads and long reads with different insert sizes are best practice for whole genome assembly (Baker, 2012; Matz, 2017), although mammalian genomes have been assembled using solely short read technology where there is sufficient coverage (Li *et al*., 2010). For plants, the *de novo* assembly of genomes using short reads is limited because of the large proportions of repeat regions in plant genomes (Veeckman *et al*., 2016). For more complete assemblies, long read technologies, particularly mate-pair sequencing and single molecule real time (SMRT) technologies (such as those developed by Pacific Biosciences and Oxford Nanopore Technologies) provide the sequences necessary to assemble over these repeat regions (Mehrotra and Goyal, 2014). Although a complete and well-annotated *de novo* assembled genome is still out of reach for many, short read assemblies that provide contigs or even scaffolds for analysis can still provide useful references. In this thesis we use short reads with a single insert size to assemble contigs for further analysis and acknowledge that this results in a highly contiguous genome.

The sole use of short reads with a single insert size to *de novo* assemble genomes makes error correction steps critical for maximising the efficiency of the assembler, since the incorporation of errors in a contig can prevent extension (Baker, 2012). Error correction steps for both *de novo* genome assembly and whole genome sequencing analysis include

trimming of adapters and low quality bases, and removal of low quality and low frequency reads (Wajid and Serpedin, 2014). At most of these stages there are trade-offs between aggressive trimming that removes as much low quality data as possible, and trimming that maximises the sequence data available to the assembler (Marçais *et al*., 2015). Sources of error in whole genome sequencing analysis, such as sequencing and software errors and mapping artefacts, necessitates quality control at every stage of the workflow (Van de Auwera *et al*., 2013). Many short read sequence assemblers have been developed, and most employ a sequencing by hybridisation graph based approach (Narzisi and Mishra, 2011). All of these short read assemblers split sequences into k-mers of specified (pre-determined) length k. These k-mers are used to construct de Bruijn graphs, whereby each k-mer derived from the sequenced reads represents a node, and nodes are connected by edges if two k-mers overlap by k-1 nucleotides. The de Bruijn graphs are then condensed and used to construct contigs in the assembly (Simpson *et al*., 2009; Compeau *et al*., 2011; Luo *et al*., 2012). The efficiency of assembly is usually measured by assembly metrics such as average contig size and the minimum contig length needed to cover 50 % of the genome (N50) (Zhang *et al*., 2011). Another important determinant of assembly efficiency is the computational resources required.

Having assembled contigs of a genome, the analysis of low coverage whole genome sequencing is conducted by mapping short read sequences from individuals to the assembled contigs using alignment tools such as Bowtie2 (Li and Durbin, 2009; Langmead, 2012) or BWA (Li and Durbin, 2010). Following the mapping stage, improvement steps such as local indel realignment and base quality score recalibration are recommended in order to reduce the number of miscalls (Van de Auwera *et al*., 2013). Genomic variant discovery can then proceed using tools such as UnifiedGenotyper in GATK (McKenna *et al*., 2010) or bcftools in samtools (Lie *et al*., 2009), both of which employ Bayesian likelihood methods to infer the likelihood of a genome variant at a position and to call the genotype. Conducting population genomic statistical analysis of divergence across the genome then facilitates identification of regions of differentiation and those that have potentially undergone selection.

1.6 Study system: *Plantago lanceolata* at a natural CO_2 spring

1.6.1 Natural CO_2 springs

Access to plants growing for multiple generations at elevated [CO_2] remains a key challenge for the study of plant adaptation, owing to the time, energy and expense of growing plants for multiple generations under these conditions. Of all methods available, FACE facilities provide the best simulation of plant responses to elevated [CO_2] in the natural environment. However, these facilities are expensive and labour intensive to build and maintain, contribute to anthropogenic emissions, and cannot provide information on population responses to elevated [CO_2] over generations in the timescale needed to prepare for climate change. To this end, the use of plants surrounding natural CO_2 springs are a precious resource to elucidate adaptation to elevated [CO_2] (Table 1.5).

The study of plant adaptation at natural CO_2 springs requires thorough characterisation and monitoring of the spring and the identification of an appropriate control site (Onoda *et al*., 2007). Previous studies of adaptation in plants at natural CO_2 springs have utilised fast growing herbaceous plants with limited pollen or seed dispersal, or primarily clonal propagation where adaptation is more likely to be found (Onoda *et al*., 2009; Nakamura *et al*., 2011; Van Loon et al, 2016; Watson-Lazowski *et al*., 2016). Where these populations are sufficiently large or sufficiently isolated this may be expected to limit gene flow and therefore promote local adaptation. However these sites have also facilitated the study of the long-lived species that are difficult to work with experimentally (Korner and Miglietta, 1994; Hättenschwiler *et al*., 1997; Rapparini *et al*., 2004), including experiments with their offspring in crossed factored experiments (Tognetti and Johnson, 1999; Polle *et al*., 2001).

Research utilising plants at natural CO_2 spring sites began in the early 1990's and peak funding and interest for using natural CO_2 springs coincides with peak funding for the field generally (Figure 1.8). This has provided a wealth of phenotypic data on a large number of plant species growing around CO_2 springs, but a decline in use of these sites means that only two papers have used new genetic techniques to study these responses at the genetic level (Nakamura *et al*., 2011; Watson-Lazowski *et al*., 2016).

Advantages	Disadvantages
✓ **Able to study multigenerational plant responses in the natural environment**	✗ **Variable gas composition of vent emissions** CO_2 spring emissions may also include SO_2, H_2S, N_2O or even radioactive gases. Composition of emissions has been analysed in well characterised spring sites.
✓ **Low cost associated with the experiments since the carbon dioxide is free**	✗ **Plants are assumed to have been exposed to elevated [CO_2] for multiple generations** This assumption is not always true and there is usually no historical record of the age of CO_2 springs. Researchers may be forced to rely on anecdotal evidence of local inhabitants.
✓ **The presence of a CO_2 concentration gradient from the spring source allows comparison of plants at different CO_2 at the same site (Vodnik *et al*. 2002)**	✗ **Gas emission may be extremely variable over time** This may vary not only during the course of a single day but also over years and months (unlike future predicted CO_2). For example salts may form blocking the soil cracks in vents, and the compaction of the soil can have a large impact on the emission.
✓ **Environmental conditions are natural**	✗ **Other environmental variations between spring and control sites** High CO_2 and H_2S concentrations may make the soil extremely acidic, there may also be large variations in temperature between spring and control sites.
✓ **Sites can be used flexibly for different experiments;** • **Existing vegetation may be used or plants may be introduced** • **Pots can be placed nearby for short term exposure studies** • **Plants may be reciprocally transplanted between spring and a control site** • **Plant material may be taken away for chamber experiments and analysis**	✗ **Analysis of natural vegetation in two sites can be limited by plant age, size and species abundance, reducing replication for statistical power**
✓ **Plant succession experiments can be performed**	✗ **High levels of CO_2 can be a health hazard for researchers working at the site** Symptoms typically begin from 5% atmospheric concentration.
✓ **Many naturally occurring CO_2 vents exist globally**	✗ **The soil around the spring is affected by CO_2 filtering up through the soil, affecting its biogeochemistry.** This is in contrast to the effect of increasing atmospheric CO_2 on the soil via a 'top down' effect.

Table 1.5: The advantages and disadvantages of using natural CO_2 springs to study long term plant adaptation to elevated [CO_2].

Although there is a large number of disadvantages to using natural CO_2 vents as opposed to FACE facilities, they can still be useful as models when limitations are minimised. Problems associated with variability in emissions require that there is frequent data collection on the composition and quantity of CO_2. It is probably for this reason that the majority of research occurs at a handful of well-characterised sites.

The first was a study of a natural population of *Plantago asiatica* at a natural CO_2 spring (Nakamura *et al*., 2011) which showed that genetic differences accompanied phenotypic differences over a [CO_2] gradient, with the implication that [CO_2] acted as the selection pressure. The second study, by Watson-Lazowski *et al*., 2016 utilised RNA-Seq in *P. lanceolata* evidenced that plants under elevated [CO_2] at a natural CO_2 spring site were adapted to this environment at the level of gene expression. They also found very little genetic difference between populations at the spring site and at a nearby control site, raising questions about the potential mechanism of this adaptation. This research provides a foundation for experimental work outlined in this thesis and highlights the value of using these sites in combination with crossed factored experiments for elevated [CO_2] research (Figure 1.9).

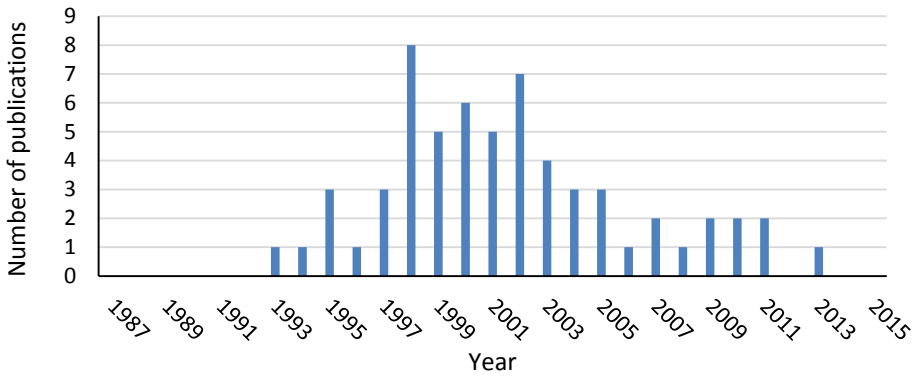
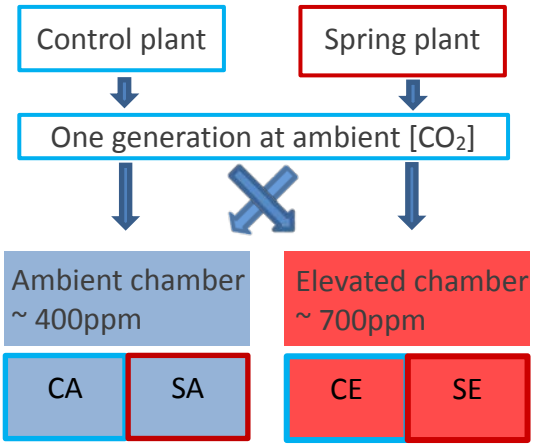


Figure 1.8: The number of publications using natural CO_2 springs to study plant adaptation to elevated [CO_2] by year.

Figure 1.9: Summary of gene expression analysis by Watson-Lazowski *et al*. (2016) in *Plantago lanceolata* at a natural CO_2 spring.

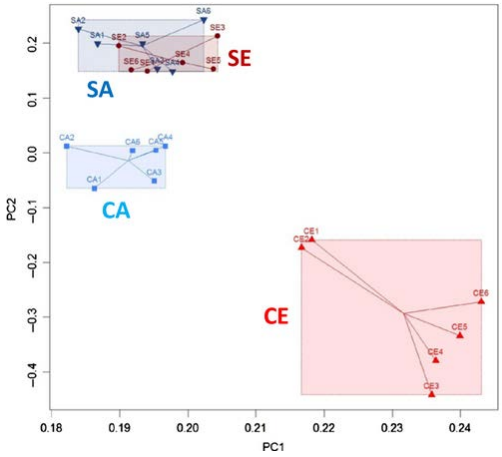
This study used a cross factored experimental design (1.9a), whereby seeds from *P. lanceolata* (1.9b) exposed to naturally elevated (400-1200 ppm) [CO_2], and a nearby ambient [CO_2] control site were grown in growth chambers. Following one generation of growth in the same environment to account for parental effects, the seeds from each population (spring and control) were grown under ambient and elevated [CO_2] conditions.

1.9b) Illustrated drawing of *P. lanceolata*

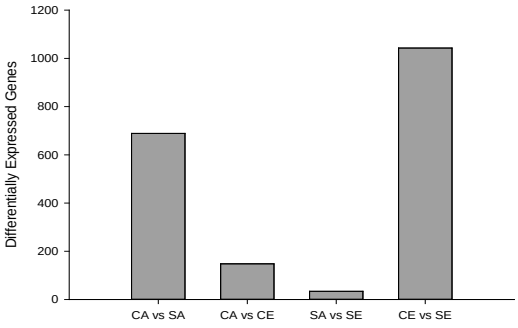


1.9a) Experimental design of crossed factored chamber experiment

From this study, the relatively modest number of differentially expressed genes between spring plants exposed to different growth [CO_2], and likewise for control plants, when compared with spring and control populations at the same growth [CO_2], evidences pre-adaptation to high [CO_2] in the spring plants (1.9c). This is further illustrated by principal component analysis (1.9d).
Population genomic analysis identified essentially zero genetic divergence. This heavily implicates epigenetic mechanisms in coordinating plant responses to elevated [CO_2] and maintaining them through generations.



1.9d) Principal component (PC) analysis for RNA-Seq expression data. PC1 separates ambient and control groups, while PC2 separates spring and control groups.

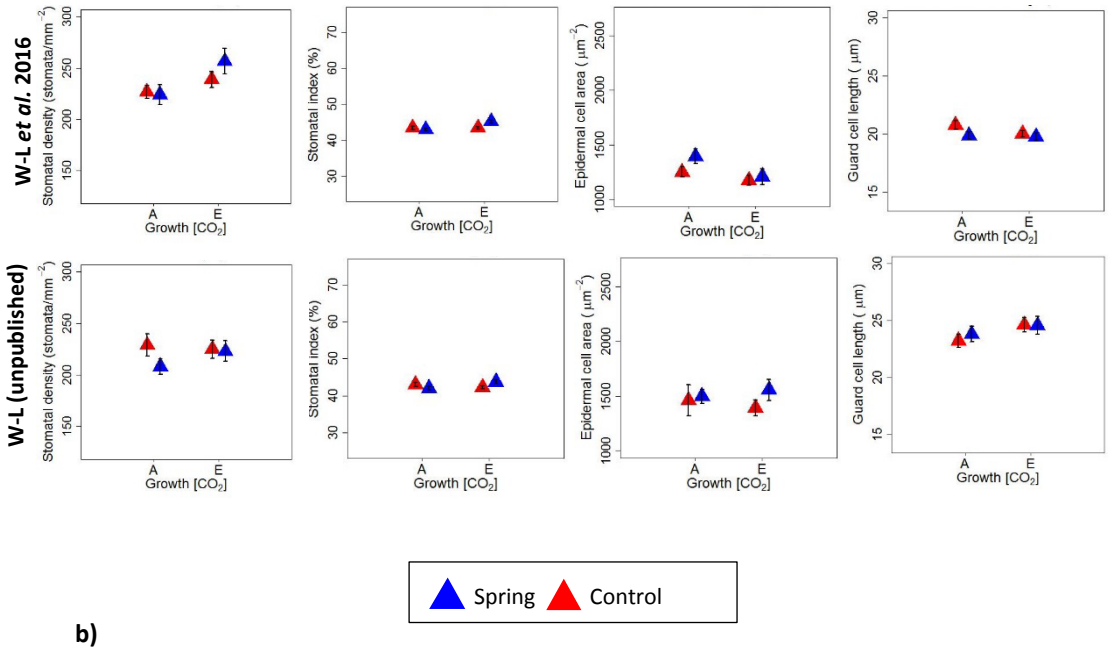


1.9c) Numbers of differentially expressed loci between each treatment group.

1.6.2 Using *Plantago lanceolata* to study multigenerational response to elevated [CO_2]

P. lanceolata growing at the Bossoleto CO_2 spring and established control site is an ideal system to study multigenerational response to elevated [CO_2]. *P. lanceolata* is a basal rosette forming perennial herb, with a gynodioecous breeding system. Although limited to *P. lanceolata* because it occurs in sufficient sample size for experimentation at the elevated [CO_2] spring site, there are also a number of benefits to utilising this species to study multigenerational plant response to elevated [CO_2]. Firstly it is an obligate outcrosser by two mechanisms, gametophytic self-incompatibility and protogynous growth syndrome. This provides assurance that any crosses will reciprocally cross, allowing the study and separation of maternal and paternal effects. Secondly, *P. lanceolata* is wind pollinated over short distances of less than 1 m (Tonsor, 1985; Young and Schmitt, 1995) and produces thousands of seeds. Although these are covered in a sticky mucilage to promote long distance epizoochoric seed dispersal (Kreitschitz *et al*., 2016), the majority of seeds will fall in the area surrounding the maternal plant. This restriction on gene flow may make transgenerational indirect genetic effects more likely to evolve (Galloway, 2005), while obligate outcrossing promotes outbreeding, providing an ideal system to study parental effects. Thirdly, the response of this species to elevated [CO_2] has been independently studied in CO_2 springs, FACE and in controlled environment chambers and is well-documented (Bettarini *et al*., 1998; Norton *et al*., 1999; Hui *et al*., 2002; Watson-Lazowski *et al*., 2016). Fourthly since this species is a cosmopolitan weed with a wide geographic range it is increasingly being studied to answer fundamental ecological questions, for example as part of the global research network PLANTPOPNET (Wardle *et al*., 2014). This may provide more genomic resources and information about population demography in future. Finally, *P. lanceolata* contains economically, agronomically and medicinally important bioactive compounds (Stewart, 1996; Beara *et al*., 2010). The seeds and leaves are a source of mucilage, and sister species *P. ovata* and *P. psyllium* are farmed commercially in India for seed husks which have uses in the pharmaceutical, food and cosmetic industries (Kotwal *et al*., 2016). Since plant response to elevated [CO_2] may involve marked changes to secondary metabolite composition and quantity of plants this will be of key importance to these industries.

a)



b)

Trait	Factor	Watson-Lazowski *et al.*, 2016	Watson-Lazowski (unpublished data, 2013)
Stomatal density (SD)	Location		
	Treatment	‡	
	Location*Treatment		
Stomatal index (SI)	Location	*	
	Treatment		
	Location*Treatment	*	**
Epidermal cell area (ECA)	Location		
	Treatment	‡	
	Location*Treatment		
Guard cell length (GCL)	Location		
	Treatment		*
	Location*Treatment		

Figure 1.10: A summary of stomatal patterning responses in previous work by Alex Watson-Lazowski using *P. lanceolata* in a crossed factored experiment

Seed from a population of plants at a natural CO_2 spring and nearby control site were grown in an experiment at ambient and elevated [CO_2] creating four treatment groups (spring-ambient, spring-elevated, control-ambient, control elevated). 1.10a) Stomatal patterning responses in experiments by Yunan Lin in 2009, published as Watson-Lazowski *et al*., 2016, and by Alex Watson-Lazowski in 2013 (unpublished data), measured across four traits, with red triangles representing plants derived from the spring site and blue triangles representing plants derived from the control site, at ambient (A) or elevated (E) [CO_2]. 1.10b) The significance of effect of location of origin (spring or control) and the effect of treatment (ambient or elevated growth [CO_2]), and their interaction on the trait response.

The responses of *P. lanceolata* to elevated [CO_2] are typical of herbaceous plant responses described in meta-analysis (Ainsworth and Rogers, 2007). In open and closed chambers, glasshouses and at natural CO_2 springs, net photosynthesis in *P. lanceolata* has universally been recorded as stimulated in elevated [CO_2] (Miglietta *et al*., 1998; Clark *et al*., 1999; Staddon *et al*., 1999; Klus *et al*., 2001; He *et al*., 2002). Stomatal conductance decreases under elevated [CO_2] (Bettarini *et al*., 1998; Clark *et al*., 1999; Klus *et al*., 2001; Marchi *et al*., 2004) and this increases water use efficiency (Klus *et al*., 2001, Clark *et al*., 1999). This increased stomatal closure facilitates conservation of water while photosynthesis is still enhanced by increased CO_2 availability. Changes in stomatal density have been recorded as increased (Marchi *et al*., 2004) and decreased (Bettarini *et al*., 1998) on the abaxial surface and decreased on the adaxial surface (Bettarini *et al*., 1998; Marchi *et al*., 2004) of *P. lanceolata* at the natural CO_2 spring. Stomatal phenotypes of *P. lanceolata* derived from natural CO_2 spring populations and utilised in previous experiments at the University of Southampton are presented in Figure 1.10.

1.6.2.1 Transgenerational indirect genetic effects in *Plantago lanceolata*

Transgenerational Indirect genetic effects have been studied in *P. lanceolata* with respect to parental temperature (Alexander and Wulff, 1985; Case *et al*., 1996; Lacey, 1996; Lacey and Herr, 2000), light environment (Schmitt *et al*., 1992; Wulff *et al*., 1994; van Hinsberg and van Tienderen, 1997), herbivory/disturbance (Fajer *et al*., 1991; Latzel *et al*., 2009; Latzel and Klimešová, 2010; Latzel *et al*., 2010) and nutrient provision (Wulff *et al*., 1994; Latzel *et al*., 2009; Latzel and Klimešová, 2010; Latzel *et al*., 2010; Latzel *et al*., 2014). These studies suggest that transgenerational indirect genetic effects in *P. lanceolata* response to the environment can flexibly alter traits such as seed weight (Alexander and Wulff, 1985; Lacey, 1996; Schmitt *et al*., 1992; Latzel *et al*., 2009) and seed germination (Alexander and Wulff, 1985; Lacey, 1996; van Hinsberg and van Tienderen, 1997), as well as offspring growth (Alexander and Wulff, 1985; Wulff *et al*., 1994; Latzel and Klimešová, 2010; Latzel *et al*., 2010) and reproductive traits (Case *et al*., 1996; Latzel *et al*., 2014).

The evolutionary significance of transgenerational indirect genetic effects depends on the ability of the effects to persist into offspring adulthood and further into the subsequent generation, and on the ability of transgenerational indirect genetic effects seen in a controlled environment setting to persist in a more heterogenous natural environment.

Although multiple studies show that parental effects in *P. lanceolata* can persist into adulthood (Alexander and Wulff, 1985; Wulff *et al*., 1994; Latzel *et al*., 2010; Latzel *et al*., 2014), these may not persist to affect survival or reproduction after the first year in the absence of the environmental stimulus (Lacey and Herr, 2000). Over multiple generations however, grandparental temperature was shown to affect leaf area, leaf allometry and percentage of plants flowering among offspring (Case *et al*., 1996).

In parents, elevated [CO_2] increases total biomass (Fajer *et al*., 1991; Staddon *et al*., 1999, Hodge and Millard, 1998; Klus *et al*., 2001; Maestre and Reynolds, 2006; He *et al*., 2002) but reproductive biomass appears to be decreased under elevated [CO_2], suggesting that resources are preferentially allocated to vegetative biomass (Fajer *et al*., 1991). Since this has not been measured over multiple generations or in a natural CO_2 spring, it is not possible to discern if the increased allocation to vegetative biomass is a strategy that increases fitness output in the following year.

Transgenerational indirect genetic effects in *P. lanceolata* response to elevated [CO_2] were studied in two papers more than 25 years ago (Fajer *et al*., 1991; Alexander and Wulff, 1985) (Table 1.3). These studies suggest that elevated [CO_2] led to reduced seed weight and number in the parents but seeds produced under elevated [CO_2] had higher quality, but these studies only focussed on maternal effects.

Parental effect trait	Effect of elevated [CO_2]	Citation
Seed weight	NS ↓(*)	Fajer, Bowers and Bazzazz (1991) Wulff and Alexander (1985)
Seed number	↓(*)	Fajer, Bowers and Bazzazz (1991)
Percentage seeds germinating	NS ↑(**)	Fajer, Bowers and Bazzazz (1991) Wulff and Alexander (1985)
Germination rate index	↑(**)	Wulff and Alexander (1985)
Mean time to germinate	NS	Fajer, Bowers and Bazzazz (1991)
Offspring Leaf area	↑(****)	Wulff and Alexander (1985)
Offspring leaf growth rate	NS	Wulff and Alexander (1985)

Table 1.6: A summary of transgenerational effects documented in two studies in *P. lancelolata* response to elevated [CO_2]

1.7 Summary

This review of multigenerational plant responses to environmental change, focussing on the progress made in understanding how plants will respond to elevated [CO_2] under climate change, has highlighted significant gaps in knowledge that impact our ability to predict plant response to future climate scenarios.

A wealth of published single generation experiments evidence that plants will plastically respond to elevated [CO_2] under climate change (Ainsworth and Long, 2007; Wang *et al*., 2012). However, due to a lack of multigenerational response studies, little is known about whether these plastic responses effectively predict the multigenerational response of plants to elevated [CO_2], or whether direct or transgenerational indirect genetic effects contribute significantly to expressed phenotype. A major challenge to studying the contribution of these mechanisms in plant response to elevated [CO_2] is the time and financial cost associated with multigenerational experiments (Paoletti *et al*., 2005). To this end we propose that plants growing in elevated [CO_2] at natural CO_2 springs provide a resource to quantify these multigenerational responses and explore the contribution of plasticity, adaptation and transgenerational indirect genetic effects through crossed factored experiments combined with HTS technologies. Trends and gaps highlighted by this review include:

1. Single generation experiments have been extensively used to inform and validate predictive models of plant response to elevated [CO_2] under climate but there is no consensus or synthesis of multigenerational plant responses recorded in natural CO_2 springs. Since natural CO_2 springs provide a proxy for the multigenerational or long-term response of plants to elevated [CO_2], and a large number of studies at these sites have been conducted in the last 30 years, a systematic synthesis of these responses would provide the first insights into the magnitude and direction of long-term and multigenerational responses.
2. The general consideration of transgenerational indirect genetic effects as statistical noise in heritability studies has hampered our understanding of how they contribute to adaptation to environmental change (Falconer *et al*., 1996). Nevertheless, increasing evidence suggests that they can directly facilitate the expression of adaptive phenotypes in plant offspring, and as such they could be a

key factor influencing plant response to climate change. Particularly they may contribute to the response to elevated [CO_2] because of the predictable nature of the change. Although elevated [CO_2] is known to effect plant phenology and seed provisioning via maternal effects, there is a lack of research into the potential role of parental and grandparental effects. Statistical models of phenotypic variation between plants in multigenerational experiments can partition the contribution of transgenerational indirect genetic effects to plant adaptation at natural CO_2 springs. The application of HTS technologies to these experiments could further elucidate the mechanistic basis of these responses, but have not readily been applied because of a lack of genomic resources for the non-model species that grow at natural CO_2 springs.

3. Methylation could play a role in coordinating plastic response to elevated [CO_2], and epigenetic inheritance of methylation changes under elevated [CO_2] could facilitate adaptation. This has not yet been explored despite evidence that methylation could be involved in coordinating responses to other abiotic environmental changes. Whether or not environmentally induced chromatin modifications can be adaptive through epigenetic inheritance remains controversial.

1.8 Aims and objectives

This work aims to address knowledge gaps about how plants respond to elevated [CO_2] over multiple generations that were identified from a review of the literature.
P. lanceolata from the Bosseleto natural CO_2 spring site in Italy are used to model plant responses to multigenerational elevated [CO_2] exposure in the experimental work of chapters three and four.

Objectives

1. **Quantify general trends in multigenerational plant response to elevated [CO_2] at natural spring sites.**
   Chapter 2; Summarise trends in plant [CO_2] response at spring sites through a systematic search of the literature and meta-analysis of all traits that are sufficiently reported. Compare these to existing FACE meta-analyses.

2. **Approximate the contribution of transgenerational indirect genetic effects and heritability to multigenerational plant response to elevated [CO_2].**
   Chapter 3; Test theoretical predictions for the evolution of transgenerational indirect genetic effects under elevated [CO_2] response by conducting a multigenerational, crossed factored experiment with *P. lanceolata* from the Bossoleto CO_2 spring. Calculate the heritability of growth and stomatal traits and the influence of parental and grandparental effects using an 'animal model'.

3. **Explore the role of sequence divergence, epigenetic inheritance and plasticity in multigenerational plant response to elevated [CO_2] by producing and interrogating a complete dataset of gene expression, phenotype, methylation pattern and sequence analysis in a crossed factored experiment.**
   Chapter 4; Sequence and assemble a reference genome for *P. lanceolata*. Sequence, assemble and analyse the methylome, genome and transcriptome of 24 plants grown in a crossed factored experiment, in order to complete the transcriptome and phenotype datasets previously acquired from an experiment by Alex Watson-Lazowski and Yunan Lin.

1.9 Statement regarding Candidate Contributions

All three data chapters (two, three and four) were prepared as manuscripts for publication. As such, all three chapters are listed with multiple authors who contributed to the work described, but the candidate was the lead author in each instance and drafted and revised all three manuscripts. For all chapters, the candidate received guidance and support from supervisors Dr Mark Chapman and Professor Gail Taylor, who are listed as co-authors.

For Chapter two the candidate undertook all of the data collection and analysis.

In the multigenerational experiment described in chapter three the candidate received seed collected by Professor Taylor from the Bossoleto spring site for the experiment. The candidate undertook all of the experimental work and analysis, and received guidance on statistics from co-author Dr Thomas Ezard and on experimental design from Dr Ezard, Dr Chapman and Professor Taylor.

For chapter four, the candidate utilised results of a crossed factored chamber experiment carried out by Dr Yunan Lin in 2009. Dr Lin grew and phenotyped all of the plants in this experiment. Dr Alex Watson-Lazowski extracted all of the RNA for sequencing and the results of his analysis of gene expression were published in Watson-Lazowski et al, 2016. The candidate re-analysed the raw RNA sequencing data gathered in this experiment, utilising increased computational resources and further re-visualised gene expression data. The candidate undertook all lab work associated with extracting and cleaning DNA for sequencing (genome assembly, whole genome and whole genome bisulfite sequencing). The candidate *de novo* assembled and preliminarily annotated the genome for *P. lanceolata*, and completed all analysis of the whole genome and whole genome bisulfite sequencing datasets from quality control to differential methylation analysis. The candidate received some initial instruction in methylation sequencing analysis from Dr Steve Eichten of the Australian National University during a three month training placement. Funding for sequencing outlined in chapter four was provided as part of a start-up fund assigned to Professor Taylor on her commencement of the position of Chair of the Department of Plant Sciences at University of California, Davis.

1.10 References

Ainsworth, E. A., Rogers, A., Nelson, R., & Long, S. P. (2004) Testing the "source–sink" hypothesis of down-regulation of photosynthesis in elevated [CO_2] in the field with single gene substitutions in *Glycine max*. *Agric. Forest Meteorol.,* **122**, 85–94.

Ainsworth, E. A., & Long, S. P. (2005) What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytologist*, **165**, 351-372.

Ainsworth, E. A., Rogers, A., Vodkin, L. O., Walter, A., & Schurr, U. (2007) The effects of elevated CO_2 concentration on soybean gene expression. An analysis of growing and mature leaves. *Plant Physiology,* **142,** 135-147.

Ainsworth, E. A. (2016) The importance of intraspecific variation in tree responses to elevated [CO_2]: breeding and management of future forests. *Tree Physiology*, **36**, 679-681.

Alexander, H. M., & Wulff, R. D. (1985) Experimental Ecological Genetics in *Plantago*: The Effects of Maternal Temperature on Seed and Seedling Characters in *Plantago lanceolata*. *Journal of Ecology,* **73,** 271-282.

Andalo, C., Godelle, B., Lefranc, M., Mousseau, M., & Tillbottraud, I. (1996) Elevated CO_2 decreases seed germination in *Arabidopsis thaliana*. *Global Change Biology,* **2,** 129-135.

Arndal, M. F., Toyler, A., Larsen, K. S., Beier, C., & Schmidt, I. K. (2018) Fine Root Growth and Vertical Distribution in Response to Elevated CO_2, Warming and Drought in a Mixed Heathland–Grassland. *Ecosystems*, **21**, 15-30.

Auge, G. A., Leverett, L. D., Edwards, B. R., & Donohue, K. (2017) Adjusting phenotypes via within-and across-generational plasticity. *New Phytologist*, **216**, 343-349.

Baker, M. 2012. *De novo* genome assembly: what every biologist should know. *Nature Methods,* **9**, 333-337.

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.

Barnes, J. D., Bettarini, I., Polle, A., Slee, N., Raines, C., Miglietta, F., Raschi, A., & Fordham, M. (1997) The impact of elevated CO_2 on growth and photosynthesis in *Agrostis canina L. ssp. monteluccii* adapted to contrasting atmospheric CO_2 concentrations. *Oecologia*, **110**, 169-178.

Barnaby, J. Y., & Ziska, L. H. (2012) Plant Responses to Elevated CO_2. *eLS*, doi:10.1002/9780470015902.a0023718.

Barnola, J. M., Raynaud, D., Lorius, C. & Barkov, N. I. (2003) Historical CO_2 record from the Vostok ice core. In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Baulcombe, D. C., & Dean, C. (2014) Epigenetic regulation in plant responses to the environment. *Cold Spring Harbor perspectives in biology,* **6,** a019471.

Bazzaz, F. A., Jasienski, M., Thomas, S. C., & Wayne, P. (1995) Microevolutionary Responses in Experimental Populations of Plants to CO_2 Enriched Environments - Parallel Results from 2 Model Systems. *Proceedings of the National Academy of Sciences of the United States of America,* **92,** 8161-8165.

Beara, I. N., Orcic, D. Z., Lesjak, M. M., Mimica-Dukic, N. M., Pekovic, B. A., & Popovic, M. R. (2010) Liquid chromatography/tandem mass spectrometry study of anti-inflammatory activity of Plantain (*Plantago* L.) species. *Journal of Pharmaceutical and Biomedical Analysis,* **52,** 701-706.

Beaumont M.A & Rannala B. (2004) The Bayesian revolution in genetics. *Nature Reviews Genetics,* **5**, 251–261.

Becker, C., Hagmann, J., Muller, J., Koenig, D., Stegle, O., Borgwardt, K., & Weigel, D. (2011) Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature,* **480,** 245-127.

Becklin, K. M., Walker, S. M., Way, D. A., & Ward, J. K. (2017) CO_2 studies remain key to understanding a future world. *New Phytologist*, **214**, 34-40.

Bernacchi, C. J., Calfapietra, C., Davey, P. A., Wittig, V. E., Scarascia-Mugnozza, G. E., Raines, C. A., & Long, S. P. (2003) Photosynthesis and stomatal conductance responses of

poplars to free-air CO_2 enrichment (PopFACE) during the first growth cycle and immediately following coppice. *New Phytologist*, **159**, 609-621.

Bettarini, I., Vaccari, F. P., & Miglietta, F. (1998) Elevated CO_2 concentrations and stomatal density: observations from 17 plant species growing in a CO_2 spring in central Italy. *Global Change Biology,* **4,** 17-22.

Bilichak, A., & Kovalchuk, I. (2016) Transgenerational response to stress in plants and its application for breeding. *Journal of Experimental Botany,* **67,** 2081-2092.

Blackman, C. J., Aspinwall, M. J., Resco De Dios, V., Smith, R. A., & Tissue, D. T. (2016) Leaf photosynthetic, economics and hydraulic traits are decoupled among genotypes of a widespread species of eucalypt grown under ambient and elevated CO_2. *Functional Ecology*, **30**, 1491-1500.

Bloom, A. J., Asensio, J. S. R., Randall, L., Rachmilevitch, S., Cousins, A. B., & Carlisle, E. A. (2012) CO_2 enrichment inhibits shoot nitrate assimilation in C_3 but not C_4 plants and slows growth under nitrate in C_3 plants. *Ecology*, **93**, 355-367.

Bond, W. J., & Midgley, G. F. (2012) Carbon dioxide and the uneasy interactions of trees and savannah grasses. *Philosophical Transactions of the Royal Society B-Biological Sciences,* **367,** 601-612.

Boyko, A., & Kovalchuk, I. (2008) Epigenetic control of plant stress response. *Environmental and Molecular Mutagenesis,* **49,** 61-72.

Brooks, S.P., & Gelman, A. (1998)General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, **7**, 434-455.

Case, A. L., Lacey, E. P., & Hopkins, R. G. (1996) Parental effects in *Plantago lanceolata* L . II Manipulation of grandparental temperature and parental flowering time. *Heredity,* **76,** 287-295.

Chen, C., Jiang, Q., Ziska, L. H., Zhu, J., Liu, G., Zhang, J., Ni, K., Seneweera, S., & Zhu, C. (2015) Seed vigor of contrasting rice cultivars in response to elevated carbon dioxide. *Field Crops Research,* 178**,** 63-68.

Compeau, P. E., Pevzner, P. A., & Teslar, G. (2011) How to apply de Bruijn graphs to genome assembly. *Nature biotechnology*, **29**, 987.

Clark, H., Newton, P., & Barker, D. (1999) Physiological and morphological responses to elevated CO_2 and a soil moisture deficit of temperate pasture species growing in an established plant community. *Journal of Experimental Botany,* **50,** 233-242.

Cotrufo, M. F., Ineson, P., & Scott, A. (1998) Elevated CO_2 reduces the nitrogen concentration of plant tissues. *Global Change Biology,* **4,** 43-54.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M. W., Gaffney, D.J., Elo, L. L., Zhang, X., & Mortazavi, A. (2016) A survey of best practices for RNA-seq data analysis. *Genome biology*, **17**, 13.

Darbah, J. N. T., Kubiske, M. E., Nelson, N., Oksanen, E., Vaapavuori, E., & Karnosky, D. F. (2007) Impacts of elevated atmospheric CO_2 and O_3 on paper birch (Betula papyrifera): Reproductive fitness. *The Scientific World Journal,* **7,** 240-246.

Das, O., & Messing, J. (1994) Variegated phenotype and developmental methylation changes of a maize allele originating from epimutation. *Genetics,* **136,** 1121-1141.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499.

Deans, C., & Maggert, K. A. (2015) What do you mean, “epigenetic”? *Genetics*, **199**, 887-896.

Debiasse, M. B., & Kelly, M. W. (2016) Plastic and evolved responses to global change: what can we learn from comparative transcriptomics? *Journal of Heredity*, ***107***, 71-81.

De Costa, W. A. J. M., Weerakoon, W. M. W., Chinthaka, K. G. R., Herath, H. M. L. K., & Abeywardena, R. M. I. (2007) Genotypic variation in the response of rice (*Oryza sativa L.*) to increased atmospheric carbon dioxide and its physiological basis. *Journal of Agronomy and Crop Science,* **193,** 117-130.

Dempewolf, H., Eastwood, R. J., Guarino, L., khoury, C. K., Muller, J. V., & Toll, J. (2014) Adapting agriculture to climate change: a global initiative to collect, conserve, and use crop wild relatives. *Agroecology and Sustainable Food Systems*, **38**, 369-377.

De Souza, A. P., Arenque, B. C., Tavares, E. Q., & Buckeridge, M. S. (2016) Transcriptomics and Genetics Associated with Plant Responses to Elevated CO_2 Atmospheric Concentrations. In D. Edwards & J. Batley (Eds.) *Plant Genomics and Climate Change* (pp. 67-83). Springer, New York, NY.

Dhami, N., Tissue, D. T., & Cazzonelli, C. I. (2018) Leaf-age dependent response of carotenoid accumulation to elevated CO_2 in Arabidopsis. *Archives of biochemistry and biophysics*, ***647***, 67-75.

Dippery, J. K., Tissue, D. T., Thomas, R. B., & Strain, B. R. (1995) Effects of Low and Elevated CO_2 on C_3 and C_4 Annuals. *Oecologia,* **101,** 13-20.

Dorazio, R. M. (2016) Bayesian data analysis in population ecology: motivations, methods, and benefits. *Population ecology*, ***58***, 31-44.

Eichten, S. R., Schmitz, R. J., & Springer, N. M. (2014) Epigenetics: Beyond Chromatin Modifications and Complex Genetic Regulation. *Plant Physiology,* **165,** 933-947.

Ekman, Å., Bülow, L., & Stymne, S. (2007) Elevated atmospheric CO_2 concentration and diurnal cycle induce changes in lipid composition in *Arabidopsis thaliana*. *New Phytologist,* **174,** 591-599.

Ellegren, H. (2014) Genome sequencing and population genomics in non-model organisms. *Trends in ecology & evolution*, **29**, 51-63.

Ellison, A. M. (2004) Bayesian inference in ecology. *Ecology letters*, **7**, 509-520.

Engineer, C. B., Hashimoto-Sugimoto, M., Negi, J., Israelsson-Nordstrom, M., Azoulay-Shemer, T., Rappel, W. J., Iba, K., & Schroeder, J. I. (2016) CO_2 sensing and CO_2 regulation of stomatal conductance: advances and open questions. *Trends in plant science*, **21**, 16-30.

English, S., Pen, I., Shea, N., Uller, T. (2015) The information value of non-genetic inheritance in plants and animals. *PLoS ONE*, **10***,* e0116996.

Ezard, T. H. G., Prizak, R., Hoyle, R. B., Marshall, D. (2014) The fitness costs of adaptation via phenotypic plasticity and maternal effects. *Functional Ecology*, **28**, 693–701.

Fajer, E. D., Bowers, M. D., & Bazzaz, F. A. (1991) Performance and Allocation Patterns of the Perennial Herb, *Plantago lanceolata*, in Response to Simulated Herbivory and Elevated CO_2 Environments. *Oecologia,* **87,** 37-42.

Falconer, D. S., Mackay, T. F., & Frankham, R. (1996) Introduction to quantitative genetics. *Trends in Genetics*, **12**, 280.

Fan, G., Cai, Q., Wang, C., Wan, J., & Zhu, J. (2005) QTL for yield and its components responded to elevated CO_2 in rice (*Oryza sativa* L.). *Acta genetica Sinica,* **32,** 1066-1073.

Farnsworth, E. J., & Bazzaz, F. A. (1995) Inter-Generic and Intra-Generic Differences in Growth, Reproduction, and Fitness of 9 Herbaceous Annual Species Grown in Elevated CO_2 Environments. *Oecologia,* **104,** 454-466.

Ferris, R., Long, L., Bunn, S. M., Robinson, K. M., Bradshaw, H. D., Rae, A. M., & Taylor, G. (2002) Leaf stomatal and epidermal cell development: identification of putative quantitative trait loci in relation to elevated carbon dioxide concentration in poplar. *Tree Physiology,* **22,** 633-640.

Forsman, A. (2015) Rethinking phenotypic plasticity and its consequences for individuals, populations and species. *Heredity*, **115**, 276.

Franks, P. J., Adams, M. A., Amthor, J. S., Barbour, M. M., Berry, J. A., Ellsworth, D. S., Farquhar, G.D., Ghannoum, O., Lloyd, J., McDowell, N., & Norby, R. J. (2013) Sensitivity of plants to changing atmospheric CO_2 concentration: from the geological past to the next century. *New Phytologist*, **197**, 1077-1094.

Frenck, G., Linden, L., Mikkelsen, T. N., Brix, H., & Jorgensen, R. B. (2013) Response to multi-generational selection under elevated [CO_2] in two temperature regimes suggests enhanced carbon assimilation and increased reproductive output in *Brassica napus* L. *Ecology and evolution*, **3**, 1163-1172.

Frick, J., Nielsen, S. S., & Mitchell, C. A. (1994) Yield and Seed Oil Content Response of Dwarf, Rapid-Cycling Brassica to Nitrogen Treatments, Planting Density, and carbon dioxide Enrichment. *Journal of the American Society for Horticultural Science,* **119,** 1137-1143.

Friedlingstein, P., Andrew, R. M., Rogelj, J., Peters, G. P., Canadell, J. G., Knutti, R., Luderer, G., Raupach, M. R., SChaeffer, M., Van Vuuren, D. P., & Le Quere, C. (2014) Persistent growth of CO_2 emissions and implications for reaching climate targets. *Nature Geoscience,* **7,** 709-715.

Furrow, R. E. (2014) Epigenetic Inheritance, Epimutation, and the Response to Selection. *Plos One,* 9, e101559.

Fusco, G., & Minelli, A. (2010) Phenotypic plasticity in development and evolution: facts and concepts. *Phil. Trans. R. Soc. B*, **365**, 547-556.

Futuyma, D. J. (2009) Natural selection and adaptation. In D. J. Futuyma (Ed.) *Evolution* (pp 279-301). Sinauer Associates, Massachusetts, USA.

Galloway, L. F. (2005) Maternal effects provide phenotypic adaptation to local environmental conditions. *New Phytologist,* **166,** 93-99.

Galloway, L. F., & Etterson, J. R. (2007) Transgenerational plasticity is adaptive in the wild. *Science,* **318,** 1134-1136.

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. *Plant, Cell & Environment*, **41**, 1233-1246.

Ge, Y., Guo, B., Cai, Y., Zhang, H., Luo, S., & Gallois, P. (2018) Transcriptome analysis identifies differentially expressed genes in maize leaf tissues in response to elevated atmospheric [CO_2]. *Journal of Plant Interactions*, **13**, 373-379.

Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology,* **21,** 394-407.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. (Eds.) (1995) *Markov chain Monte Carlo in practice* (pp. 5-8). Chapman & Hall/CRC press, London, UK.

Gill, J. (2014) Bayesian methods: A social and behavioral sciences approach (Vol. 20). CRC press, Florida, USA.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., AdiconiS, X., Fan, L., Raychowdhury, R., & Zeng, Q. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology,* **29,** 644-652.

Gratani, L. (2014) Plant Phenotypic Plasticity in Response to Environmental Factors. *Advances in* Botany, 2014, http://dx.doi.org/10.1155/2014/208747.

Gray, J. E., Holroyd, G. H., Van Der Lee, F. M., Bahrami, A. R., Sijmons, P. C., Woodward, F. I., Schuch, W., & Heterington, A. M. (2000) The HIC signalling pathway links CO_2 perception to stomatal development. *Nature,* **408,** 713-716.

Grossman, J. D., & Rice, K. J. (2014) Contemporary evolution of an invasive grass in response to elevated atmospheric CO_2 at a Mojave Desert FACE site. *Ecology letters*, **17**, 710-716.

Grünzweig, J. M., & Dumbur, R. (2012) Seed traits, seed-reserve utilization and offspring performance across pre-industrial to future CO_2 concentrations in a Mediterranean community. *Oikos,* **121,** 579-588.

Gupta, P., Duplessis, S., White, H., karnosky, D. F., Martin, F., & Podila, G. K. (2005) Gene expression patterns of trembling aspen trees following long-term exposure to interacting elevated CO_2 and troposphere O_3. *New Phytologist,* **167,** 633-633.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., & Lieber, M. (2013) *De novo* transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. *Nature protocols,* **8**, 1494-1512.

Hadfield, J.D., (2010) MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software,* **33**, 1-22.

Hättenschwiler, S., Miglietta, F., Raschi, A., & Körner, C. (1997) Thirty years of *in situ* tree growth under elevated CO_2: a model for future forest responses? *Global Change Biology*, **3**, 463-471.

He, J. S., Bazzaz, F. A., & Schmid, B. (2002) Interactive effects of diversity, nutrients and elevated CO_2 on experimental plant communities. *Oikos,* **97,** 337-348.

Hendry, A. P. (2015) Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. *Journal of Heredity*, **107**, 25-41.

Herman, J. J., & Sultan, S. E. (2011) Adaptive transgenerational plasticity in plants: case studies, mechanisms, and implications for natural populations. *Frontiers in Plant Science,* **2**, doi: 10.3389/fpls.2011.00102.

Herman, J. J., & Sultan, S. E. (2016) DNA methylation mediates genetic variation for adaptive transgenerational plasticity. *Proc. R. Soc. B*, **283**, 20160988.

Herman, J. J., Sultan, S. E., Horgan-Kobelski, T., & Riggs, C. (2012) Adaptive transgenerational plasticity in an annual plant: grandparental and parental drought stress enhance performance of seedlings in dry soil. *Integrative and Comparative Biology,* **52**, 77-88.

Heard, E., & Martienssen, R. A. (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell,* **157,** 95-109.

Hill, W. G. (2010) Understanding and using quantitative genetic variation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 73-85.

Hodge, A., & Millard, P. (1998) Effect of elevated CO_2 on carbon partitioning and exudate release from *Plantago lanceolata* seedlings. *Physiologia Plantarum,* **103,** 280-286.

Hoyle, R. B., & Ezard, T. H. G. (2012) The benefits of maternal effects in novel and in stable environments. *Journal of the Royal Society Interface,* **9**, 2403–2413.

Hui, D. F., Sims, D. A., Johnson, D. W., Cheng, W. X., & Luo, Y. Q. (2002) Effects of gradual versus step increases in carbon dioxide on *Plantago* photosynthesis and growth in a microcosm study. *Environmental and Experimental Botany,* **47,** 51-66.

Hussain, M., Kubiske, M. E., & Connor, K. F. (2001) Germination of CO_2-enriched *Pinus taeda* L. seeds and subsequent seedling growth responses to CO_2 enrichment. *Functional Ecology,* **15,** 344-350.

IPCC (2014) Summary for Policymakers. In: C. B. Field *et al.,* (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. Cambridge, United Kingdom, and New York, NY, USA: Cambridge University Press. pg 29.

Jablonski, L. M., Wang, X. Z., & Curtis, P. S. (2002) Plant reproduction under elevated CO_2 conditions: a meta-analysis of reports on 79 crop and wild species. *New Phytologist,* **156,** 9-26.

Jones, A. G., Scullion, J., Ostle, N., Levy, P. E., & Gwynn-Jones, D. (2014) Completing the FACE of elevated CO_2 research. *Environment International*, **73**, 252-258.

Jump, A. S., Penuelas, J. (2005) Running to stand still: adaptation and the response of plants to rapid climate change. *Ecol. Lett.* **8**, 1010–1020.

Klus, D. J., Kalisz, S., Curtis, P. S., Teeri, J. A., & Tonsor, S. J. (2001) Family- and population-level responses to atmospheric CO_2 concentration: Gas exchange and the allocation of C, N, and biomass in *Plantago lanceolata* (Plantaginaceae). *American Journal of Botany,* **88,** 1080-1087.

Kotwal, S., Kaul, S., Sharma, P., Gupta, M., Shankar, R., Jain, M., & Dhar, M. K. (2016) *De Novo* Transcriptome Analysis of Medicinally Important *Plantago ovata* Using RNA-Seq. *Plos One,* **11**, e0150273.

Kreitschitz, A., Kovalev, A., & Gorb, S. N. (2016) "Sticky invasion" - the physical properties of *Plantago lanceolata* L. seed mucilage. *Beilstein Journal of Nanotechnology,* **7,** 1918-1927.

Korner, C. (2006) Plant CO_2 responses: an issue of definition, time and resource supply. *New Phytologist,* **172,** 393-411.

Korner, C., & Miglietta, F. (1994) Long term effects of naturally elevated CO_2 on Mediterranean grassland and forest trees. *Oecologia*, **99**, 343-351.

Kruuk, L. E., Slate, J., & Wilson, A. J. (2008) New answers for old questions: The evolutionary quantitative genetics of wild animal populations. Annual Review of Ecology, Evolution, and Systematics, ***39***, 525–548.

Kuijper, B., & Hoyle, R. B. (2015) When to rely on maternal effects and when on phenotypic plasticity? *Evolution*, **69**, 950–968.

Kumar, S., Kumari, R., Sharma, V., & Sharma, V. (2013) Roles, and establishment, maintenance and erasing of the epigenetic cytosine methylation marks in plants. *Journal of genetics,* **92,** 629-666.

Kumar, S., Sreeharsha, R. V., Mudalkar, S., Sarashetti, P. M., & Reddy, A. R. (2017) Molecular insights into photosynthesis and carbohydrate metabolism in *Jatropha curcas* grown under elevated CO_2 using transcriptome sequencing and assembly. *Scientific reports*, **7**, e11066.

Lacey, E. P. (1996) Parental effects in *Plantago lanceolata* L. I: a growth chamber experiment to examine pre-and postzygotic temperature effects. *Evolution,* **50,** 865-878.

Lacey, E. P., & Herr, D. (2000) Parental effects in *Plantago lanceolata* L. III. Measuring parental temperature effects in the field. *Evolution,* **54,** 1207-1217.

Lande, R. (2009) Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *Journal of Evolutionary Biology,* **22,** 1435-1446.

Langmead, B., & Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods*, ***9***, 357.

Latzel, V. (2015) Pitfalls in ecological research–transgenerational effects. *Folia Geobotanica,* **50,** 75-85.

Latzel, V., Hájek, T., Klimešová, J., & Gómez, S. (2009) Nutrients and disturbance history in two *Plantago* species: maternal effects as a clue for observed dichotomy between resprouting and seeding strategies. *Oikos,* **118,** 1669-1678.

Latzel, V., Janeček, Š., Doležal, J., Klimešová, J., & Bossdorf, O. (2014) Adaptive transgenerational plasticity in the perennial *Plantago lanceolata*. *Oikos,* **123,** 41-46.

Latzel, V., & Klimešová, J. (2010) Year-to-year changes in expression of maternal effects in perennial plants. *Basic and Applied Ecology,* **11,** 702-708.

Latzel, V., Klimešová, J., Hájek, T., Gómez, S., & Šmilauer, P. (2010) Maternal effects alter progeny's response to disturbance and nutrients in two *Plantago* species. *Oikos,* **119,** 1700-1710.

Lau, J. A., Shaw, R. G., Reich, P. B., Shaw, F. H., & Tiffin, P. (2007) Strong ecological but weak evolutionary effects of elevated CO_2 on a recombinant inbred population of *Arabidopsis thaliana*. *New Phytologist,* **175,** 351-362.

Law, J. A., & Jacobsen, S. E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature reviews. Genetics,* **11,** 204.

Leakey, A. D., Ainsworth, E. A., Bernacchi, C. J., Roger, A., Long, S. P., & Ort, D. R. (2009) Elevated CO_2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. *Journal of experimental botany*, **60**, 2859-2876.

Leakey, A. D., & Lau, J. A. (2012) Evolutionary context for understanding and manipulating plant responses to past, present and future atmospheric [CO_2]. *Phil. Trans. R. Soc. B*, **367**, 613-629.

Lehmann, P., & OR, D. (2015) Effects of stomata clustering on leaf gas exchange. New Phytologist, **207**, 1015-1025.

Leimar, O., McNamara, J. M. (2015) The evolution of transgenerational integration of information in heterogeneous environments. American Naturalist, 185, E55–E69.

Li, H., & Durbin, R. (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.

Li, L., Zhang, Y. B., Luo, J. X., Korpelainen, H., & Li, C. Y. (2013) Sex-specific responses of *Populus yunnanensis* exposed to elevated CO_2 and salinity. *Physiologia Plantarum,* **147,** 477-488.

Li, R., Fan, W., Tian, G., Zhu, H., He, L., Cai, J., Huang, Q., Cai, Q., Li, B., Bai, Y., & Zhang, Z. (2010) The sequence and *de novo* assembly of the giant panda genome. *Nature*, **463**, 311.

Lindroth, R. L., Roth, S., & Nordheim, E. V. (2001) Genotypic variation in response of quaking aspen (*Populus tremuloides*) to atmospheric CO_2 enrichment. *Oecologia,* **126,** 371-379.

Little, R. J. A., & Rubin, D.B. (2002) Statistical analysis with missing data, 2nd edn. John Wiley and Sons, New York, USA.

Lin, Y. B., Shen, C. G., Lin, E. D., Hao, X. Y., & Han, X. (2016) Transcriptome response of wheat Norin 10 to long-term elevated CO_2 under high yield field condition. *Journal of integrative agriculture*, **9**, 2142-2152.

Liu, Y., Zhou, J., & White, K. P. (2013) RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics,* **30,** 301-304.

Long, S. P., Ainsworth, E. A., Leakey, A. D. B., Nosberger, J., & Ort, D. R. (2006) Food for thought: Lower-than-expected crop yield stimulation with rising CO_2 concentrations. *Science*, **312**, 1918-1921.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., & Tang, J. (2012) SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience*, **1**, 18.

Luthi, D., Le Floch, M., Bereiter, B., Blunier, T., Barnola, J. M., Siegenthaler, U., Raynaud, D., Jouzel, J., Fischer, H., Kawamura, K., & Stocker, T. F. (2008) High-resolution carbon dioxide concentration record 650,000-800,000 years before present. *Nature*, **453**, 379-382.

MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., Van Ommen, T., Smith, A., & Elkins, J. (2006) Law Dome CO_2, CH_4 and N_2O ice core records extended to 2000 years BP. *Geophysical Research Letters*, **33**, L14810 .

Maestre, F. T., & Reynolds, J. F. (2006) Nutrient availability and atmospheric CO_2 partial pressure modulate the effects of nutrient heterogeneity on the size structure of populations in grassland species. *Annals of Botany*, **98**, 227-235

Marchi, S., Tognetti, R., Vaccari, F. P., Lanini, M., Kaligaric, M., Miglietta, F., & Raschi, A. (2004) Physiological and morphological responses of grassland species to elevated atmospheric CO_2 concentrations in FACE-systems and natural CO_2 springs. *Functional Plant Biology*, **31**, 181-194.

Marty, C., & Bassirirad, H. (2014) Seed germination and rising atmospheric CO_2 concentration: a meta-analysis of parental and direct effects. *New phytologist*, **202**, 401-414.

Matz, M. V. (2017) Fantastic Beasts and How To Sequence Them: Ecological Genomics for Obscure Model Organisms. *Trends in Genetic,* **34**, 121-132.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & Depristo, M. A. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, **20**, 1297-1303.

McPeek, T. M., & Wang, X. (2007) Reproduction of dandelion (*Taraxacum officinale*) in a higher CO_2 environment. *Weed science*, **55**, 334-340.

Mehrotra, S., & Goyal, V. (2014) Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. *Genomics Proteomics Bioinformatics*, **12**, 164-171.

Metz, J., Von Oppen, J., & Tielborger, K. (2015) Parental environmental effects due to contrasting watering adapt competitive ability, but not drought tolerance, in offspring of a semi-arid annual *Brassicaceae*. *Journal of Ecology*, **103**, 990-997.

Meyer, K. (2007) WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *Journal of Zhejiang University Science B*. **8**, 815-821.

Miglietta, F., Bettarini, I., Raschi, A., Korner, C., & Vaccari, F. P. (1998) Isotope discrimination and photosynthesis of vegetation growing in the Bossoleto CO_2 spring. *Chemosphere*, **36**, 771-776.

Monnin, E., Indermuhle, A., Dallenbach, A., Fluckiger, J., Stauffer, B., Stocker, T. F., Raynaud, D. & Barnola, J. M. (2001) Atmospheric CO2 concentrations over the last glacial termination. *Science*, **291**, 112-114.

Moran, E. V., Hartig, F., & Bell, D. M. (2016) Intraspecific trait variation across scales: implications for understanding global change responses. *Global Change Biology*, **22**, 137-150.

Mousseau, T. A., & Fox, C. W. (1998) The adaptive significance of maternal effects. *Trends in Ecology & Evolution*, **13**, 403-407.

Murren, C. J., Auld, J. R., Callahan, H., Ghalambor, C. K., Handelsman, C. A., Heskel, M. A., Kingsolver, J. G., Maclean, H. J., Masel, J., Maughan, H., & Pfennig, D. W. (2015) Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. *Heredity*, **115**, 293.

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., Carlisle, E., Dietterich, L. H., Fitzgerald, G., & Hasegawa, T. (2014) Increasing CO_2 threatens human nutrition. *Nature*, **510**, 139-142.

Nair, S. S., Luu, P. L., Qu, W., Maddugoda, M., Huschtscha, L., Reddel, R., Chenevix-Trench, G., Toso, M., Kench, J.G., Horvath, L.G., & Hayes, V. M. (2018) Guidelines for whole genome bisulphite sequencing of intact and FFPET DNA on the Illumina HiSeq X Ten. *Epigenetics & chromatin*, **11**, https://doi.org/10.1186/s13072-018-0194-0.

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 CO_2 concentration. *Oecologia*, **165**, 809-818.

Narzisi, G., & Mishra, B. (2011) Comparing *de novo* genome assembly: the long and short of it. *PloS one*, **6**, e19175.

Neftel, A., H. Friedli, E. Moor, H. Lötscher, H. Oeschger, U. Siegenthaler & B. Stauffer. (1994) Historical CO_2 record from the Siple Station ice core. *In* Trends: A compendium of data on global change. *Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy,* Oak Ridge, Tennessee, USA.

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius, U., Poot, P., Purugganan, M.D., Richards, C.L., Valladares, F., & Van Kleunen, M. (2010) Plant phenotypic plasticity in a changing climate. *Trends in plant science*, **15**, 684-692.

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**, 149-156.

Noble, D. W., McFarlane, S. E., Keogh, J. S., & Whiting, M. J. (2014) Maternal and additive genetic effects contribute to variation in offspring traits in a lizard. *Behavioral Ecology*, **25**, 633-640.

Norton, L. R., Firbank, L. G., & Blum, H. (1999) Effects of free-air CO_2 Enrichment (FACE) on experimental grassland communities. *Functional Ecology*, **13**, 38-44.

Onoda, Y., Hirose, T., & Hikosaka, K. (2007) Effect of elevated CO_2 levels on leaf starch, nitrogen and photosynthesis of plants growing at three natural CO_2 springs in Japan. *Ecological research*, **22**, 475-484.

Onoda, Y., Hirose, T., & Hikosaka, K. (2009) Does leaf photosynthesis adapt to CO_2-enriched environments? An experiment on plants originating from three natural CO_2 springs. *New Phytologist*, **182**, 698-709.

O'Hara, R. B., Cano, J. M., Ovaskainen, O., Teplitsky, C. A. J. S., & Alho, J. S. (2008) Bayesian approaches in evolutionary quantitative genetics. *Journal of evolutionary biology*, **21**, 949-957.

Ovaskainen, O., Cano, J.M., & Merila, J. A. (2008) Bayesian framework for comparative quantitative genetics. *Proceedings of the Royal Society B: Biological Sciences,* **275**, 669-678.

Paszkowski, J., & Grossniklaus, U. (2011) Selected aspects of transgenerational epigenetic inheritance and resetting in plants. *Current Opinion in Plant Biology*, **14**, 195-203.

Pearson, P. N., & Palmer, M. R. (2000) Atmospheric carbon dioxide concentrations over the past 60 million years. *Nature*, **406**, 695.

Pecinka, A., & Scheid, O. M. (2012) Stress-Induced Chromatin Changes: A Critical View on Their Heritability. *Plant and Cell Physiology*, **53**, 801-808.

Pigliucci, M. (2003) Phenotypic integration: studying the ecology and evolution of complex phenotypes. *Ecology Letters*, **6**, 265-272.

Platt, A., Gugger, P. F., Pellegrini, M., & Sork, V. L. (2015) Genome-wide signature of local adaptation linked to variable CpG methylation in oak populations. *Molecular Ecology*, **24**, 3823-3830.

Polle, A., McKee, I., & Blaschke, L. (2001) Altered physiological and growth responses to elevated [CO_2] in offspring from holm oak (*Quercus ilex* L.) mother trees with lifetime exposure to naturally elevated [CO_2]. *Plant, Cell & Environment*, **24**, 1075-1083.

Potvin, C., & Tousignant, D. (1996) Evolutionary consequences of simulated global change: genetic adaptation or adaptive phenotypic plasticity. *Oecologia*, **108**, 683-693.

Quadrana, L., & Colot, V. (2016) Plant transgenerational epigenetics. *Annual review of genetics*, **50**, 467-491.

R Core Team, 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://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 CO_2. Plant Cell and Environment, **29**, 1730-1741.

Rapparini, F., Baraldi, R., Miglietta, F., & Loreto, F. (2004) Isoprenoid emission in trees of *Quercus pubescens* and *Quercus ilex* with lifetime exposure to naturally high CO_2 environment. *Plant, Cell & Environment*, **27**, 381-391.

Rasse, D. P., Peresta, G., & Drake, B. G. (2005) Seventeen years of elevated CO_2 exposure in a Chesapeake Bay Wetland: sustained but contrasting responses of plant growth and CO_2 uptake. *Global Change Biology*, **11**, 369-377.

Reich, P. B., Hobbie, S. E., Lee, T. D., & Pastore, M. A. (2018) Unexpected reversal of C_3 versus C_4 grass response to elevated CO_2 during a 20-year field experiment. *Science*, **360**, 317-320.

Resco De Dios, V., Mereed, T. E., Ferrio, J. P., Tissue, D. T., & Voltas, J. (2016) Intraspecific variation in juvenile tree growth under elevated CO_2 alone and with O_3: a meta-analysis. *Tree physiology*, **36**, 682-693.

Roach, D. A., & Wulff, R. D. (1987) Maternal effects in plants. *Annual review of ecology and systematics*, **18**, 209-235.

Romero, I. G., Pai, A. A., Tung, J., & Gilad, Y. (2014) RNA-seq: impact of RNA degradation on transcript quantification. *BMC biology*, **12**, 42.

Sage, R. F. (1994) Acclimation of photosynthesis to increasing atmospheric CO_2: the gas exchange perspective. *Photosynthesis research*, **39**, 351-368.

Sanchez-Gomez, D., Mancha, J. A., Cervera, M. T., & Aranda, I. (2017) Inter-genotypic differences in drought tolerance of maritime pine are modified by elevated [CO_2]. *Annals of botany*, **120**, 591-602.

Schmid, B. (1996) Genetic variation in the response of plant populations to elevated CO_2 in a nutrient-poor, calcareous grassland. In C. Korner (Ed.) *Carbon dioxide, populations, and communities* (pp. 31-50). Academic Press, San Diego, California, USA.

Schmitt, J., Niles, J., & Wulff, R. D. (1992) Norms of reaction of seed traits to maternal environments in *Plantago lanceolata*. *The American Naturalist*, **139**, 451-466.

Schurch, N. J., Schofield, P., Gierlinski, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G. G., & Blaxter, M. (2016) How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *Rna*, **22**, 839-851.

Sekhar, K. M., Sreeharsha, R. V., & Reddy, A. R. (2015) Differential responses in photosynthesis, growth and biomass yields in two mulberry genotypes grown under elevated CO_2 atmosphere. *Journal of Photochemistry and Photobiology B: Biology*, **151**, 172-179.

Seneweera, S., Makino, A., Hirotsu, N., Norton, R., & Suzuki, Y. (2011) New insight into photosynthetic acclimation to elevated CO_2: The role of leaf nitrogen and ribulose-1, 5-bisphosphate carboxylase/oxygenase content in rice leaves. *Environmental and Experimental Botany*, **71**, 128-136.

Siegenthaler, U., Stocker, T. F., Monnin, E., Luthi, D., Schwander, J., Stauffer, B., Raynaud, D., Barnola, J. M., Fischer, H., Masson-Delmotte, V. & Jouzel, J. (2005) Stable carbon cycle-climate relationship during the late Pleistocene. Measured at the University of Bern. *Science*, **310**, 1313-1317.

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**, 1117-1123.

Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. J. (2003) WinBUGS Version - 1.4 User Manual. MRC Biostatistics Unit, Cambridge. URL http://www.mrc-bsu.cam.ac.uk/bugs/.

Springer, C. J., & Ward, J. K. (2007) Flowering time and elevated atmospheric CO_2. *New Phytologist*, **176**, 243-255.

Staddon, P. L., Fitter, A. H., & Robinson, D. (1999) Effects of mycorrhizal colonization and elevated atmospheric carbon dioxide on carbon fixation and below-ground carbon partitioning in *Plantago lanceolata*. *Journal of Experimental Botany*, **50**, 853-860.

Steinger, T., Stephan, A., & Schmid, B. (2007) Predicting adaptive evolution under elevated atmospheric CO_2 in the perennial grass *Bromus erectus*. *Global Change Biology*, **13**, 1028-1039.

Stewart, A. (1996) Plantain (*Plantago lanceolata*)-a potential pasture species. Proceedings of the Conference-New Zealand Grassland Association, **58**, 77-86.

Tallis, M. J., Lin, Y., Rogers, A., Zhang, J., Street, N. R., Miglietta, F., Karnosky, D. F., De Angelis, P., Calfapietra, C., & Taylor, G. (2010) The transcriptome of Populus in elevated CO_2 reveals increased anthocyanin biosynthesis during delayed autumnal senescence. *New Phytologist*, **186**, 415-428.

Taylor, G., Street, N. R., Tricker, P. J., Sjodin, A., Graham, L., Skogstrom, O., Calfapietra, C., Scarascia-Mugnozza, G., & Jansson, S. (2005) The transcriptome of Populus in elevated CO_2. *New Phytologist,* **167**, 143-154.

Tans, P. & Keeling, R. (1958-2014) Annual mean CO_2 concentration, Mauna Loa [Online]. Available: ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2_annmean_mlo.txt [Accessed 19th Jan 2015].

Teng, N., Jin, B., Wang, Q., Hao, H., Ceulemans, R., Kuang, T., & Lin, J. (2009) No detectable maternal effects of elevated CO_2 on *Arabidopsis thaliana* over 15 generations. *PLoS One*, **4**, e6035.

Temme, A. A., Liu, J. C., Cornwell, W. K., Cornelissen, J. H., & Aerts, R. (2015) Winners always win: growth of a wide range of plant species from low to future high CO_2. *Ecology and evolution*, **5**, 4949-4961.

Thomas, J. M. G., Boote, K. J., Allen, L. H., Gallo-Meagher, M., & Davis, J. M. (2003) Elevated temperature and carbon dioxide effects on soybean seed composition and transcript abundance. *Crop Science*, **43**, 1548-1557.

Thurig, B., Korner, C., & Stocklin, J. (2003) Seed production and seed quality in a calcareous grassland in elevated CO_2. *Global Change Biology*, **9**, 873-884.

Todd, E. V., Black, M. A., & Gemmell, N. J. (2016) The power and promise of RNA-seq in ecology and evolution. *Molecular ecology*, **25**, 1224-1241.

TognettI, R., & Johnson, J. D. (1999) Responses to elevated atmospheric CO_2 concentration and nitrogen supply of *Quercus ilex* L. seedlings from a coppice stand growing at a natural CO_2 spring. *Annals of forest science*, **56**, 549-561.

Tonsor, S. J. (1985) Leptokurtic Pollen-Flow, Non-Leptokurtic Gene-Flow in a Wind-Pollinated Herb, *Plantago lanceolata* L. *Oecologia*, **67**, 442-446.

Tran, H., Wu, X., Tithi, S., Sun, M., Xie, H., & Zhang, L. (2016) A Bayesian Assignment Method for Ambiguous Bisulfite Short Reads. *PLoS ONE*, **11**, e0151826.

Tripati, A. K., Roberts, C. D., & Eagle, R. A. (2009) Coupling of CO_2 and Ice Sheet Stability Over Major Climate Transitions of the Last 20 Million Years. *Science*, **326**, 1394-1397.

Tylianakis, J.M., Didham, R. K., Bascompte, J., & Wardle, D.A. (2008) Global change and species interactions in terrestrial ecosystems. *Ecology Letters*, **11**, 1351–1363.

Uller, T. (2008) Developmental plasticity and the evolution of parental effects. *Trends in ecology & evolution*, **23**, 432-438.

Unamba, C. I., Nag, A., & Sharma, R. K. (2015) Next generation sequencing technologies: the doorway to the unexplored genomics of non-model plants. *Frontiers in Plant Science*, **6**, 1074.

Uprety, D. C., Sen, S., & Dwivedi, N. (2010) Rising atmospheric carbon dioxide on grain quality in crop plants. *Physiology and molecular biology of plants*, **16**, 215-227.

Van De Auwera, G. A., Caneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., & BANKS, E. (2013) From

FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current protocols in bioinformatics*, **43**, 11-10.

Van Loon, M. P., Riekerk, M., Dekker, S. C., Hikosaka, K., Ueda, M. U., & Anten, N. P. (2016) Plant–plant interactions mediate the plastic and genotypic response of *Plantago asiatica* to CO_2: an experiment with plant populations from naturally high CO_2 areas. *Annals of botany*, **117**, 1197-1207.

Van Hinsberg, A., & Van Tienderen, P. (1997) Variation in growth form in relation to spectral light quality (red/far-red ratio) in *Plantago lanceolata* L. in sun and shade populations. *Oecologia*, **111**, 452-459.

Veeckman, E., Ruttink, T., & Vandepoele, K. (2016) Are we there yet? Reliably estimating the completeness of plant genome sequences. *The Plant Cell*, **28**, 1759-1768.

Villemereuil, P., Gimenez, O., & Doligez, B. (2013) Comparing parent–offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations: a simulation study for Gaussian and binary traits. *Methods in Ecology and Evolution*, **4**, 260-275.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008) Heritability in the genomics era—concepts and misconceptions. *Nature reviews genetics*, **9**, 255.

Voelckel, C., Gruenheit, N., & Lockhart, P. (2017) Evolutionary transcriptomics and proteomics: insight into plant adaptation. *Trends in plant science*, **22**, 462-471.

Wajid, B., Serpedin, E. (2012) Review of general algorithmic features for genome assemblers for next generation sequencers. *Genomics Proteomics Bioinformatics*, **10**, 58–73.

Wand, S. J. E., Midgely, G., Jones, M. H., & Curtis P. S. (1999) Responses of wild C4 and C3 grass (*Poaceae*) species to elevated atmospheric CO_2 concentration: a meta-analytic test of current theories and perceptions. *Global Change Biology*, **5**, 723-741.

Wang, X. Z., & Curtis, P. S. (2001) Gender-specific responses of *Populus tremuloides* to atmospheric CO_2 enrichment. *New Phytologist*, **150**, 675-684.

Wang, Z., Gerstein, M., & Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews genetics*, **10**, 57.

Wang, D., Heckathorn, S. A., Wang, X. Z., & Philpott, S. M. (2012) A meta-analysis of plant physiological and growth responses to temperature and elevated CO_2. *Oecologia*, **169**, 1-13.

Wang, X., Liu, Q., & Zhang, B. (2014) Leveraging the complementary nature of RNA-Seq and shotgun proteomics data. *Proteomics*, **14**, 2676-2687.

Ward, J., & Strain, B. (1997) Effects of low and elevated CO_2 partial pressure on growth and reproduction of *Arabidopsis thaliana* from different elevations. *Plant, Cell & Environment*, **20**, 254-260.

Ward, J. K., Antonovics, J., Thomas, R. B., & Strain, B. R. (2000) Is atmospheric CO_2 selective agent on model C_3 annuals? *Oecologia*, **123**, 330-341.

Wardle, G., Buckley, Y., & PlantPopNet steering group (2014) *PlantPopNet: A Spatially Distributed Model System for Population Ecology*. Ecology Society of Australia 2014 Annual Conference, Alice Springs NT, Australia. Available: http://plantago.plantpopnet.com.

Watson-Lazowski, A., Lin, Y., Miglietta, F., Edwards, R. J., Chapman, M. A., & Taylor, G. (2016) Plant adaptation or acclimation to rising CO_2? Insight from first multigenerational RNA-Seq transcriptome. *Global Change Biology*, **22**, 3760-3773.

Wheeler, T., & Von Braun, J. (2013) Climate Change Impacts on Global Food Security. *Science*, **341**, 508-513.

Wieneke, S., Prati, D., Brandl, R., Stocklin, J., & Auge, H. (2004) Genetic variation in *Sanguisorba* minor after 6 years *in situ* selection under elevated CO_2. *Global Change Biology*, **10**, 1389-1401.

Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L.E., & Nussey, D. H. (2010) An ecologist's guide to the animal model. *Journal of Animal Ecology*, **79**, 13-26.

Wolak, M. E., & Keller, L. F. (2014) Dominance genetic variance and inbreeding in natural populations. In A. Charmantier *et al*., (Eds.) *Quantitative Genetics in the Wild* (pp. 104-127). Oxford University Press, Oxford, UK.

Wolf, J. B., Brodie III, E. D., Cheverud, J. M., Moore, A. J., & Wade, M. J. (1998) Evolutionary consequences of indirect genetic effects. *Trends in ecology & evolution*, **13**, 64-69.

Wolf, J. B., & Wade, M. J. (2009) What are maternal effects (and what are they not)? *Philosophical Transactions of the Royal Society B-Biological Sciences*, **364**, 1107-1115.

Wulff, R. D., & Alexander, H. M. (1985) Intraspecific variation in the response to CO_2 enrichment in seeds and seedlings of *Plantago lanceolata* L. *Oecologia*, **66**, 458-460.

Wulff, R., Caceres, A., & Schmitt, J. (1994) Seed and seedling responses to maternal and offspring environments in *Plantago lanceolata*. *Functional Ecology*, **8**, 763-769.

Wu, H., Wang, C., & Wu, Z. (2014) PROPER: comprehensive power evaluation for differential expression using RNA-seq. *Bioinformatics*, **31**, 233-241.

Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S., Huang, W., He, G., Gu, S., & Ll, S. (2014) SOAPdenovo-Trans: *de novo* transcriptome assembly with short RNA-Seq reads. *Bioinformatics*, **30**, 1660-1666.

Young, K. A., & Schmitt, J. (1995) Genetic-Variation and Phenotypic Plasticity of Pollen Release and Capture Height in *Plantago lanceolata*. *Functional Ecology*, **9**, 725-733.

Zhao, H. X., Xu, X. A., Zhang, Y. B., Korpelainen, H., & Li, C. Y. (2011) Nitrogen deposition limits photosynthetic response to elevated CO_2 differentially in a dioecious species. *Oecologia*, **165**, 41-54.

Ziska, L. H. (2003a) Evaluation of the growth response of six invasive species to past, present and future atmospheric carbon dioxide. *Journal of Experimental Botany*, **54**, 395-404.

Ziska, L. H. (2003b) The impact of nitrogen supply on the potential response of a noxious, invasive weed, Canada thistle (*Cirsium arvense*) to recent increases in atmospheric carbon dioxide. *Physiologia Plantarum*, **119**, 105-112.

Ziska, L. H., & Bunce, J. A. (2000) Sensitivity of field-grown soybean to future atmospheric CO_2: selection for improved productivity in the 21st century. *Australian Journal of Plant Physiology*, **27**, 979-984.

Chapter 2 FACE facts hold for multiple generations; Evidence from natural CO_2 springs

Jasmine M. Saban[a], Mark A. Chapman[a] & Gail Taylor[a,b]

[a]Biological Sciences, University of Southampton, Life Sciences, Building 85, Highfield Campus, Southampton, SO17 1BJ, UK

[b]Department of Plant Sciences, University of California, Davis, CA 95616, USA

Correspondence: Gail Taylor, Tel: +1- 530-752-4362. E-mail address: gtaylor@ucdavis.edu

Opinion article

Key words: Natural CO_2 spring, plant physiology, plant adaptation, plasticity, climate change, meta-analysis

2.1 Abstract

Rising atmospheric CO_2 concentration ($[CO_2]$) is a key driver of enhanced global greening, thought to account for up to 70% of increased global vegetation in recent decades. CO_2 fertilization effects have further profound implications for ecosystems, food security and biosphere-atmosphere feedbacks. However, it is also possible that current trends will not continue, due to ecosystem level constraints and as plants acclimate to future CO_2 concentrations. Future predictions of plant response to rising $[CO_2]$ are often validated using single generation short-term FACE (Free Air CO_2 Enrichment) experiments, however whether this accurately represents vegetation response over decades is unclear. Whether adaptation occurs (and if so by what mechanism) has yet to be elucidated. Here, we propose that naturally occurring high CO_2 springs provide a proxy to quantify the multigenerational and long-term impacts of rising $[CO_2]$ in herbaceous and woody species respectively, such that plasticity and genetic adaptation can be quantified together in these systems. In this first meta-analysis of responses to elevated $[CO_2]$ at natural CO_2 springs, we show that the magnitude and direction of change in eight of nine functional plant traits is consistent between spring and FACE experiments. We found increased photosynthesis (49.8 % in spring experiments, comparable to 32.1 % in FACE experiments) and leaf starch (58.6 % spring, 84.3 % FACE), decreased stomatal conductance (g_s, 27.2 % spring, 21.1 % FACE), leaf nitrogen content (6.3 % spring, 13.3 % FACE) and Specific Leaf Area (SLA, 9.7 % spring, 6.0 % FACE). These findings not only validate the use of these sites for studying multigenerational plant response to elevated $[CO_2]$, but additionally suggest that long-term positive photosynthetic response to rising $[CO_2]$ are likely to continue as predicted by single generation exposure FACE experiments.

2.2 Introduction

Average atmospheric global [CO_2] is now consistently above 400 ppm for the first time in around 23 million years of evolutionary time (Pearson and Palmer, 2000). Increased atmospheric [CO_2] will be a key feature of future climates, and although there is clear resolve to cap atmospheric [CO_2] to below 530 ppm in order to avoid catastrophic ecosystem change under global warming, it remains unclear whether these [CO_2] targets will be met (Stocker, 2013). Despite the profound impact of [CO_2] on plant functioning, future predictions of plant responses to elevated [CO_2] are predominantly validated using experimental data derived from single generation experiments, which model only plant phenotypic plasticity. These plastic responses have been extensively quantified in experimental systems ranging from small controlled environment studies to large ecosystem experiments using FACE, and generalised through meta-analyses that are used to inform or validate models and predictions (Ainsworth and Long, 2005; Dybzinski *et al.,* 2015; Vanuytrecht and Thorburn, 2017). While these experiments have played a pivotal role in informing short term projections of, for example, food security (Wheeler and Von Braun, 2013; Myers *et al*., 2014) and the likely distribution of plant ecotones in a changing climate (Barnaby and Ziska, 2012; Smith *et al*., 2016; Forkel *et al*., 2016), extrapolating to predict consequences of climate change for the end of the century may be precarious.

Beyond single generation plastic plant responses to elevated [CO_2] there is some evidence for adaptation (the inheritance of derived characteristics that enhance fitness in a given environment) but a lack of conclusive evidence that elevated [CO_2] could act as a selective agent on either genetic or epigenetic variation under climate change in the natural environment (Ward *et al.,* 2000; Frenck *et al.,* 2013; Leakey and Lau, 2012). Regardless, there is a wealth of evidence to suggest that transgenerational indirect genetic effects

can and do contribute to plant response to elevated [CO_2] over multiple generations (Jablonski *et al.,* 2002; Springer and Ward, 2007; Johnston and Reekie, 2008).

Multigenerational experiments are a key challenge for the study of plant adaptation, owing to the time, energy and expense of growing plants under such conditions long-term, especially for long-lived and large plant species. Facilities are expensive and labour intensive to build and maintain, contribute to anthropogenic emissions, and cannot provide information on population responses to elevated [CO_2] over generations in the timeframe needed to prepare for climate change. To this end, the use of plants surrounding natural CO_2 springs are a precious resource to further elucidate evolutionary adaptation and long-term response to elevated [CO_2]. Plants growing at natural CO_2 springs have previously been utilised to study physiological response to rising [CO_2] but have largely been abandoned due to concerns about CO_2 emission variability over time and contamination by other exhaust gases. Here we propose that as with other systems, provided these limitations are appropriately managed, spring sites represent a valuable resource that can contribute to our understanding of multigenerational plant response to elevated [CO_2] in combination with other systems. In this first meta-analysis of natural CO_2 spring plant response to elevated [CO_2] we highlight sites at which research has been conducted and synthesise available data, comparing responses to those in FACE experiments.

2.3 Materials and methods

2.3.1 Systematic search

To evaluate research at CO_2 springs, we captured available data through a systematic search of the literature on 3rd July 2017. Using a structured string search and standard systematic review methodology, 3,294 studies were collated from Web of Science and screened according to strict inclusion criteria to provide a database of studies measuring traits in plants at natural CO_2 springs compared to an ecologically similar control site in close proximity. These inclusion criteria are outlined in Appendix A. 1 and include (among others) that there must be a difference in [CO_2] of at least 100 ppm between spring and control sites, and that sites are only included where contamination by [H_2S] < 0.02 ppm and [SO_2] < 0.015 ppm.

To avoid non-independence as a result of multiple measurements of a trait being reported in a single publication, only one data point was taken for a trait for each species in each study. The data point extracted was decided on a trait by trait basis, for example photosynthetic measurements were taken at midday and during summer months if they were measured multiple times. In order to calculate effect sizes, mean, sample size and standard deviation were obtained from the text, tables or extracted from figures using DATATHIEF (Tummers, 2006). Authors were contacted if there was insufficient data reported for inclusion in the meta-analysis and many authors kindly provided additional data.

Ultimately we analysed data from 16 sufficiently replicated traits across 39 species in 26 papers (Appendix A. 2 and Appendix Table A. 1). This represents a subset of studies that have ever been used to study plant response at natural CO_2 springs because we were

unable to include traits (and therefore studies) where fewer than five species or studies measured the trait across the database.

2.3.2 Statistical analysis

2.3.2.1 Effect size calculation

To compare trait differences between spring (elevated [CO_2]) and control (ambient [CO_2]) groups we calculated the log response ratio (lnR) for each trait under elevated [CO_2] as a metric for analysis. Log response ratio quantifies the proportional difference in population mean for a trait under elevated [CO_2] at the spring site relative to ambient [CO_2] at the control site. The log transformation is used to linearise the relationship between the two variables and to obtain residuals that are approximately symmetrically distributed where the sampling distribution may otherwise be skewed (particularly in small samples) (Hedges *et al*., 1999). Log response ratio was calculated as:

$$lnR = \ln \frac{\bar{x}_{Spring}}{\bar{x}_{Control}} = \ln(\bar{x}_{Spring}) - \ln(\bar{x}_{Control})$$

where $\bar{x}_{Spring}$ is the mean trait value for plants growing under elevated [CO_2] at the spring site and $\bar{x}_{Control}$ is the mean trait value for plants growing in ambient [CO_2] at the control site. For more intuitive presentation, the log response ratio is converted to percentage difference using the formula [(R-1) x 100]. All statistical analyses were performed in R version 3.2.2 (R Core Team, 2015).

2.3.2.2 Meta-analysis

A random effects model was applied to calculate overall effect of elevated [CO_2] on populations at the spring site relative to the control populations. Random effects models were used to account for environmental variation by assuming that true effect size varies

between studies forming a distribution of effect sizes. The studies within the analysis are assumed to be a random sample of this distribution and the overall summary effect of a random effects model estimates the mean of the distribution of true effect sizes. The null hypothesis is that the mean of the distribution of effects is zero. The effect size of each species from each study was weighted using the inverse of its variance. All models used restricted maximum likelihood estimation. If a 95 % confidence interval for a trait did not overlap zero then a significant response was considered in plants exposed to elevated [CO_2] relative to their ambient counterparts at control sites.

2.3.3 Assessing heterogeneity between studies

We examined variation between studies, partitioning it from within study error using the heterogeneity statistic Q and subsequently I^2 using the formula $I^2 = 100\ \% \times (Q\text{-}df)/Q$ (Higgins and Thompson, 2002). The I^2 statistic describes the percentage of variation across studies that is due to heterogeneity rather than chance. Of the sixteen traits that were measured, the Q and I^2 statistics indicated that thirteen traits showed a significant degree of between-study heterogeneity and effect sizes were calculated using a random effects model to account for this (Appendix Table A. 2). For three traits (V_{cmax}, J_{max} and leaf carbon: nitrogen ratio) we found Q with $p>0.05$ and/or an I^2 statistic <50 % suggesting the variation in findings is compatible with chance alone (homogeneity) and therefore a fixed effect model was used to calculate these effect sizes.

Significant heterogeneity between studies existed for all traits analysed, suggesting that almost all of the variability in estimates was due to variation between samples rather than sampling error. This is common among ecological studies where an average I^2 of 83-92 % were reported in an analysis of ecological meta-analyses (Senior *et al*., 2016). Given that individual samples come from a diverse array of global sites and from multiple

functional groups, this heterogeneity is to be expected, but it is also useful to explore the basis of this heterogeneity by modelling potential moderator variables. Subgroup analysis was performed to examine trait changes in functional groups where sample size permitted (as trees, including both deciduous and evergreen trees, and herbs, including grasses, with forbs also analysed separately for stomatal conductance for comparison to FACE analyses), and a random effects meta-regression model with defined moderator variables was fitted to the data to examine the effect of these moderator variables in the R package glmulti (Calcagno and De Mazancourt, 2010). Plant functional group and climate zone were used as moderator variables for meta-regression analysis. For categorical variables the category was considered an important predictor if the 95 % confidence intervals of the category estimate did not overlap those of the overall effect size. Photosynthetic rate at growth [CO_2] was the only trait where either of these categorical predictors were considered significant in predicting the estimate under meta-regression. For this trait we further decomposed the categorical variable 'climate zone' to two continuous variables; average maximum daily temperature and annual precipitation for meta-regression. Variance explained by a predictor variable was calculated through ANOVA of the model containing only this predictor variable versus the null model.

2.3.4 Publication bias

In ecological studies there may be a bias towards publishing positive and significant results, and studies with larger sample size have more power to detect significant differences, indeed Haworth *et al*., (2016) have suggested that publication bias has resulted in a significant over-estimation of the impacts of elevated [CO_2] on plants in FACE study meta-analyses. Publication bias was quantified using weighted regression with multiplicative dispersion using standard error as the predictor to detect funnel plot

asymmetry (the classical Egger's test), using the *regtest* function in the METAFOR package (Viechtbauer, 2010), by examining plots of the data and by estimating the fail-safe number (Appendix Table A. 3; Rosenberg, 2005). From analyses of these tests and examination of the normal Q-Q and funnel plots, we acknowledge that publication bias and the presence of outliers reduce confidence in the model estimates of summary effect for adaxial stomatal density, leaf chlorophyll content and leaf carbon content. Our interpretation of these results is duly cautious.

We additionally performed sensitivity analysis by applying weight functions to the effect sizes of studies to determine the impact of moderate publication bias. Assuming moderate selection of publication bias on the gathered dataset, we estimate that effect sizes in this study may be inflated by 6-13%. This is similar in magnitude to the estimated inflation of FACE study effect sizes by 5-15% due to moderate reporting bias (Haworth *et al*., 2016).

2.4 Results

A systematic search of the literature revealed CO_2 springs that have previously been utilised for this research occur extensively across the globe and range in latitude, temperature and rainfall (Figure 2.1). Significant differences in vegetation types and species present at each site are apparent, including many long-lived tree species that are difficult to work with experimentally. The most comprehensively studied and characterised springs are located in Italy and Japan (Figure 2.1**,** Appendix Table A. 1)

Photosynthetic rate at growth [CO_2] was significantly enhanced, by 49.8 % (±10.6 %) in spring versus control sites (Figure 2.2**,** Appendix Figure A. 1). This is comparable to the 31 % enhancement observed in a meta-analysis of plants at FACE facilities (Figure 2.3; Ainsworth and Long, 2005). Climate classification explained 60.9 % of the variation in photosynthetic rate response to elevated [CO_2], while functional group did not significantly explain variation. Much of the variation was attributable to studies at a site in Venezuela, where very high [CO_2] was measured at the vents (27,000-35,000 ppm) and there was a lack of vertical characterisation of [CO_2] at the study site. This was also the only site in the tropical biome, highlighting that plant responses to elevated [CO_2] in the tropics is a clear gap in our understanding of plant responses to elevated [CO_2] globally (Jones *et al*., 2014). Using meta-regression average yearly maximum temperature was identified as a key component of photosynthetic response to elevated [CO_2]. On average each 1 °C increase in average maximum daily temperature increased the effect of elevated [CO_2] on photosynthetic rate by 4.8 % over the range of temperatures measured, a finding that is well supported by existing research (Ainsworth and Long, 2005; Wang *et al.,* 2012). The impact of elevated [CO_2] on maximum carboxylation rate (V_{cmax}) and maximum rate of electron transport (J_{max}) were measured in fewer studies

than photosynthetic rate. Effect sizes were calculated at -17.3 % (±4.0 %) and -9.4 % (±3.4 %) in spring versus control, respectively (Figure 2.2). The greater reduction in V_{cmax} relative to J_{max} suggests that where acclimation of photosynthesis occurs in these plants it is likely through a reduction in ribulose bisphosphate carboxylase (RuBisCO) content or activity.

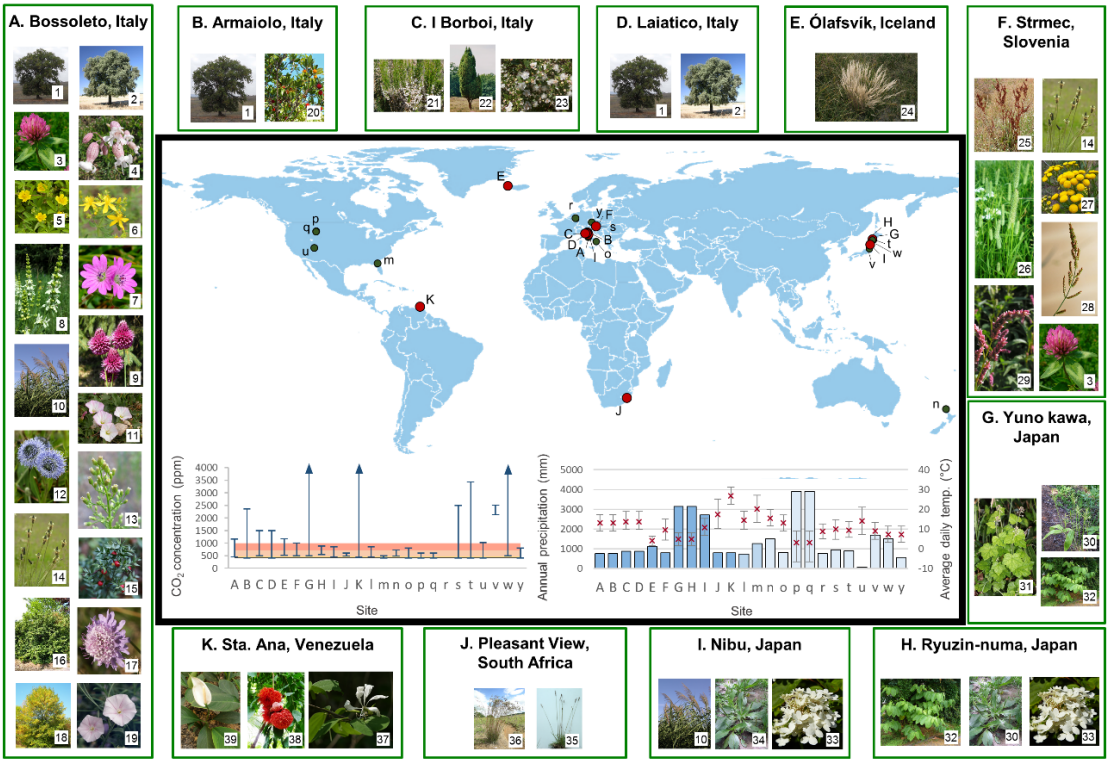


Figure 2.1: Sites of naturally elevated [CO_2] that have been used to study plant adaptation to elevated [CO_2].

21 sites are identified. Sites indicated by a red dot, and denoted with a capital letter were analysed in this meta-analysis. Sites indicated by a green dot and denoted by lower case letters were not used by studies included in this meta-analysis but studies at these sites have been published. Graphs show [CO_2] and climatic conditions of each site, graph of [CO_2] has predicted scenarios for the end of the century coloured from yellow to orange (IPCC., 2014). Green boxes for each of the sites used in the meta-analysis show images of species represented in the meta-analysis. Images were acquired from Google Images, labelled for reuse. Sites (l-u) l. Solfatara, Italy, m. Ichetucknee springs, USA, n. Hakanoa springs, New Zealand, o. Orciatico, Italy, p. Ochre springs, USA, q. Mammoth upper terrace, USA, r. Laacher See, Germany, s. Rihtarovci, Slovenia, t. Tashiro, Japan, u. Burning hills, USA, v. Asahi, Japan, w. Kosaka, Japan, y. Plesná stream, Czech Republic. Species; 1. *Quercus pubescens*, 2. *Quercus ilex*, 3. *Trifolium pratense*, 4. *Silene vulgaris*, 5. *Potentilla reptans*, 6. *Hypericum perforatum*, 7. *Gerranium molle*, 8. *Stachys recta*, 9. *Allium sphaerocephalon*, 10. *Phragmites australis*, 11. *Convolvulus arvensis*, 12. *Globularia punctata*, 13.

Conyza candensis, 14. *Plantago lanceolata*, 15. *Ruscus aculeates*, 16. *Buxus sempervirens*, 17. *Scabiosa columbaria*, 18. *Fraxinus ornus*,19. *Convolvulus cantabrica*, 20. *Arbutus unedo*, 21. *Erica arborea*, 22. *Juniperus communis*, 23. *Myrtus communis*, 24. *Nardus stricta*, 25. *Rumex crispus*, 26. *Phleum pratense*, 27. *Tanacetum vulgaris*, 28, *Echinochloa crus-galli*, 29. *Polygonum hydropiper*, 30. *Sasa kurilensis*, 31. *Tiarella polyphylla*, 32. *Polygonum sachalinense*,

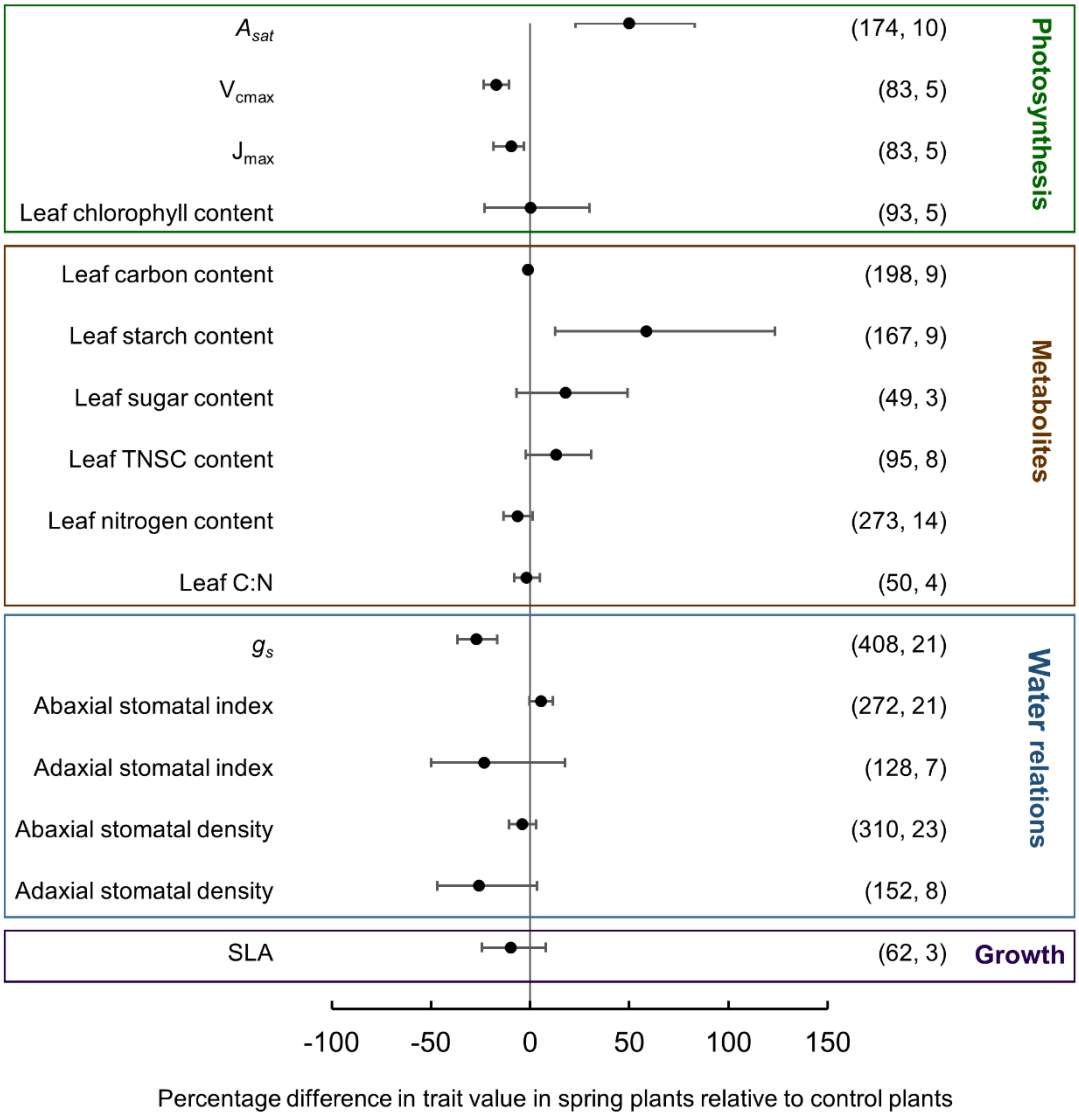


Figure 2.2: Trait percentage difference between plants at elevated and ambient [CO_2] at naturally occurring CO_2 springs; Meta-analysis summary effect sizes.

Traits: Photosynthetic rate at growth [CO_2] (A), maximum carboxylation rate (V_{cmax}), maximum rate of electron transport (J_{max}), leaf carbon content, leaf sugar content, leaf starch content, leaf total non-structural carbohydrate (TNSC) content, leaf nitrogen content, leaf carbon to nitrogen ratio (C:N), stomatal conductance (gs), abaxial stomatal index ((stomatal density/(stomatal density + epidermal cell density)) x 100), adaxial stomatal index, abaxial stomatal density (stomata per unit area), adaxial stomatal density and specific leaf area (SLA). Symbols represent the percentage difference at elevated [CO_2] and their 95 % confidence intervals. Total sample size (n) followed by the number of species included for each variable appear in parentheses after the symbol.

A large and significant increase in starch content +58.6 % (±19.1 %) indicates that excess photosynthate from enhanced photosynthesis are increasingly converted to starch for storage for spring grown plants in response to elevated [CO_2]. Leaf total non-structural carbohydrates (TNSC) were not significantly increased +13.1 % (±7.6 %), and neither was leaf sugar content +17.9 % (±12.7 %). Additionally, no difference was seen in total carbon content in the leaves of plants at natural CO_2 springs but with publication bias in this trait reducing confidence in the estimated effect size -1.6 % (±0.7 %).

Trait	A_{sat}	V_{cmax}	J_{max}	g_s	Leaf starch	Leaf sugar	Leaf nitrogen	SLA	chlorophyll
Spring	**+49.8%** [22.7, 82.8]	**-17.2%** [-23.4, -10.6]	**-9.4%** [-18.6, -3.3]	**-27.2%** [-36.5, -16.6]	**+58.6%** [12.5, 123.6]	**+17.9%** [-6.8, 49.1]	**-6.3%** [-13.3, 1.2]	**-9.7%** [-24.3, 7.7]	**0.1%** [-22.8, 30.0]
FACE	**+32.1%** [29.0, 35.2]	**-10.1%** [-13.0, -7.1]	**-3.5%** [-5.6, -1.5]	**-21.2%** [-19.3, -23.0]	**+84.3%** [61.3, 107.3]	**+31.9%** [0.2, 47.9]	**-13.3%** [-16.8, -9.8]	**-6.0%** [-8.2, -3.6]	**-17.1%** [-27.2, -5.9]

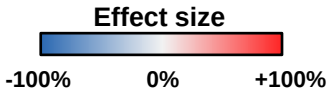


Figure 2.3: Comparison of response to elevated [CO_2] in this CO_2 spring meta-analysis to FACE meta-analyses (Ainsworth and Long, 2005, Ainsworth and Rogers, 2007).
Average percentage difference between plants growing at elevated versus ambient CO_2 is given in bold with 95% confidence lower and upper boundaries given in square brackets. Squares are coloured according the percentage difference as shown in the colour scale. Traits are photosynthetic rate at growth CO_2 (A_{sat}), maximum carboxylation rate (V_{cmax}), maximum rate of electron transport (J_{max}), stomatal conductance (g_s), leaf starch content, leaf sugar content, leaf nitrogen content, and specific leaf area (SLA).

When a global effect size was calculated, leaf nitrogen content did not differ between CO_2 spring and control sites -6.3 % (±4.0 %), although the magnitude and direction of the effect size was consistent with those observed in FACE meta-analyses. Spring sites typically have acidic soils (with pH 3.3-6.8, where recorded, at sites in this study, Appendix Table A2.1) and relatively anaerobic conditions which would predict higher soil concentrations of ammonium and reduced nitrate availability which could in part explain

the apparent lack of photosynthetic acclimation seen in plants at CO_2 springs (Onoda *et al*., 2007; Bloom *et al*., 2010). When functional groups were analysed separately in subgroup analysis, trees showed a significant decrease in leaf nitrogen -10.4 % (±3.6 %), while there was no significant difference in herbaceous plants or the global affect size across both functional groups (Figure 2.4). However, our estimation of leaf nitrogen content in trees was limited by the lack of replication across sites, with five species being measured in only two sites (Figure 2.4).

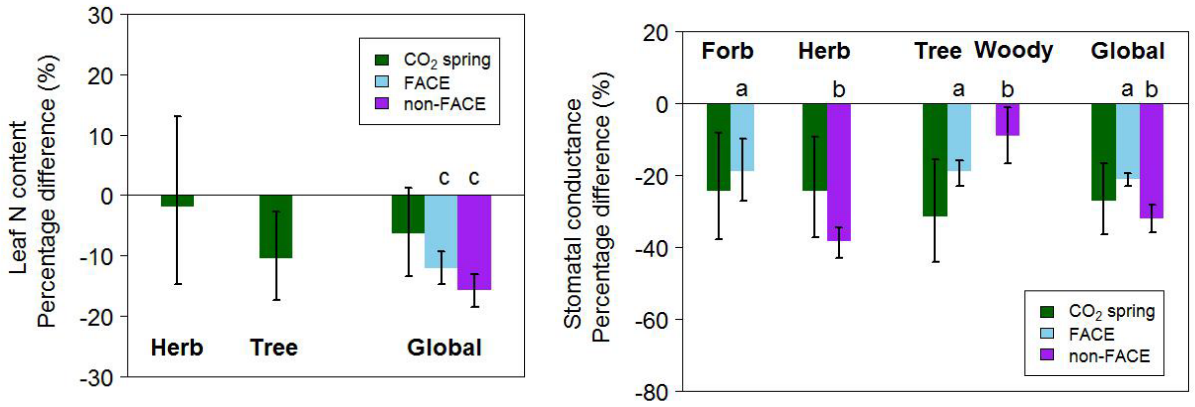


Figure 2.4: Percentage difference in stomatal conductance and leaf nitrogen content of plants growing at elevated [CO_2] relative to ambient [CO_2].

Global effect size is presented, as well as effect size when divided by functional group for this meta-analysis of plants at natural CO_2 springs (green), with comparison to a meta-analysis of plants at FACE (blue) and non-FACE (purple) facilities. Meta-analyses referenced are a) Ainsworth and Long, 2005 b) Wang *et al*., 2012 and c) Loladze, 2014.

A significant reduction in stomatal conductance of -27.2 % (±7.2 %) in plants at spring versus control plants suggests water savings through reduced transpiration, and this was of a similar magnitude to reduced stomatal conductance measured in FACE meta-analysis (Figure 2.3, Appendix Figure A. 2). Although we acknowledge that our comparison to FACE and semi- or closed design (non-FACE) meta-analyses are confounded by differences in average CO_2 concentration of studies (Table 2.1) this directional response is consistent across functional groups and experimental designs (Figure 2.4). There were no consistent responses in stomatal density (SD) or stomatal index (SI) to elevated [CO_2] in springs (Figure 2.2). A decrease in SD may be observed more frequently for species exposed to

elevated [CO_2] in controlled environment (Woodward and Kelly, 1995) and FACE studies (Ainsworth and Rogers, 2007), with ~60 % of studies in both analyses evidencing decreased SD under elevated [CO_2]. In this meta-analysis fewer than 50% of observations had decreased SD in spring sites, with average effect size of -4.0 % (±3.7 %), comparable to the non-significant 5 % decrease in FACE meta-analysis (Ainsworth and Rogers, 2007).

Although SLA did not differ significantly between spring and control populations, the magnitude and direction of the effect size -9.7 % (±9.41 %) was consistent with FACE meta-analyses (Figure 2.3). Since estimating increases in Leaf Area Index (LAI) to predict global greening and evapotranspiration under climate change depend upon changes in SLA, robust estimates of SLA response to elevated [CO_2] based on empirical data is crucial to these predictions (De Kauwe *et al*., 2014). Meta-analysis of SLA in plants at natural CO_2 springs tends to support the decline in SLA in FACE meta-analyses used to inform these models, but additionally suggests that some plant species may increase SLA under elevated [CO_2], and this requires further investigation.

Meta-analysis	Experimental designs analysed	Average [CO_2] of elevated treatments
This meta-analysis	Natural CO_2 springs	791 ppm
Ainsworth and Long, 2005	FACE	~560 ppm
Ainsworth and Rogers, 2007	FACE	567 ppm
Wang *et al*., 2012	Semi-open and closed systems	702 ppm
Loladze, 2014	FACE	560 ppm
	Semi-open and closed systems	732 ppm

Table 2.1: A comparison of the average [CO_2] of experiments included in five meta-analyses

Across nine traits that had been measured in both this, the first meta-analysis of response at spring sites, and comparable meta-analyses of responses at FACE sites, eight traits were consistent in direction and magnitude (Figure 2.3). Leaf chlorophyll content was the only trait that was inconsistent in direction between the two meta-analyses, however the sample size of this trait for meta-analysis at CO_2 springs was small (with only five species

studied) and was affected by publication bias. Other traits, such as leaf sugar content and SLA, although consistent in direction and magnitude showed larger variability than in FACE meta-analyses. Whether this is solely an artefact of our small sample size compared to the large data availability for FACE meta-analyses, or whether this is a result of comparing wild plants with the traditionally greater proportion of crop plants in FACE meta-analyses is not discernible from this data set.

2.5 Discussion

Here we report the first meta-analysis for data collected from plants in natural CO_2 springs. Although these sites were initially suggested to study multigenerational plant response to elevated [CO_2] in the early 1990's, this research was largely focussed on physiological and biochemical analysis since until recently, genomic technologies were unavailable for wild non-model plant species such as those found at spring sites. We propose that they should now be re-examined given the potential of new sequencing technologies to provide insight into future adaptive response to increased atmospheric [CO_2]. Through meta-analysis we show that long-term and multigenerational responses of plants to elevated [CO_2] at natural CO_2 springs are remarkably consistent with those measured in single-generation FACE studies with eight of a panel of nine traits showing consistency. This is a key finding since it suggests that the magnitude and direction of long-term response of plants to elevated [CO_2] may be adequately predicted by single generation experiments, regardless of the mechanisms coordinating this response. The consequences of this finding may be wide-ranging in supporting predictions of ecosystem change from models that have been parameterised with FACE data, for example the maintenance of positive photosynthetic rate which combined with other environmental factors may lead to the maintenance of global greening. Additionally our results suggest that these sites are valuable to disentangle the role of indirect and direct genetic effects, and environmental constraints in the multigenerational response. This is particularly timely given the rapid recent progress in reduced cost of sequencing and software development for *de novo* genome and transcriptome assembly in non-model organisms (Moreton *et al*., 2016; Li and Harkess, 2018).

A panel of eight traits in this study highlighted consistent response of FACE and spring-grown plants. Altered gas exchange and photosynthetic rate are key features of the multigenerational response to elevated [CO_2] and these trait differences were slightly enhanced relative to those at FACE sites (Ainsworth and Rogers, 2007). This may reflect the higher CO_2 concentrations at spring study sites (800 – 1000 ppm, representative of the 'worst case' RCP8.5 climate scenario) relative to those across FACE sites (530-580 ppm, representative of the more moderate stabilisation pathway RCP4.5) (IPCC, 2014) but suggests that photosynthetic rate is likely to be maintained despite environmental constraints and resource limitations, and over multiple generations. The magnitude of reduced stomatal conductance supports conclusions from FACE experiments that stomatal conductance does not acclimate to elevated [CO_2] (Leakey *et al*., 2009) even over multiple generations, whether plastically coordinated or as the result of genetic assimilation or accommodation (Grossman and Rice, 2014). It is increasingly recognised that there is large variation in stomatal density (SD) response to elevated [CO_2] both within and between species, and with significant dependence on other environmental factors (Haworth *et al*., 2010; Haworth *et al*., 2015; Yan *et al*., 2017). In accordance with FACE meta-analyses, our data provide no conclusive evidence that there is a general reduction in stomatal density in CO_2 spring sites (Ainsworth and Rogers, 2007). Increased abaxial stomatal index was observed for some species but there was large variation across species, with a non-significant mean effect size of 5.4 % (±7.2 %), which may indicate that decreases in SD result from expanding epidermal cells rather than a decline in stomatal initiation. Adaxial stomatal density and index were measured in fewer species and showed large variation. However, comparison between this meta-analysis and the response of plants to elevated [CO_2] in FACE experiments were limited because meta-analyses of stomatal density (SD) response to elevated [CO_2] in other systems (and many

of the papers from which they take data) did not explicitly state whether SD was measured from the abaxial or adaxial leaf surface (Woodward and Kelly, 1995; Ainsworth and Rogers, 2007). Since the mechanisms of stomatal patterning on these surfaces are independent this is an important distinction, particularly because the ratio of stomata on these surfaces (and thus their role in gas exchange) is highly variable between species.

Although the sample size of this meta-analysis was small, the study of plants growing *in situ* at natural CO_2 springs meant that there was large diversity in plant species studied, which included functional groups such as trees that are difficult to study experimentally. Subgroup analysis of functional groups on traits evidenced that herbs growing at natural CO_2 springs had enhanced photosynthetic rate, reduced stomatal conductance and no difference in nitrogen content of the leaves relative to control plants. Trees in contrast showed similarly enhanced photosynthetic rate and reduced stomatal conductance but a significant decrease in nitrogen content of the leaves at spring sites. These differences in leaf nitrogen content response between functional groups could be due to several factors not quantified here, including differences in nitrogen allocation, differential biotic interactions such as the association of mycorrhiza to trees versus herbs, or abiotic factors such as differential light availability or soil accessibility (Osada *et al*., 2010; Ueda *et al*., 2017).

Interpretation of plant responses at CO_2 springs would clearly be improved by further characterisation of soil properties across the sites including nitrogen source (ammonium and nitrate availability), pH (characterised in just under half of sites globally) and soil CO_2 concentration (Pfanz *et al*., 2007; Ueda *et al*., 2017). For example, there is limited information available on soil nitrogen at natural CO_2 springs, but where quantified, total nitrogen pools have generally been found to be larger in spring than control soils (Newton

et al., 1996; Ross *et al*., 2000; Ueda *et al*., 2017). Of total soil nitrogen content, smaller inorganic nitrogen pools in CO_2 spring sites may be indicative of increased uptake by plants under elevated [CO_2] (Ueda *et al*., 2017), though nitrogen content of leaf litter returning to the soil generally shows decreased or unchanged nitrogen content at natural CO_2 springs (Ross *et al*., 1996; Gahrooee, 1998; Coûteaux *et al*., 1999; Cotrufo *et al*., 1999) suggesting changes in plant nitrogen allocation that may impact plant-soil nitrogen cycling (see Gamage *et al*., 2018). Where investigated, and likely as a result of anaerobic and acidic soil conditions characteristic of natural CO_2 springs, ammonium is the predominant form of inorganic nitrogen (Onoda *et al*., 2007; Osada *et al*., 2010; Ueda *et al*., 2017), which may facilitate the positive response of spring plant photosynthetic rate to elevated [CO_2], since plants primarily utilising ammonium as an inorganic nitrogen source will be less impacted by inhibition of nitrate assimilation by elevated [CO_2] than plants utilising nitrate (Bloom, 2015; Rubio-Asensio and Bloom, 2016). Soil properties also influence the occurrence of soil microorganisms with impact on plant-soil nutrient cycling which may well be key to understanding ecosystem response to long-term CO_2 exposure at natural CO_2 springs. Microorganism populations including arbuscular mycorrhizal fungi (Rillig *et al*., 2000; Maček *et al*., 2011; Maček *et al*., 2012; Maček, 2013), archea (Krüger *et al*., 2011; Šibanc *et al*., 2014), bacteria (Videmšek *et al*., 2009; Krüger *et al*., 2011; Frerichs *et al*., 2013; Šibanc *et al*., 2014), yeast (Šibanc *et al*., 2018), collembola (Hohberg *et al*., 2015) and nematodes (Hohberg *et al*., 2015; Pilz and Hohberg, 2015) show significant shifts in abundance and diversity at natural CO_2 springs, especially towards acidophilic and anaerobic microorganisms (Krüger *et al*., 2011; Šibanc *et al*., 2014; Šibanc *et al*., 2018). This highlights the need for further characterisation of soil properties and plant-soil interactions at natural CO_2 springs in order to interpret plant responses to elevated

[CO_2] at these sites and relate them to plant response to elevated [CO_2] under climate change.

The potential for adaptation mediated by genetic change in plant populations exposed to elevated [CO_2] is not well understood at present. Although genetic variation in traits responsive to elevated [CO_2] has been evidenced in a wide range of plant taxa (Ziska and Bunce, 2000; Lindroth *et al*., 2001; Wieneke *et al*., 2004; De Costa *et al*., 2007; Nakamura *et al*., 2011) and this variation has been shown to be heritable in some studies (Schmid *et al*., 1996; Case *et al*., 1998), there remains significant debate over whether the strength of the elevated [CO_2] signal is sufficient to induce an evolutionary response. Studies that have utilised reciprocal transplant or crossed factored experimental designs with natural populations of plants growing around CO_2 springs have largely concluded that [CO_2] can act as a selective agent because of significant differences in traits of spring and control plants when grown in ambient versus elevated [CO_2] (Barnes *et al*., 1997; Polle *et al*., 2001; Onoda *et al*., 2009; Nakamura *et al*., 2011; Watson-Lazowski *et al*., 2016), though this finding is not universal (Van Loon *et al*., 2016). A natural extension of research utilising gradients and crossed factored experiments at natural CO_2 springs is to combine this approach with High Throughput Sequencing (HTS) tools to further elucidate the role of adaptation and plasticity in the multigenerational response (Watson-Lazowski *et al*., 2016). In addition epigenetic mechanisms have previously been highlighted as playing a role in coordinating plastic responses to elevated [CO_2] (May *et al*., 2013) and the potential contribution of epigenetics to transgenerational indirect genetic effects under elevated [CO_2] has not been explored, where natural CO_2 springs can be combined with HTS tools such as methylation sequencing to provide insight.

The use of natural CO_2 springs as a model for plant response to elevated [CO_2] has largely fallen out of favour in the past two decades because of concerns about variability of gas emission over time and contamination with exhaust gases such as hydrogen sulphide (H_2S) and sulphur dioxide (SO_2). As a result, increasing emphasis on quantifying potential contaminants in sites that are actively used for research with the exclusion of those that do not meet requirements, is evident in the literature (see Miglietta *et al*., 2012). In this meta-analysis we restricted the inclusion of data to springs with H_2S contamination below thresholds that could affect plant functioning and those with recorded SO_2 concentrations of below 0.015 ppm (Appendix Table A. 1). Although this threshold of [SO_2] exceeds the minimal concentration expected to affect plant growth (0.01 ppm), it is less than concentrations recorded in and around industrialised cities globally (De Kok, *et al*., 2007). As with potential ethylene contamination of industrial CO_2 in FACE sites, it is necessary to record and report these gas concentrations, both in initial site characterisation and over time to continually evaluate the suitability of the site as a model.

2.6 Conclusions

This first meta-analysis of long-term and multigenerational plant physiological responses to elevated [CO_2] at natural CO_2 springs has shown consistency in direction and magnitude with earlier observations in FACE, for eight traits related to gas exchange and physiology in a panel of nine traits. This suggests that predictions of plant response to rising [CO_2] from single generation FACE studies are robust over multiple generations in short-lived species and over long-term exposure in long-lived species, while highlighting that the role of ecological and evolutionary feedback in this response requires further investigation. This analysis supports the critical insights drawn from predictive models that incorporate empirical FACE data with relevance to food security, conservation and

ecosystem change under climate change. Dissecting whether multigenerational responses are solely plastic, have an epigenetic basis, and/or if adaptive genetic accommodation or assimilation occurs, will require reciprocal transplant and crossed factored experiments (Nakamura *et al.,* 2011; Watson-Lazowski *et al.,* 2016) which combined with newly accessible genomic technologies should provide crucial insight into the mechanistic basis of plant adaptation to elevated [CO_2] in the near future. Nevertheless, our results suggest that single generation experiments have provided robust insight of wide-ranging significance.

2.7 Acknowledgements

We thank the authors of these studies for their published work, which provide the data and foundation of knowledge for this meta-analysis. We are deeply grateful to those authors that provided additional data, answered our queries and gave advice. We also thank Fabrizia Ratto, Rebecca Spake and Zoe Harris for their invaluable guidance on meta-analyses, and anonymous reviewers for their constructive and insightful feedback. This work was funded by the National Environmental Research Council (NERC), through a SPITFIRE DTP PhD studentship to JMS [grant number NE/L002531/1], and through funding to GT from UC Davis and The British Council.

2.8 Author Contributions

J.M. Saban and G. Taylor conceived the study. J.M. Saban collected data and completed all statistical analyses. J.M. Saban and G. Taylor drafted the manuscript and M.A. Chapman commented on the manuscript.

2.9 References

Ainsworth, E. A., & Long, S. P. (2005) What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytologist*, **165**, 351-372.

Ainsworth, E. A., & Rogers, A. (2007) The response of photosynthesis and stomatal conductance to rising [CO_2]: mechanisms and environmental interactions. *Plant, Cell & Environment*, ***30***, 258-270.

Barnaby, J. Y., & Ziska, L. H. (2012) Plant Responses to Elevated CO_2. *eLS*, doi:10.1002/9780470015902.a0023718.

Barnes, J. D., Bettarini, I., Polle, A., Slee, N., Raines, C., Miglietta, F., Raschi, A., & Fordham, M. (1997) The impact of elevated CO_2 on growth and photosynthesis in *Agrostis canina* L. ssp. *monteluccii* adapted to contrasting atmospheric CO_2 concentrations. *Oecologia*, ***110***, 169-178.

Bloom, A. J., Burger, M., Asensio, J. S. R., & Cousins, A. B. (2010) Carbon dioxide enrichment inhibits nitrate assimilation in wheat and *Arabidopsis. Science*, **328**, 899-903.

Calcagno, V., & de Mazancourt, C. (2010) glmulti: an R package for easy automated model selection with (generalized) linear models. *Journal of statistical software*, **34**, 1-29.

Case, A., Curtis, P., & Snow, A. (1998) Heritable variation in stomatal responses to elevated CO_2 in wild radish, *Raphanus raphanistrum* (Brassicaceae). *American Journal of Botany*, **85**, 253-253.

Cotrufo, M. F., Raschi, A., Lanini, M., & Ineson, P. (1999) Decomposition and nutrient dynamics of Quercus pubescens leaf litter in a naturally enriched CO_2 Mediterranean ecosystem. *Functional Ecology*, **13**, 343–351.

Coûteaux, M. M., Kurz, C., Bottner, P., & Raschi, A. (1999) Influence of increased atmospheric CO_2 concentration on quality of plant material and litter decomposition. *Tree Physiology*, **19**, 301–311.

De Costa, W. A. J. M., Weerakoon, W. M. W., Chinthaka, K. G. R., Herath, H. M. L. K., & Abeywardena, R. M. I. (2007) Genotypic variation in the response of rice (*Oryza sativa* L.)

to increased atmospheric carbon dioxide and its physiological basis. *Journal of Agronomy and Crop Science*, **193**, 117-130.

De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., Wang, Y. P., Luo, Y., Jain, A.K., El-Masri, B., Hickler, T., & Wårlind, D. (2014) Where does the carbon go? A model–data intercomparison of vegetation carbon allocation and turnover processes at two temperate forest free-air CO_2 enrichment sites. *New Phytologist*, **203**, 883-899.

De Kok, L. J., Durenkamp, M., Yang, L., & Stulen, I. (2007) Atmospheric sulfur. In Hawkesford, M. J., & De Kok, L. J. (Ed.), Sulfur in plants an ecological perspective (pp. 17 91–106). Dordrecht, the Netherlands: Springer.

Dybzinski, R., Farrior, C. E., & Pacala, S. W. (2015) Increased forest carbon storage with increased atmospheric CO_2 despite nitrogen limitation: a game-theoretic allocation model for trees in competition for nitrogen and light. *Global Change Biology*, **21**, 1182-1196.

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R., Heimann, M., Thonicke, K., Zaehle, S., & Reichstein, M. (2016) Enhanced seasonal CO_2 exchange caused by amplified plant productivity in northern ecosystems. *Science*, **351**, 696-699.

Frenck, G., Linden, L., Mikkelsen, T. N., Brix, H., & Jørgensen, R. B. (2013) Response to multi-generational selection under elevated [CO_2] in two temperature regimes suggests enhanced carbon assimilation and increased reproductive output in *Brassica napus* L. *Ecology and evolution*, **3**, 1163-1172.

Frerichs, J., Oppermann, B. I., Gwosdz, S., Möller, I., Herrmann, M., & Krüger, M. (2013) Microbial community changes at a terrestrial volcanic CO_2 vent induced by soil acidification and anaerobic microhabitats within the soil column. *FEMS Microbiology Ecology*, **84**, 60–74.

Gahrooee, F. R. (1998) Impacts of elevated atmospheric CO_2 on litter quality, litter decomposability and nitrogen turnover rate of two oak species in a Mediterranean forest ecosystem. *Global Change Biology*, **4**, 667–677.

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 & Environment*, **41**, 1233–1246.

Grossman, J. D., & Rice, K. J. (2012) Evolution of root plasticity responses to variation in soil nutrient distribution and concentration. *Evolutionary Applications*, **5**, 850-857.

Haworth, M., Heath, J., & McElwain, J. C. (2010) Differences in the response sensitivity of stomatal index to atmospheric CO_2 among four genera of Cupressaceae conifers. *Annals of Botany*, **105**, 411-418.

Haworth, M., Killi, D., Materassi, A., & Raschi, A. (2015) Coordination of stomatal physiological behavior and morphology with carbon dioxide determines stomatal control. *American journal of botany*, **102**, 677-688.

Haworth, M., Hoshika, Y., & Killi, D. (2016) Has the impact of rising CO_2 on plants been exaggerated by meta-analysis of free air CO_2 enrichment studies? *Frontiers in Plant Science*, **7**, 1153.

Hedges, L. V., Gurevitch, J., & Curtis, P. S. (1999) The meta-analysis of response ratios in experimental ecology. *Ecology*, **80**, 1150-1156.

Higgins, J., & Thompson, S. G. (2002) Quantifying heterogeneity in a meta-analysis. *Statistics in medicine*, **21**, 1539-1558.

Hohberg, K., Schulz, H. J., Balkenhol, B., Pilz, M., Thomalla, A., Russell, D. J., & Pfanz, H. (2015). Soil faunal communities from mofette fields: Effects of high geogenic carbon dioxide concentration. *Soil Biology and Biochemistry*, **88**, 420–429.

IPCC. (2014) Summary for Policymakers. In: C. B. Field *et al.,* (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. Cambridge, United Kingdom, and New York, NY, USA: Cambridge University Press. pg 29.

Jablonski, L. M., Wang, X., & Curtis, P. S. (2002) Plant reproduction under elevated CO_2 conditions: a meta-analysis of reports on 79 crop and wild species. *New Phytologist*, **156**, 9-26.

Johnston, A., & Reekie, E. (2008) Regardless of whether rising atmospheric carbon dioxide levels increase air temperature, flowering phenology will be affected. *International Journal of Plant Sciences*, **169**, 1210-1218.

Jones, A. G., Scullion, J., Ostle, N., Levy, P. E., & Gwynn-Jones, D. (2014) Completing the FACE of elevated CO_2 research. *Environment international*, **73**, 252-258.

Krüger, M., Jones, D., Frerichs, J., Oppermann, B. I., West, J., Coombs, P., Green, K., Barlow, T., Lister, R., Shaw, R., Strutt, M., & Möller, I. (2011) Effects of elevated CO_2 concentrations on the vegetation and microbial populations at a terrestrial CO_2 vent at Laacher See, Germany. *International Journal of Greenhouse Gas Control*, **5**, 1093–1098.

Leakey, A. D., Ainsworth, E. A., Bernacchi, C. J., Rogers, A., Long, S. P., & Ort, D. R. (2009) Elevated CO_2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. *Journal of experimental botany*, **60**, 2859-2876.

Leakey, A. D., & Lau, J. A. (2012) Evolutionary context for understanding and manipulating plant responses to past, present and future atmospheric [CO_2]. *Philosophical Transactions of the Royal Society B*, **367**, 613-629.

Li, F. W., & Harkess, A. (2018) A guide to sequence your favorite plant genomes. *Applications in plant sciences*, **6**, e1030.

Lindroth, R. L., Roth, S., & Nordheim, E. V. (2001) Genotypic variation in response of quaking aspen (*Populus tremuloides*) to atmospheric CO_2 enrichment. *Oecologia*, **126**, 371-379.

Loladze, I. (2014) Hidden shift of the ionome of plants exposed to elevated CO_2 depletes minerals at the base of human nutrition. *Elife*, **3**, e02245.

Maček, I. (2013) A decade of research in mofette areas has given us new insights into adaptation of soil microorganisms to abiotic stress. *Acta agriculturae Slovenica*, **101**, 209-217.

Maček, I., Dumbrell, A. J., Nelson, M., Fitter, A. H., Vodnik, D., & Helgason, T. (2011) Local adaptation to soil hypoxia determines the structure of an arbuscular mycorrhizal fungal community in roots from natural CO_2 springs. *Applied and Environmental Microbiology*, **77**, 4770–4777.

Maček, I., Kastelec, D., & Vodnik, D. (2012) Root colonization with arbuscular mycorrhizal fungi and glomalin-related soil protein (GRSP) concentration in hypoxic soils in natural CO_2 springs. *Agricultural and Food Science*, **21**, 62–71.

May, P., Liao, W., Wu, Y., Shuai, B., McCombie, W. R., Zhang, M. Q., & Liu, Q. A. (2013) The effects of carbon dioxide and temperature on microRNA expression in *Arabidopsis* development. *Nature communications*, **4**, 2145.

Miglietta, F., Badiani, M., Bettarini, I., van Gardingen, P., Selvi, F., & Raschi, A. (2012) Natural CO_2 vents. *Global Change and Mediterranean-Type Ecosystems*, **117**, 102.

Moreton, J., Izquierdo, A., & Emes, R. D. (2016) Assembly, assessment, and availability of *de novo* generated eukaryotic transcriptomes. *Frontiers in genetics*, **6**, 361.

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., Carlisle, E., Dietterich, L. H., Fitzgerald, G., Hasegawa, T., & Holbrook, N. M. (2014) Increasing CO_2 threatens human nutrition. *Nature*, **510**, 139.

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 CO_2 concentration. *Oecologia*, **165**, 809-818.

Newton, P. C. D., Bell, C. C., & Clark, H. (1996) Carbon dioxide emissions from mineral springs in Northland and the potential of these sites for studying the effects of elevated carbon dioxide on pastures. *New Zealand Journal of Agricultural Research*, **39**, 33–40.

Onoda, Y., Hirose, T., & Hikosaka, K. (2007) Effect of elevated CO_2 levels on leaf starch, nitrogen and photosynthesis of plants growing at three natural CO_2 springs in Japan. *Ecological research*, **22**, 475-484.

Osada, N., Onoda, Y., & Hikosaka, K. (2010) Effects of atmospheric CO_2 concentration, irradiance, and soil nitrogen availability on leaf photosynthetic traits of *Polygonum sachalinense* around natural CO_2 springs in northern Japan. *Oecologia*, **164,** 41-52.

Pearson, P. N., & Palmer, M. R. (2000) Atmospheric carbon dioxide concentrations over the past 60 million years. *Nature*, **406**, 695.

Petit, J. R., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J. M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, M., Kotlyakov, V. M., Legrand, M., Lipenkov, V. Y., Lorius, C., Pepin, L., Ritz, C., Saltzman, E. & Stievenard, M. (1999) Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature*, **399**, 429-436.

Polle, A., McKee, I., & Blaschke, L. (2001) Altered physiological and growth responses to elevated [CO_2] in offspring from holm oak (*Quercus ilex* L.) mother trees with lifetime exposure to naturally elevated [CO_2]. *Plant, Cell & Environment*, ***24***, 1075-1083.

Pfanz, H., Vodnik, D., Wittmann, C., Aschan, G., Batic, F., Turk, B., & Macek, I. (2007) Photosynthetic performance (CO_2-compensation point, carboxylation efficiency, and net photosynthesis) of timothy grass (*Phleum pratense* L.*)* is affected by elevated carbon dioxide in post-volcanic mofette areas. *Environmental and Experimental Botany*, **61**, 41–48.

Pilz, M., & Hohberg, K. (2015) Laboratory experiments on nematodes from natural CO_2 springs reveal species-specific tolerance to extreme CO_2 concentrations. *Soil Biology and Biochemistry*, **88**, 165–168.

R-Core-Team. (2015) *R: A language and environment for statistical computing* [Online]. R Foundation for Statistical Computing, Vienna, Austria. Available: https://www.R-project.org/.

Rillig, M. C., Hernandez, G. Y., & Newton, P. C. D. (2000) Arbuscular mycorrhizae respond to elevated atmospheric CO_2 after long-term exposure: Evidence from a CO_2 spring in New Zealand supports the resource balance model. *Ecology Letters*, **3**, 475–478.

Rosenberg, M. S. (2005) The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution*, **59**, 464-468.

Ross, D. J., Tate, K. R., Newton, P. C. D., & Clark, H. (2002) Decomposability of C_3 and C_4 grass litter sampled under different concentrations of atmospheric carbon dioxide at a natural CO_2 spring. *Plant and Soil*, **240**, 275-286.

Ross, D. J., Tate, K. R., Newton, P. C. D., Wilde, R. H., & Clark, H. (2000) Carbon and nitrogen pools and mineralization in a grassland gley soil under elevated carbon dioxide at a natural CO_2 spring. *Global Change Biology*, **6**, 779-790.

Rubio-Asensio, J. S., & Bloom, A. J. (2016) Inorganic nitrogen form: A major player in wheat and *Arabidopsis* responses to elevated CO_2. *Journal of Experimental Botany*, **68**, 2611–2625.

Schmid, B. (1996) Genetic variation in the response of plant populations to elevated CO_2 in a nutrient-poor, calcareous grassland. *Carbon dioxide, populations, and communities*, 31-50.

Senior, A. M., Grueber, C. E., Kamiya, T., Lagisz, M., O'Dwyer, K., Santos, E. S., & Nakagawa, S. (2016) Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. *Ecology*, **97**, 3293-3299.

Šibanc, N., Dumbrell, A. J., Mandić-Mulec, I., & Maček, I. (2014) Impacts of naturally elevated soil CO_2 concentrations on communities of soil archaea and bacteria. *Soil Biology and Biochemistry*, **68**, 348–356.

Šibanc, N., Zalar, P., Schroers, H. J., Zajc, J., Pontes, A., Sampaio, J. P., & Macek, I. (2018) *Occultifur mephitis* fa, sp. nov. and other yeast species from hypoxic and elevated CO_2. *International Journal of Systematic and Evolutionary Microbiology*, **68**, 1-14.

Smith, W. K., Reed, S. C., Cleveland, C. C., Ballantyne, A. P., Anderegg, W. R., Wieder, W. R., Liu, Y.Y., & Running, S. W. (2016) Large divergence of satellite and Earth system model estimates of global terrestrial CO_2 fertilization. *Nature Climate Change*, **6**, 306.

Springer, C. J., & Ward, J. K. (2007) Flowering time and elevated atmospheric CO_2. *New Phytologist*, **176**, 243-255.

Stocker, T. F. (2013) The closing door of climate targets. *Science*, **339**, 280-282.

Tummers, B. (2006) *DataThief III* [Online]. Available: <http://datathief.org/>

Ueda, M. U., Onoda, Y., Kamiyama, C., & Hikosaka, K. (2017) Decades-long effects of high CO_2 concentration on soil nitrogen dynamics at a natural CO_2 spring. *Ecological research*, **32**, 215-225.

Van Loon, M. P., Rietkerk, M., Dekker, S. C., Hikosaka, K., Ueda, M. U., & Anten, N. P. (2016) Plant–plant interactions mediate the plastic and genotypic response of *Plantago asiatica* to CO_2: an experiment with plant populations from naturally high CO_2 areas. *Annals of botany*, ***117***, 1197-1207.

Vanuytrecht, E., & Thorburn, P. J. (2017) Responses to atmospheric CO_2 concentrations in crop simulation models: a review of current simple and semicomplex representations and options for model development. *Global Change Biology*, **23**, 1806-1820.

Videmšek, U., Hagn, A., Suhadolc, M., Radl, V., Knicker, H., Schloter, M., & Vodnik, D. (2009) Abundance and diversity of CO_2-fixing bacteria in grassland soils close to natural carbon dioxide springs. *Microbial ecology*, **58**, 1-9.

Viechtbauer, W. (2010) Conducting meta-analyses in R with the metafor package. *J Stat Softw*, **36**, 1-48.

Wang, D., Heckathorn, S. A., Wang, X., & Philpott, S. M. (2012) A meta-analysis of plant physiological and growth responses to temperature and elevated CO_2. *Oecologia*, **169**, 1-13.

Ward, J. K., Antonovics, J., Thomas, R. B., & Strain, B. R. (2000) Is atmospheric CO_2 a selective agent on model C_3 annuals? *Oecologia*, **123**, 330-341.

Watson-Lazowski, A., Lin, Y., Miglietta, F., Edwards, R. J., Chapman, M. A., & Taylor, G. (2016) Plant adaptation or acclimation to rising CO_2? Insight from first multigenerational RNA-Seq transcriptome. *Global Change Biology*, **22**, 3760-3773.

Wheeler, T., & Von Braun, J. (2013) Climate change impacts on global food security. *Science*, **341**, 508-513.

Wieneke, S., Prati, D., Brandl, R., Stöcklin, J., & Auge, H. (2004) Genetic variation *in Sanguisorba minor* after 6 years *in situ* selection under elevated CO_2. *Global Change Biology*, **10**, 1389-1401.

Woodward, F. I., & Kelly, C. K. (1995) The influence of CO_2 concentration on stomatal density. *New Phytologist*, **131**, 311-327.

Yan, W., Zhong, Y., & Shangguan, Z. (2017) Contrasting responses of leaf stomatal characteristics to climate change: a considerable challenge to predict carbon and water cycles. *Global Change Biology*, **23**, 3781-3793.

Ziska, L. H., & Bunce, J. A. (2000) Sensitivity of field-grown soybean to future atmospheric CO_2: selection for improved productivity in the 21st century. *Functional Plant Biology*, **27**, 979-984.

Chapter 3 Maternal and parental effects contribute to the multigenerational plant response to elevated [CO_2] at a natural CO_2 spring site

Jasmine M. Saban[a], Thomas H. G. Ezard[a, b], Mark A. Chapman[a] & Gail Taylor[a,c]

[a]Biological Sciences, University of Southampton, Life Sciences, Building 85, Highfield Campus, Southampton, SO17 1BJ, UK

[b]Ocean and Earth Science, National Oceanography Centre Southampton, Southampton, SO14 3ZH, UK

[c]Department of Plant Sciences, University of California, Davis, CA 95616, USA

In preparation for submission to *New Phytologist*

3.1 Abstract

- Parental and grandparental effects may be expected to evolve as part of the plant response to elevated CO_2 concentrations ($[CO_2]$) because of the predictability of seasonal $[CO_2]$ cycles and the continued rise in mean annual $[CO_2]$. The potential role of paternal and grandparental effects in particular have yet to be fully explored in plant response to elevated $[CO_2]$ and may play a role in facilitating adaptation.
- Here we conducted a multigenerational experiment using seed from a population of *Plantago lanceolata* L. growing at a natural $[CO_2]$ spring. We modelled components of phenotypic variation that were attributed to single and multigenerational elevated $[CO_2]$ exposure, heritability and parental or grandparental effects using a generalised linear mixed model (GLMM) incorporating pedigree information.
- Our data confirm that both maternal and paternal effects contribute to the multigenerational plant response to elevated $[CO_2]$ in a species with limited pollen and seed dispersal. We also find a persistent increased biomass in offspring of CO_2 spring plants, but the mechanism to explain this effect remains elusive.
- Our results demonstrate that paternal effects may play a significant role in plant response to elevated $[CO_2]$ and warrant further investigation as a mechanism to facilitate adaptation in future climates.

3.2 Introduction

Since plants are sessile organisms, the ability to sense and respond to environmental changes through the expression of adaptive phenotypic plasticity is a crucial determinant of their fitness (Nicotra *et al*., 2010; Gratani, 2014). Plants may also affect the expression of phenotypes of other individuals through indirect genetic effects, defined as the causal effect of an individual's genotype or phenotype on the phenotype of an individual in the same species (Wolf *et al*., 1998). The best studied example of indirect genetic effects are maternal effects, which may be adaptive in providing a rapid and flexible mechanism for transgenerational adjustment of offspring phenotype to environmental conditions experienced by the mother (Uller, 2008). Other transgenerational indirect genetic effects such as paternal and grandparental effects can also evolve but have been less comprehensively studied.

Extending theoretical predictions, transgenerational indirect genetic effects might be expected to evolve in plant responses to increasing atmospheric [CO_2] under climate change, because seasonal cycles in [CO_2] are predictable and relatively slow, and because annual changes in [CO_2] are generally directional and occur with consistent spatial distribution (Mousseau and Fox, 1998; Galloway, 2005; Galloway and Etterson, 2007; Marshall and Uller, 2007; Ezard *et al*., 2014; Tans and Keeling, 2018). Empirical data largely supports this prediction, but only provides evidence for a role of maternal effects in facilitating plant adaptation to elevated [CO_2] (Jablonski *et al*., 2002; Poorter and Navas, 2003), while paternal and grandparental effects have not been considered. Addressing this gap in knowledge provides an opportunity to both better understand the mechanisms of adaptation to elevated [CO_2], and to test theoretical predictions of scenarios in which paternal and grandparental effects may be expected to evolve. Partitioning sources of variance to indirect genetic effects in the plant response to elevated [CO_2] also facilitates the study of adaptation to elevated [CO_2] by direct genetic effects. Understanding the mechanisms of adaptation to elevated [CO_2] is key to predicting ecosystem change, biosphere processes and food security under climate change (Myers *et al*., 2014; Becklin *et al*., 2017).

In single generation exposure experiments, elevated [CO_2] increases photosynthetic rate in plants and this can further affect plant functioning by, for example, increasing plant

water and nitrogen use efficiency and increasing biomass and yield (Leakey *et al*., 2009; Reddy *et al*., 2010; Prior *et al*., 2011). The contribution of maternal effects to the multigenerational response to elevated [CO_2] was first highlighted by Wulff and Alexander (1985) in *Plantago lanceolata*, with evidence that growth at elevated [CO_2], decreased seed weight but increased germination percentage relative to those grown under ambient [CO_2]. Theoretical and empirical work has since demonstrated that maternal effects can facilitate adaptation, and that although both maternal and offspring fitness are affected, selection largely acts to maximise maternal fitness (Mousseau and Fox, 1998; Galloway, 2005; Marshall and Uller, 2007). Maternal effects are now well documented in the plant response to elevated [CO_2], evidenced through altered seed size, weight and number (Jablonski et al, 2002), as well as altered germination rate (Marty and BassiriRad, 2014), flowering time (Springer and Ward, 2007) and seed resource composition (Grünzweig and Dumbur, 2012; Myers *et al*., 2014), with considerable variation in direction and magnitude of the effect between species.

Paternal effects are generally weaker than maternal effects in plants because their effect on offspring is prezygotic only, whereas maternal plants can influence offspring in postzygotic stages (Lacey, 1996; Etterson and Galloway, 2002). In plants, offspring phenotype can be influenced by paternal environment including light quality and nutrient levels (Galloway, 2001), temperature (Lacey, 1996), soil type (Schmid and Dolt, 1994), arbuscular mycorrhizal fungi status (Varga and Soulsbury, 2017) and water status (Diggle *et al*., 2010). Paternal effects are more likely to evolve in species where there is limited pollen and seed dispersal, and where this distance is shorter than the distance to environmental heterogeneity, such that the paternal as well as maternal environment is predictive of the offspring environment (Galloway, 2005; English *et al*., 2015). The potential for paternal effects to play a role in plant adaptation to elevated [CO_2] have been comparatively neglected.

Grandparental effects have never been quantified in plant response to elevated [CO_2] but grandparental temperature (Alexander and Wulff, 1985; Case *et al*., 1996; Whittle *et al*., 1999; Groot *et al*., 2017), drought stress (Herman *et al*., 2012), nutrient environment (Wulff *et al*., 1999; Kou *et al*., 2011), and salt stress (Groot *et al*., 2016) have all been shown to influence progeny phenotypes in experimental and natural settings. Grandparental effects are expected to be weaker than parental effects because they

cannot directly influence offspring phenotype and because their predictive power is likely diminished by the increased temporal separation compared to parental environment (Leimar and McNamara, 2015; Groot *et al*., 2017). Increasing [CO_2] as measured across multiple generations may provide a more reliable prediction of offspring [CO_2] environment (Herman *et al*., 2014) but grandparental effects would be expected to be weaker than parental effects and therefore difficult to detect. Since paternal effects, and to a greater extent grandparental effects do not have direct influence on the environment experienced by progeny, epigenetic inheritance has been highlighted as a potentially key mechanism by which these (as well as maternal) effects could influence offspring phenotype (Curley *et al*., 2011; Rando, 2012; Holeski *et al*., 2012; Baker *et al*., 2018).

Direct genetic effects could also contribute to plant adaptation to elevated [CO_2] over many generations. Indeed genetic adaptation may become increasingly important in determining population responses to elevated [CO_2] as photosynthetic rates reach [CO_2] saturation, which is predicted to occur at around 1000 ppm in those plants without a carbon concentrating mechanism (85% of higher plant species) (Yamori et al, 2014; Ward and Strain, 1997; Zheng *et al*., 2018). This is akin to the concentrations predicted for the end of the century under worst case emissions scenarios (IPCC, 2014). Decreased responsiveness of plants to future elevated [CO_2], relative to present day [CO_2] and the lower [CO_2] of the past so far supports the prediction of reduced responsivity (Dippery *et al*., 1995; Ward and Strain, 1997; Temme *et al*., 2015). The potential for plasticity to be exhausted by these conditions, along with the identification of heritable genetic variation that is responsive to elevated [CO_2] (Schmid, 1996; Case *et al*., 1998) supports a role for genetic adaptation in plant response to elevated [CO_2] under climate change scenarios. Yet, there is only limited and controversial evidence that elevated [CO_2] is a strong enough selection pressure to induce population level genotypic changes under natural selection (Leakey and Lau, 2012). Even in selection experiments, evidence that elevated [CO_2] can drive genetic adaptation is variable and limited (Potvin & Tousignant 1996; Ward *et al*. 2000; Wieneke *et al*., 2004; Teng *et al*., 2009; Frenck *et al*., 2013; Grossman and Rice, 2014).

Experiments to study the role of adaptation in multigenerational plant responses to elevated [CO_2] are challenging because of the considerable cost of growing plants under CO_2 enrichment, and by the time required to generate multiple generations (Paoletti *et*

al., 2005). To circumvent these obstacles, plants growing at natural CO_2 springs have been used to study physiological plant responses to multiple generations of growth at elevated [CO_2] for the last 30 years or more, and provide populations of plants in the natural environment that have been exposed to elevated [CO_2] for potentially hundreds of generations.

Here, we use the perennial herb *P. lanceolata* at the Bossoleto natural CO_2 spring in Italy to study components of plant adaptation to elevated [CO_2]. *P. lanceolata* is an ecologically and commercially important species (Stewart, 1996; Navarette *et al*., 2016) and has been independently studied in controlled environment chambers allowing comparison of [CO_2] response across experimental designs (Bettarini *et al*., 1998, Norton *et al*., 1999, Hui *et al*., 2002, Watson-Lazowski *et al*., 2016). Parental and grandparental effects have also previously been identified as contributing to *P. lanceolata* response to other environmental factors (Case *et al*., 1996; Latzel and Klimešová 2009; Latzel *et al*., 2009; Latzel *et al*., 2010; Latzel *et al*., 2014). As an obligate outcrosser with a limited pollen (1.5 m) and seed (0.08 m) dispersal distance (Bos *et al*., 1986; Tonsor *et al*., 1993), we show that both maternal and paternal effects (but no evidence of grandparental effects) facilitate the response of *P. lanceolata* to elevated [CO_2] at the Bossoleto spring. In partitioning components of phenotypic variation associated with indirect genetic effects, heritability of traits and the effect of single and multigenerational [CO_2] exposure, we are also able to assess evidence for genetic adaptation to elevated [CO_2] in this species.

3.3 Materials and Methods

3.3.1 Natural CO_2 spring site

Seed was collected from *Plantago lanceolata* L. plants at the field site on 2nd September 2014. *P. lanceolata* is a wind-pollinated, basal rosette forming perennial herb, with a gynodioecious breeding system. The field site is located in Bossoleto, Tuscany, Italy (Lat. 43° 17', Long. 11° 35') and has been previously described in detail (Van Gardingen *et al*., 1997). Briefly, the elevated [CO_2] spring site is an area of calcareous grassland at the top of a basin-shaped depression, at the bottom of which several vents emit high concentration dry CO_2. The daytime concentration of CO_2 measured at this site is in the range of 400-1200 ppm with an average of 1000 ppm (Scholefield *et al*., 2004). The emissions are low in pollutants such as H_2S and SO_2 (Schulte *et al*., 2002; Herschbach *et al*., 2010). The Bossoleto natural CO_2 spring is anecdotally more than 100 years old and has been used for ecological research since at least 1998, when the population of *P. lanceolata* was first documented (Bettarini *et al.,* 1998). The age and isolation, as well as the limited pollen and seed dispersal range of this species (Bos *et al*., 1986; Tonsor *et al*., 1993), predicts that transgenerational indirect genetic effects will evolve in this system (Galloway, 2005; Marshall and Uller, 2007; Ezard *et al*., 2013), and also gives confidence that these populations have experienced elevated [CO_2] for many generations of growth. In order to draw conclusions about the response of *P. lanceolata* to elevated [CO_2] at the natural CO_2 spring site, plants growing at an established site approximately 120 m away with comparable flora, soil properties and climatic conditions but ambient [CO_2] were used as a control in experimental manipulations (Körner and Miglietta, 1994; Van Gardingen *et al*., 1997).

3.3.2 Experimental design

In order to quantify the contribution of heritability, parental and grandparental effects to trait variation in *Plantago lanceolata* we grew plants from the Bossoleto sites in the same environment (ambient [CO_2] growth chambers) for multiple generations. To then assess the extent to which *P. lanceolata* plastically responds to elevated [CO_2] and whether there were any fixed (potentially adaptive) components to spring plant phenotype that could result from adaptation to elevated [CO_2] over many generations of growth at the

spring, we grew plants in either ambient or elevated [CO_2] as a third generation (Figure 3.1). Growth of plants from both sites at ambient [CO_2] prior to the crossed factored generation reduced the potential for estimates of adaptation to elevated [CO_2] in the spring plants to be confounded by parental effects. It was therefore not necessary to grow plants in generations one and two under elevated [CO_2] to either assess the contribution of heritability and parental effects to phenotype expression, or to quantify plastic or adaptive responses to elevated [CO_2].

Seed from seven maternal plants from the spring and seven maternal plants from the control site were grown for two generations at ambient [CO_2] in environmental control rooms at the University of Southampton, UK (25 °C day/18 °C night, 16 hour day 06:00-22:00 with PAR 400-500 µmol m^2 s^{-1} and 60 % humidity). Controlled reciprocal crosses of paired plants were used to generate seed for each subsequent generation. For the third generation, seed produced in generation two was sown in a crossed factored design in controlled CO_2 environment chambers, such that seed from plants originally from the spring and control sites could be grown at ambient or elevated [CO_2] (Figure 3.1). In this generation, seed were germinated and grown under one of two treatment CO_2 concentrations; ambient with a median (interquartile range) [CO_2] of 366 ppm (344 - 389 ppm) or elevated [CO_2] of 890 ppm (785 – 994 ppm), with 25 °C day/19 °C night, 16 hour day 06:00-22:00 with PAR 400-500 µmol m^2 s^{-1} and ~60 % humidity. After two weeks temperature was increased to 28 °C day/19 °C night to simulate conditions at the Bossoleto site during the summer months. In all generations plants were grown in 10 cm diameter pots in trays. Plants were randomised within and between trays every two days, and for generation three trays were swapped between all eight chambers (with the treatments swapped where necessary) every three days to minimise chamber effects. The number of generations analysed in this experiment were constrained by the generation time of *P. lanceolata* in these conditions, which was approximately nine months seed to seed.

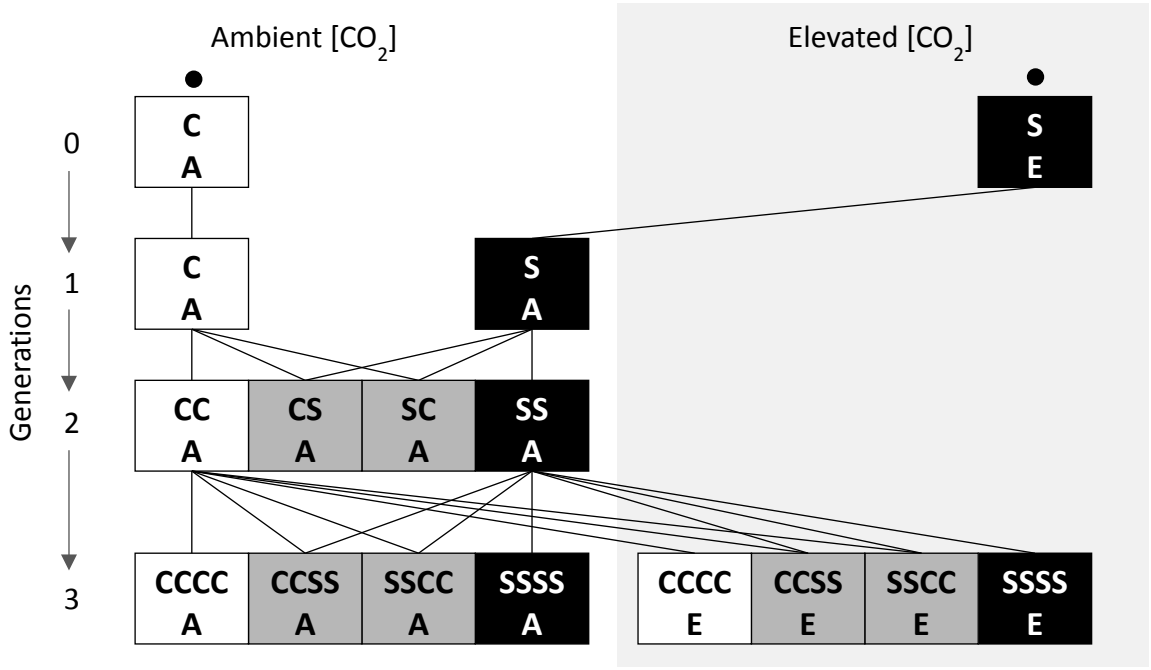


Figure 3.1: Schematic overview of this experiment.

Seed were obtained from plant populations growing in naturally elevated [CO_2] at the Bossoleto spring (S) and from plants at ambient [CO_2] in the nearby control site (C) as generation zero. Black dots indicate that this generation growing *in situ* is expected to be acclimated to their environmental [CO_2]. These seeds were grown for two generations at ambient [CO_2] in environmentally controlled rooms, and in generation one and two plants were crossed to also produce hybrid plants. In generation three plants were grown at ambient or elevated [CO_2]. Data was collected from sample sizes n=120, 194 and 220 for generations one, two and three respectively.

3.3.3 Phenotypic measurements

We collected phenotypic measurements at the same developmental stage in all three experimental generations. In generation three the increased [CO_2] in the elevated [CO_2] treatment and 3 °C increase in temperature across all treatments meant that the vegetative to floral transition occurred in around half the time of the previous generations and we adjusted data collection dates accordingly with the assumption (informed by preliminary experiments) that all aspects of phenology are accelerated equally. Abaxial stomatal conductance was measured on the fourth oldest leaf 5-7 hours after the start of the light period in LD conditions using a porometer (AP4, Delta-T Devices, Cambridge, UK) at equivalent development stages in all generations. In generation one and two this was day 77 after germination, while for generation 3 this was day 39 due to the accelerated phenology associated with growth conditions. The oldest

leaf was taken for leaf area and dry weight measurements the following day with the third oldest leaf taken for adaxial and abaxial epidermal imprints. Single leaf area was measured using ImageJ (Image J 1.42q; Wayne Rasband, Bethesda, MA, USA) from images taken at the time of harvest with a 30 cm scale bar. Specific leaf area ($cm^2\ g^{-1}$) was calculated as the ratio of single leaf to single leaf dry weight for this leaf. Abaxial imprints were taken using nail varnish painted on the surface of the leaf from half way down the length of the leaf (Potvin and Tousignant, 1996). Five images were taken across the length of this imprint and stomatal and epidermal numbers were counted in a cropped field of view of 1.6 mm^2 at 10x magnification. Epidermal cell area was averaged from 25 epidermal cells per imprint, and guard cell length was averaged from 25 stomata per imprint. Stomatal density (SD) and stomatal index (SI) were calculated using standard equations (Taylor *et al*., 2003). The number of rosettes were counted on day 102 after germination in generations one and two, and day 51 following germination in generation three.

Mature seeds were collected at monthly intervals in generations one and two and every 10 days in generation three. Destructive biomass samples were taken only in generation three, at 71 days after germination. Plant biomass was partitioned by cutting at 5 mm above ground. Roots were washed to remove soil and dried at 105 °C for 48 hours. Above ground biomass was further partitioned into reproductive and leaf biomass. Flower heads were processed to obtain seeds and these were weighed for fresh seed weight, and then dried with the remaining reproductive biomass at 105 °C for 48 hours. The number of leaves of each plant at harvest were counted before leaf biomass was dried at 105 °C for 48 hours.

3.3.4 *In situ* data collection

Stomatal conductance was measured on a mature leaf between 11 am and 1 pm (5.5-7.5 hours after sunrise) using a porometer (AP4, Delta-T Devices, Cambridge, UK). Leaf area images and stomatal imprints were taken from mature leaves of 20 individuals in each of the spring and control sites. The trait value means for these plants are described as '*in situ*' rather than generation zero and were not used in multigenerational statistics but rather to visualise trends in the spring and control sites for comparison to this experiment.

3.3.5 Statistics

We partitioned variation in plant reproductive, fitness and stomatal patterning traits by fitting generalised linear mixed models (GLMM) using MCMCglmm (Hadfield, 2010) in R v3.4.1 (R Core Team, 2013). MCMCglmm uses a Bayesian framework with Markov chain Monte Carlo simulations. This approach has greater power than parent offspring regression because more relatedness information can be incorporated into models through the construction of a pedigree (Villemereuil *et al*., 2012). The mixed models fitted in MCMCglmm are 'animal models' which explicitly take into account a pedigree structure of the experimental population or a phylogeny to model phenotype as a function of random genetic, parental and grandparental random effects as well as fixed effects that are the core of experimental manipulations (Henderson, 1976, Hadfield, 2010). The pedigree here contains 452 individuals with 120 distinct maternal plants and 105 paternal plants with a depth of four generations.

In this model three main fixed effects were defined. To model adaptation to elevated [CO_2] at the natural CO_2 spring we incorporated ancestral site (population) as a factor with three levels; spring, control or hybrid. [CO_2] treatment (either ambient or elevated) was defined as a second fixed effect as a measure of plastic within-generation response to [CO_2]. A third fixed effect "Generation" (generational environment) was used to account for differences in temperature and other variables between generations and was fitted as a four level factor. Generation was included as a fixed effect so that we could estimate random effects that were marginal compared to the large inter-annual differences in trait values, and this meant that traits were estimated relative to the mean for that generation (McFarlane *et al*., 2014). There were 3-8 replications per cross but the inclusion of this fixed effect likely meant that our estimation of heritability and indirect genetic effects were conservative, and we additionally modelled traits without this effect to quantify its impact (Appendix Table B. 1 & B. 2). We tested the interaction of population and [CO_2] treatment for all traits (Appendix Table B. 3 & B. 4) and have included them in presented models only where they had a significant effect on trait variation (days to first spike, number of rosettes and seeds per spike).

We fitted univariate animal models for each trait with the fixed effects outlined above and three (heritability and parental effects) or five (heritability, parental and

grandparental effects) random effects. To calculate heritability we estimated additive genetic variance (σ^2_a) of traits measured. In addition we estimated the component of variation in phenotype attributed to the identity of each parent (maternal effects σ^2_{mat} and paternal effects σ^2_{pat}) (Kruuk *et al*., 2008). We calculated grandparental effects for two traits, as the component of phenotypic variation attributed to grandparental identity (maternal grandmaternal effects σ^2_{grmat} and paternal grandpaternal effects σ^2_{grpat}). Total phenotypic variance was calculated as the sum of all variance components and narrow sense heritability, parental and grandparental effects were calculated as the ratio of that variance component to the total phenotypic variance for that trait. For example heritability for the trait was calculated as (except for the two traits where we also calculated grandparental effects, which had σ^2_{grpat} and σ^2_{grmat} as additional components on the denominator):

$$h^2 = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2_{mat} + \sigma^2_{pat} + \sigma^2_r}$$

where σ^2_r is the residual variance.

For all but three traits (with number of rosettes, number of seeds per spike and seed weight the exceptions discussed later) we log transformed the response variable to approximate a normal distribution and used a Gaussian error distribution. For these traits we used parameter expanded priors for random effects (V=1, ν=1, α.μ=0, α.V=1000) with inverse gamma priors (V=1, ν=0.002) for fixed effects, and estimates were robust to changes in prior specification (Appendix Table B. 5 & B. 6). For three traits we categorised the response variable and used an ordinal error distribution to model the response with parameter expanded priors for random effects and V=1 with residual variance fixed at 1 for fixed effects. For the trait 'number of rosettes' a large proportion of plants had one rosette so we designated response to one of four categories representing one rosette, two or three rosettes, four or five rosettes, or six or more rosettes (see Appendix Figure B. 1 for impact of discretization on trait distribution). Number of seeds per spike and seed weight were recorded as an average of the maternal plant. Each individual was assigned to one of three categories, based on whether the response of their maternal plant was in the bottom third of the range of values for each generation (low light), the middle third (medium) or the top third (high/heavy). We also tested the effect of categorising based on trait values below a standard deviation from the mean within each generation, within

a standard deviation from the mean or more than a standard deviation above the mean, and found this categorisation did not significantly affect the estimation of effects (Appendix Table B. 7 & B. 8). We additionally modelled the random effects using inverse gamma priors and this did not significantly impact the estimates (Appendix Table B. 9).

We determined if models had adequate mixing by examining trace plots of sampled posterior distributions for the mean and variance components, and additionally ensuring that levels of autocorrelation were <0.1 and effective size was >2000. Chain convergence of all models was examined with Gelman-Rubin-Brooks plots (Brooks and Gelman, 1998) and quantified using Heidelberger and Welch's convergence diagnostic (Heidelberger and Welch, 1983). We ran all models with 5,000,000 iterations, discarding the first 1,000,000 iterations and subsampling every 1000. All statistical analyses were conducted in R v 3.4.1 (R Core Team, 2013).

For destructive biomass measurements in the third generation, all models were fitted using the *glm* function in "lme4" (Bates *et al*., 2015) in R version 3.4.2 (R Core Team, 2013) with [CO_2] treatment and population site of origin (spring, control or hybrid) fitted as fixed effects. Normality was assessed by examination of Q-Q plots (Appendix Figure B. 2) and model fit was assessed by testing if the residual deviance fitted a chi-squared distribution. Linear models with Gaussian error distribution were fitted for leaf, root and total biomass. Since some plants did not reach the vegetative to floral transition in generation three we modelled reproductive biomass in two parts. Firstly, we modelled the weight of plants that did reach the floral transition with a glm with inverse Gaussian distribution and link function $1/mu^2$ and then back transformed coefficients using the R package "emmeans" (Lenth, 2018) to facilitate interpretation. Secondly, we modelled whether or not the plants reached the transition by date of harvest using a binomial error distribution (Appendix Table B. 10).

To identify correlations between traits, Spearman's rank correlation coefficient was calculated in R v3.4.1 base packages (R Core Team, 2013). Trait correlation in generation three are visualised in Appendix Figure B. 3 using R package "ggcorrplot" (Kassambara, 2016).

3.4 Results

3.4.1 Heritability estimates, parental and grandparental effects in *P. lanceolata*

Estimates of the amount of phenotypic variation explained by narrow-sense heritability, parental effects and grandparental effects are visualised in Figure 3.2. Broadly, this demonstrates that both maternal and paternal effects contributed to the variation in specific reproductive traits, while grandparental effects did not significantly contribute to variation in the two traits for which they were estimated.

Two reproductive traits were significantly impacted by maternal effects; offspring seed weight and number of seeds per spike (Figure 3.2). Maternal effects explained 66.4 % of the variation in seed weight (95 % Credible Intervals (CI) [53, 95]) and 55.6 % of the variation in number of seeds per spike (95 % CI [47, 88]). There was no evidence of a trade-off between seed weight and seed number, as a lack of correlation between the two traits ($\rho(192)=0.03$, ns). However, without data on the total number of seeds produced or the number of spikes produced it is impossible to discern if there was a trade-off between the total number of seeds produced per plant and seed weight.

Paternal effects contributed significantly to the variation of only one trait, with 68.7 % (95 % CI [48, 85]) of the variation in days to first spike accounted for by paternal effects (Figure 3.2). Parental effects did not significantly contribute to the variation in days to germination but across generation two and three (where both traits were measured) there was a positive correlation between days to germination and days to first spike ($\rho(398)=0.21$, $P<0.05$). This suggests that the paternal effect on date to vegetative to floral transition could be mediated in part through altered days to germination.

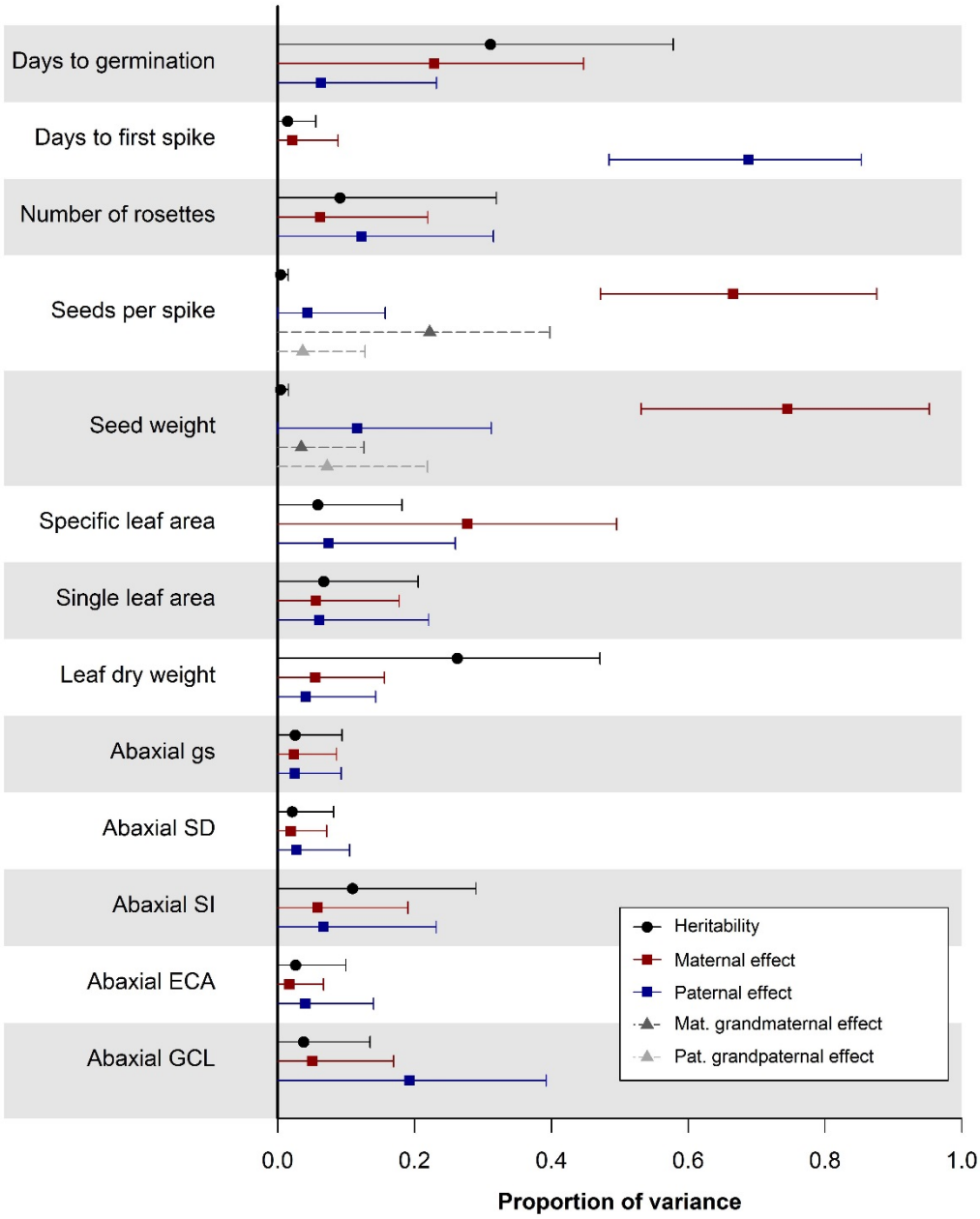


Figure 3.2: Estimates of heritability and transgenerational indirect genetic effects.

Estimates of heritability ($h^2 = \sigma^2_a/\sigma^2_p$), maternal effects ($\sigma^2_{mat}/\sigma^2_p$) and paternal effects ($\sigma^2_{pat}/\sigma^2_p$) for all traits, and maternal grandmaternal ($\sigma^2_{grmat}/\sigma^2_p$) and paternal grandpaternal ($\sigma^2_{grpat}/\sigma^2_p$) effects for number of seeds per spike and seed weight. Error bars are 95% Bayesian credible intervals. Abbreviations; SD – stomatal density, SI - stomatal index, ECA - epidermal cell area, GCL - guard cell length.

The fixed effect of generation was included for all multigenerational models and was a significant predictor of the estimate in 11 of the 13 traits, the two exceptions being seed weight and single leaf dry weight (Appendix Table B.11). This highlights that generation-

specific environmental variables significantly impacted the expression of plant traits and this can be seen in plots of the raw data (Appendix Figure B.4 and Appendix Figure B.5). When modelled without generation as a fixed effect, a paternal effect in days to germination, and maternal effects in seed weight and number of seeds per spike were still identified, but an additional paternal effect was estimated to affect the expression of abaxial SD (Appendix Table B.1).

Narrow-sense heritability was low to moderate for all traits with point estimates of random effects robust to the specification of alternate priors (parameter expanded half Couchy and inverse gamma, Appendix Table B.5 and B.6). Point estimates for heritability explained between 1.4 % and 31.1 % of the variance across the traits, but no trait's expression depended significantly on heritable variation (Figure 3.2).

3.4.2 Plasticity and adaptation to elevated [CO_2] in *P. lanceolata* reproductive traits

For simplicity, the effect of growth [CO_2] and population on the expression of traits discussed here are visualised as reaction norm plots in the third generation (Figure 3.3 and Figure 3.4). However, MCMC GLMMs used data for all generations to quantify the impact of fixed effects, and Figures 3.3 and 3.4 also only provide *P*-values for a subset of fixed effects included in the models. For plots of raw trait values across generations and full model outputs see Appendix Figure B.4, Appendix Figure B.5 and Appendix Table B.11.

In all five measured reproductive traits there was no effect of population on trait expression, suggesting that there are no fixed differences in the expression of these traits induced by multiple generations of growth at elevated [CO_2] (Figure 3.3, Appendix Table B.11). Although a plastic response to elevated [CO_2] was not statistically significant at the $P<0.05$ threshold for these traits, trends ($P<0.1$) were seen in three of five reproductive traits (days to germination, days to first spike and number of seeds produced per spike). Plants grown in elevated [CO_2] tended to germinate earlier and flower earlier (Figure 3.3). Since paternal effects contribute to the variation in days to first spike, the plastic response of days to first spike to elevated [CO_2] could contribute to multigenerational adaptation to elevated [CO_2] through this paternal effect. There was a significant interaction effect of growth [CO_2] and population on the number of seeds per spike (Figure 3.3, Appendix Table B.11), suggesting some local adaptation in CO_2 spring versus

control site populations. Thus even if the maternal effect on number of seeds per spike contributes to adaptation to elevated [CO_2], other genetic or epigenetic differences between the two populations likely also contribute to the differential expression of this trait in ambient and elevated [CO_2].

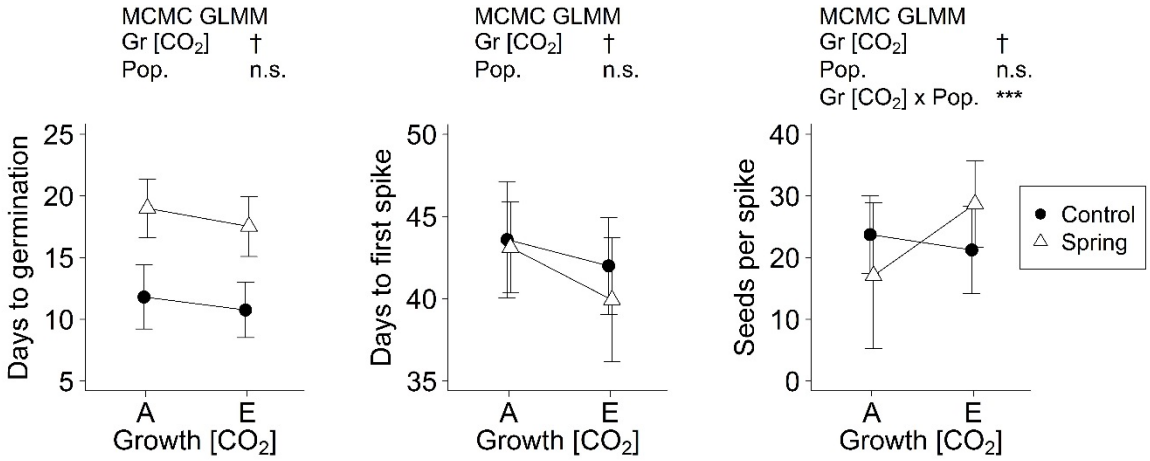


Figure 3.3: Reaction norm plots of fitness traits in *Plantago lanceolata* grown in ambient and elevated [CO_2] in generation three, for traits where population or growth [CO_2] had an effect on trait variation (*P*<0.1).

Labels on the x-axis denote growth [CO_2] (A - ambient, E - elevated). Progeny of plants originating from the spring (S) or control (C) site populations are identified by a white triangle or black circle respectively. Trait responses are given as mean with 95 % confidence intervals. P-value significance for the effect of growth [CO_2] (Gr [CO_2]) and population site of origin (Pop.) are indicated, these represent only a subset of the fixed effects included in MCMC GLMMs and full model outputs can be found in Appendix Table B.11. P-value significance indicated as: *, P < 0.05; **, P < 0.01; ***, P < 0.001; †, P < 0.1.

3.4.3 Plasticity and adaptation to elevated [CO_2] in *P. lanceolata* leaf traits

Across the three leaf-size related traits there was evidence of both a plastic response to elevated [CO_2] and of fixed differences in the expression of traits between plants derived from spring versus control site populations (Figure 3.4, Appendix Table B.11). Figure 3.5 visualises the reaction norms for these traits in ambient and elevated [CO_2] in the third generation. Growth [CO_2] had a significant effect on single leaf dry weight (*P*<0.01) and an effect at *P*<0.1 on single leaf area, where on average in generation three, single leaf dry weight was 16.2 % lower in elevated [CO_2] and single leaf area was 5.8 % higher (Figure 3.4a, Appendix Table B.11). Although there was no significant effects of population on either single leaf area or single leaf dry weight, specific leaf area was significantly affected by population site of origin (*P* <0.001), with spring plants having a 28.6 % higher specific

leaf area in generation three (Figure 3.5). Specific leaf area was also higher in hybrids relative to those from the control site population (Appendix Table B.11). Although specific leaf area was higher in spring plants, there was a trend towards lower specific leaf area (P<0.1) under elevated growth [CO_2], likely a result of a larger effect of elevated [CO_2] on single leaf dry weight than on single leaf area. This suggests that multigenerational response in spring plant specific leaf area occurs in the opposite direction to the plastic effect of elevated relative to ambient [CO_2] on specific leaf area (Figure 3.5).

A significant 37.4 % reduction of abaxial stomatal conductance when plants were grown under elevated [CO_2] suggested reduced stomatal opening (Fig. 5, Appendix Table B.11). Stomatal patterning traits were largely unaffected by single or multigenerational exposure to elevated [CO_2] with only a trend towards decreased abaxial SD in spring derived plants noted (Figure 3.4b).

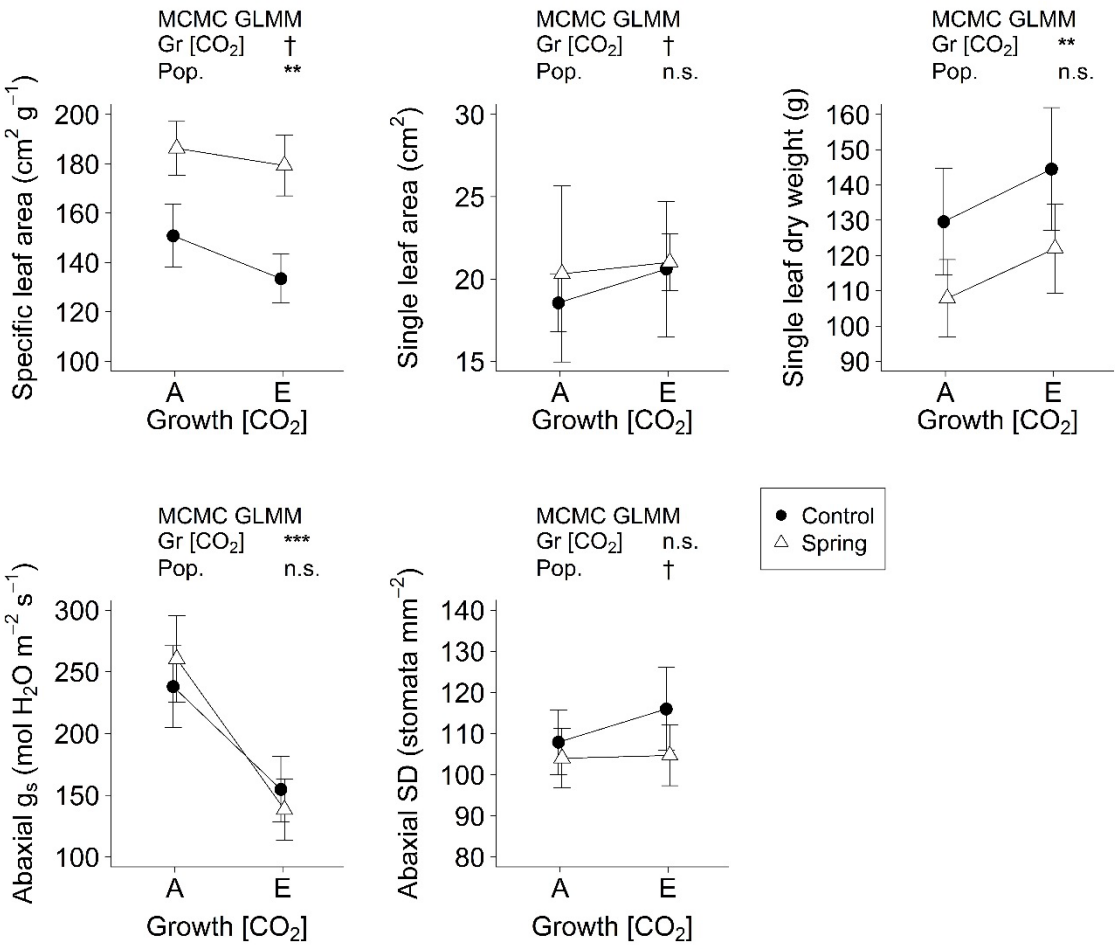


Figure 3.4: Reaction norms plots of leaf-size (a) and leaf functional (b) traits in *Plantago lanceolata* grown in ambient and elevated [CO_2] in generation three, for traits where population or growth [CO_2] had an effect on trait variation (P<0.1).

Labels on the x-axis denote growth [CO_2] (A - ambient, E - elevated). Trait responses are given as mean with 95 % confidence intervals. P-value significance for the effect of growth [CO_2] (Gr [CO_2]) and population site of origin (Pop.) are indicated, these represent only a subset of the fixed effects assessed in MCMC GLMMs and full model outputs can be found in Appendix Table B.11. P-value significance indicated as: *, P < 0.05; **, P < 0.01; ***, P < 0.001; †, P < 0.1.

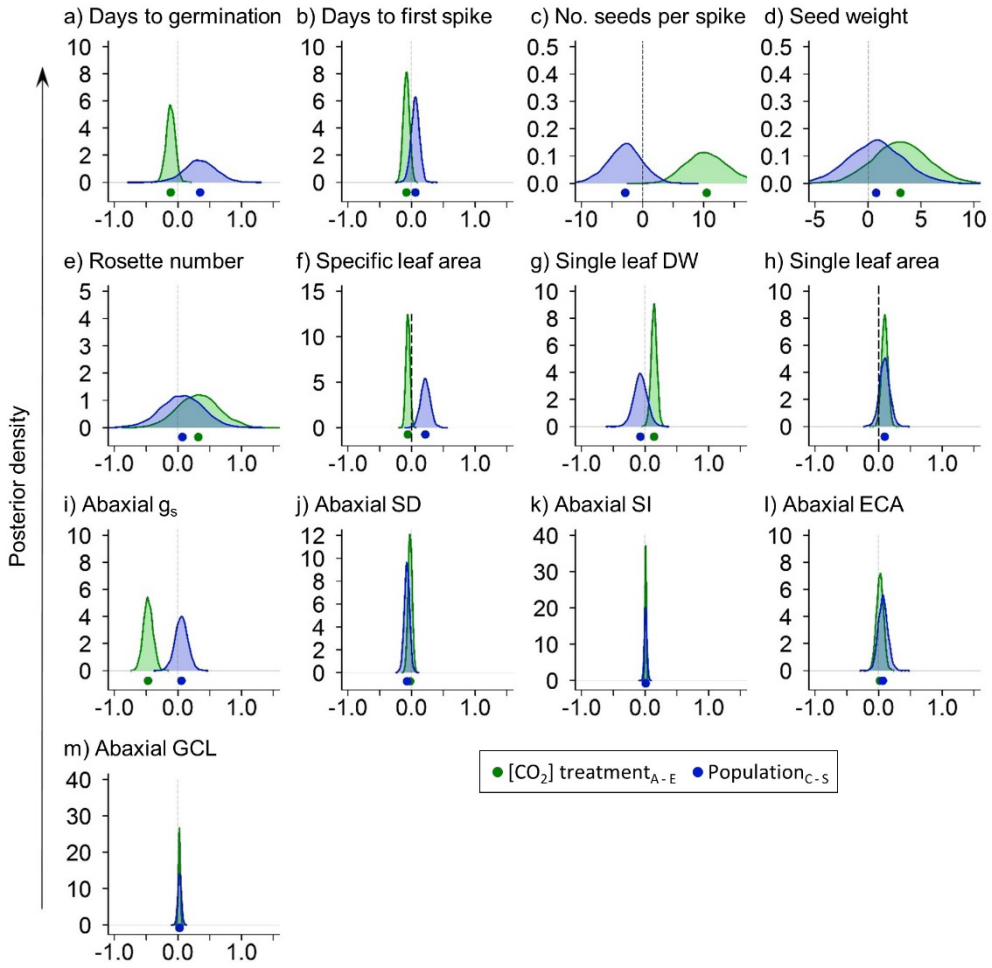


Figure 3.5: Posterior distributions of fixed effects [CO_2] treatment $_{AMB\ -\ ELEV}$ (blue) and Population $_{CONTROL\ -\ SPRING}$ (green) for all traits measured.

Colour filled curves define the 95% credible intervals of each distribution with dots representing the median of the posterior distribution.

3.4.4 Plasticity and adaptation in *P. lanceolata* biomass

There were large differences in biomass and biomass partitioning of plants derived from the CO_2 spring versus control site population but no significant plastic response of plant biomass to single generation elevated [CO_2] (Table 1). Generation three spring plant progeny had an estimated increase in biomass of 20 % (P<0.05) relative to control plant progeny across both treatments (Fig. 6, Table 1). Plant biomass of control derived plants in generation three, averaged across growth CO_2 treatments was partitioned as 40 : 45 : 16 %, root : leaf : reproductive tissue, while spring plant progeny had higher root and leaf biomass but had allocated less resources to reproductive tissues (49 : 44 : 7 %, root : leaf : reproductive tissue) (Fig.6 and Appendix Figure B.6). Figure 6 shows that the higher whole plant biomass observed in CO_2 spring derived plants were largely driven by a 47 % higher root biomass (P<0.001), highlighting the increased allocation of resources to underground nutrient acquisition in these plants.

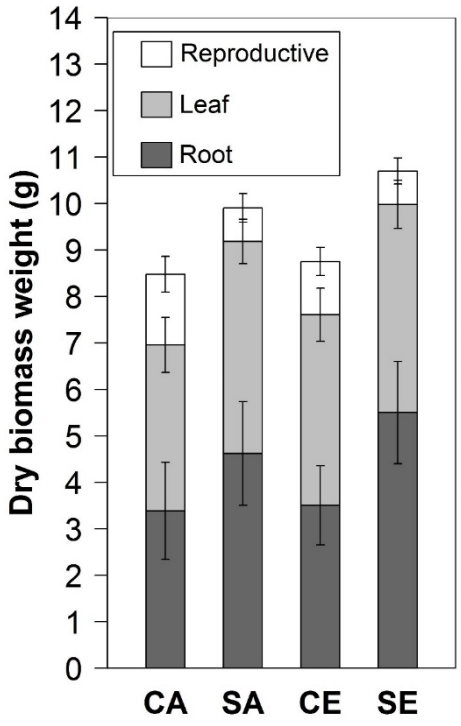


Figure 3.6: Mean biomass components of *P. lanceolata* originating from spring and control site populations when grown in ambient or elevated [CO_2] following two generations of growth at ambient [CO_2].

The four treatment groups are; SA - spring ambient, SE - spring elevated, CA – control ambient, CE – control elevated. For simplicity hybrids are not shown. Stacked bars show mean biomass weight of each component (leaf, reproductive and root) and error bars are 95% confidence intervals.

Leaf biomass was on average 18 % (P<0.05) larger in generation 3 spring plant progeny relative to control and this was likely a result of the production of more leaves, with a trend toward 13 % (P<0.1) more leaves in spring plant progeny (Table 1). Although single leaf dry weight was increased under elevated [CO_2], it was not different in spring relative to control plant progeny (Fig. 3, Appendix Table B.11). The number of rosettes produced by a plant was not significantly different between CO_2 spring and control progeny, indicating that higher leaf number was not dependent on the production of more rosettes in spring plants (Table 1).

In contrast with higher root and leaf biomass in spring versus control population progeny, reproductive biomass was 46 % lower in spring plants at the time of harvest. This is most likely to be a result of reduced allocation to the production of spikes, since seed weight, number of seeds per spike and days to flowering were not affected by population. Hybrid plant progeny also produced 35 % less reproductive biomass by harvest date than control plant progeny (Appendix Figure B.6). However, delayed phenology as well as reduced production of spikes may contribute to this lower reproductive biomass in hybrid plant progeny since there is a trend towards delayed flowering in hybrids (P<0.1) (Appendix Table B.11). Trends towards more seeds per spike and heavier seeds in hybrids, also implicate reduced spike production in the lower reproductive biomass of hybrid plants in the absence of direct evidence (Appendix Table B.11).

Seed weight of a maternal plant in generation two was positively correlated to the average total biomass of offspring in generation three ($\rho(69)=0.21$, P<0.1). This highlights that maternal effects on seed weight can still influence offspring traits into adulthood. However, since growth at elevated [CO_2] did not affect seed weight, it is unlikely that the higher biomass observed in spring versus control derived plants is a result of adaptation to elevated [CO_2] via maternal effects, and is more likely to result from direct genetic or epigenetic adaptation.

	Estimate	Standard Error	T value	
Leaf biomass [a]				
Intercept	3.72	0.23	16.22	***
$[CO_2]$ treatment $_{AMB - ELEV}$	0.22	0.24	0.91	
Population $_{CONTROL - SPRING}$	0.69	0.28	2.48	*
Population $_{CONTROL - HYBRID}$	0.03	0.30	0.08	
Root biomass [a]				
Intercept	3.14	0.41	7.66	***
$[CO_2]$ treatment $_{AMB - ELEV}$	0.62	0.48	1.46	
Population $_{CONTROL - SPRING}$	1.61	0.50	3.2	**
Population $_{CONTROL - HYBRID}$	0.27	0.54	0.50	
Reproductive biomass [b]				
Intercept	1.40	0.13	2.79	***
$[CO_2]$ treatment $_{AMB - ELEV}$	-0.14	0.14	1.05	
Population $_{CONTROL - SPRING}$	-0.6	0.16	1.53	***
Population $_{CONTROL - HYBRID}$	-0.48	0.18	2.65	**
Total biomass [a]				
Intercept	8.23	0.60	13.70	***
$[CO_2]$ treatment $_{AMB - ELEV}$	0.63	0.63	1.00	
Population $_{CONTROL - SPRING}$	1.71	0.73	2.33	*
Population $_{CONTROL - HYBRID}$	-0.32	0.80	2.33	
Total number of leaves on harvest date [a]				
Intercept	32.61	1.8	17.93	***
$[CO_2]$ treatment $_{AMB - ELEV}$	1.05	1.90	0.56	
Population $_{CONTROL - SPRING}$	4.29	2.21	1.94	†
Population $_{CONTROL - HYBRID}$	-0.45	2.41	-0.19	

Table 3.1: Effect estimates of $[CO_2]$ treatment and population site of origin (Population) on biomass traits of *Plantago lanceolata* when grown in elevated or ambient $[CO_2]$, following two generations of growth at ambient $[CO_2]$.

Biomass traits were modelled with a glm with [a] Gaussian error distribution or [b] inverse gamma distribution where coefficients were back transformed to facilitate interpretation. Subscripts describe the contrast analysed. *P*-value significance is indicated as: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; †, $P < 0.1$. Abbreviations; A – ambient, E – elevated, C – control, S – spring, H - hybrid.

3.5 Discussion

In this study we tested theoretical predictions for the evolution of transgenerational indirect genetic effects in plant response to elevated [CO_2]. Understanding how plants respond to elevated [CO_2] over multiple generations of exposure is of critical importance to predict and manage future climate change (Becklin *et al*., 2017). Through a multigenerational experiment with plants that had grown under naturally elevated [CO_2] at a CO_2 spring site, we partition phenotypic variance into components attributed to parental and grandparental effects, environmental factors and heritability using a type of mixed effect model called an 'animal model' (Wilson *et al*., 2010). We show that as hypothesised both maternal and paternal effects contribute to the multigenerational response to elevated [CO_2] in a species with limited pollen and seed dispersal. We also identified adaptation in pathways coordinating biomass responses that are likely to be the result of genetic or epigenetic adaptation.

3.5.1 Parental effects contribute to the response of *P. lanceolata* to elevated [CO_2]

Transgenerational indirect genetic effects may be predicted to evolve as part of the plant response to elevated [CO_2] because they are largely predictable from parental environment, both because there are seasonal cycles of [CO_2] and because annual differences in atmospheric [CO_2] have been directional for large periods of geological history (Berner, 2006; Tans and Keeling, 2018). A significant role for maternal effects in the multigenerational response of plants to elevated [CO_2] has previously been inferred from the effect of single generation elevated [CO_2] exposure on seed characteristics and phenology (Jablonski *et al*., 2002; Springer and Ward, 2007). However maternal effects have been only weak or absent in the response of plants to elevated [CO_2] where experimental design has tested for them over multiple generations (Teng *et al*., 2009; Steinger *et al*., 2000). Paternal effects have generally been assumed to be negligible in these studies but they may evolve in species with limited seed and pollen dispersal (Galloway, 2005; English *et al*., 2015). In this study we demonstrate that for *P. lanceolata* this is indeed the case and to our knowledge this is the first experiment to examine the role of both parental and grandparental effects in coordinating the multigenerational response of plants to elevated [CO_2]. Paternal environment was shown to significantly

affect time to first spike in this experiment, but the mechanism remains unknown. Paternal effects have previously been shown to have a significant impact on offspring phenotype expression in response to temperature in *P. lanceolata*, where prezygotic effects were more frequently transmitted paternally than maternally and low paternal temperatures accelerated the onset of flowering in progeny (Lacey, 1996). Currently there is no research into the role of paternal environment in coordinating flowering time, but examples of mechanisms that could facilitate this include paternal provisioning of RNA or proteins that affect development (Bayer *et al*., 2009; Wang and Chekanova, 2016), or epigenetic inheritance (Heard and Martienssen, 2014; Richards *et al*., 2017).

Grandparental effects did not significantly affect either trait in which they were measured: seed weight and number of seeds per spike. However, the use of an ordinal model to partition the variance associated with grandparents in these traits may have provided less resolution than if it was feasible to weigh each seed that germinated in each generation. Additional generations and the measurement of more traits would be required to examine grandparental effects more thoroughly. However, we do note that since parental effects were only identified in three of thirteen traits, grandparental effects may not be expected to play a significant role in the multigenerational response to elevated [CO_2] in this species. Grandparental effects have previously been shown to affect offspring phenotype in *P. lanceolata* response to temperature so their contribution to the multigenerational response to elevated [CO_2] warrants further investigation (Lacey *et al*., 1996; Case *et al*., 1996).

In this study we acknowledge that our sample size and the differences in environments between generations likely reduced the power to detect components contributing to the variation of phenotype, particularly indirect and direct genetic effects. For some traits the raw data appeared to suggest a trend that was not supported by statistical analysis. For example visualisation of the raw data for the time to germination appeared to suggest that germination was delayed in plants originating from the spring site. However, there was no evidence for adaptation in this trait, either as indirect genetic effects or as direct genetic effects which could suggest they are underestimated. Future experiments to determine the role of grandparental effects in plant response to elevated [CO_2] would further benefit from an increased number of experimental generations in environmentally controlled [CO_2].

3.5.2 A significant role for plasticity in the multigenerational response to elevated [CO_2]

In all of the traits studied narrow-sense heritability of traits was low with 95% credible intervals that included zero. This range of heritabilities in plant functional traits is supported by the literature (Geber and Griffin, 2000) and likely a product of predominantly plastic intraspecific variation with low genetic variation between the two populations (Siefert *et al*., 2015; Gáspár *et al*., 2016). The significant effect of the experimental generation on trait variation in almost all traits supports the significant impact of stochastic events experienced throughout the lifetime on trait expression (Steiner *et al*., 2018). Certainly, this effect was more pervasive in determining individual trait variation than either the plastic response to elevated [CO_2] or of the population site of origin which highlights the importance of integrating interactions between [CO_2] responses and other abiotic factors in predictions of plant response to climate change.

One trait that demonstrated plastic response to growth [CO_2] was time to germination. Seed sown in elevated [CO_2] in generation three germinated earlier than those sown in ambient [CO_2]. This is consistent with the increased germination rate (though with considerable interspecific variability) in meta-analysis of germination response to elevated [CO_2] (Marty and BassiriRad, 2014). Although this response is well-documented, it is not known how elevated [CO_2] is sensed by a seed and then influences plant germination, potential mechanisms include direct regulation of metabolism for example via ethylene interaction (Ziska and Bunce, 1993; Andalo *et al*., 1998) and indirect effects of soil properties or microbial community composition.

3.5.3 The potential for genetic or epigenetic adaptation to elevated [CO_2] in biomass traits

In this experiment, progeny derived from the population at the natural CO_2 spring grew larger than those derived from the ambient [CO_2] control site, even after two generations of growth at ambient [CO_2]. There was no significant plastic effect of growth [CO_2] in this experiment, although an increase in *P. lanceolata* total biomass has been reported in a number of single generation exposure studies with different experimental designs (Fajer *et al*., 1991; Staddon *et al*., 1999; Klus *et al*., 2001; He *et al*., 2002; Maestre and Reynolds, 2006), with a decrease in reproductive dry weight also supported (Fajer *et al*., 1991). The increased allocation of biomass to vegetative tissue rather than reproductive tissue observed in this experiment may result from increased growth rate and delayed flowering

(Bernacchi *et al*., 2000), or could be a potentially adaptive active redistribution of resources, for example that increases survival and competitive ability over winter resulting in increased reproductive output the following year (Fajer *et al*. 1991).

Although parental effects were not quantified in biomass traits in this experiment, the increased biomass observed in plants from the spring site population was unlikely to be the result of parental effects. This is because although biomass was correlated to seed weight, spring plants did not produce significantly heavier seeds than control plants across the experiment. Thus adaptation in the coordination of plant biomass is more likely to be a result of genetic or epigenetic adaptation. It is difficult to predict where in the pathway this adaptation could occur without photosynthetic rate measurements, which would discern whether biomass accumulation was a downstream consequence of a fixed increase in photosynthetic rate. This increased biomass phenotype is enigmatic because it is still observed when spring plants are grown at ambient [CO_2] for two generations, despite that there is less substrate available for photosynthesis. This requires further investigation with interest to commercial farmers of *Plantago* species, but may have wider applications if it is possible to extrapolate to crop breeding (de Freitas Lima *et al*., 2017).

Analysis of gene expression in a crossed factored experiment with *P. lanceolata* from the spring and control site populations has previously suggested that spring plant progeny are adapted to elevated [CO_2] at the level of gene expression (Watson-Lazowski *et al*., 2016). In this study we found that hybrids between the two populations expressed different phenotypes to their parental populations which seems to suggest some regulatory incompatibility between the two populations that could be the result of fixed gene expression differences.

Experimental studies have traditionally considered direct genetic changes and/or the influence of maternal plants on their offspring as mechanisms of plant adaptation to elevated [CO_2] (Wieneke *et al*., 2004; Teng *et al*., 2009; Van Loon et al, 2016). Further, maternal effects are largely investigated by measuring only offspring traits that are directly influenced by the mother (i.e. seed provisioning or effect of flowering time) (Springer and Ward, 2000; Jablonski *et al*., 2002). Although maternal effects may be most likely to evolve in these traits, an increasing awareness of specialised mechanisms of

transgenerational indirect genetic effects that do not involve direct influence on offspring environment, suggests that these may not be the only traits that can be affected by indirect genetic effects (Miska *et al*., 2016). Further these effects on traits may be confounded with genetic adaptation since although experiments looking to quantify genetic adaptation often grow plants in a common environment for a generation prior to applying a treatment, this may not be sufficient to negate indirect genetic effects such as epigenetic inheritance (Latzel, 2015). In environmental changes where transgenerational indirect genetic effects are predicted to evolve, such as in response to [CO_2], it is therefore imperative to better understand the contribution of paternal and grandparental effects and the mechanisms by which they occur in order to also understand the potential for adaptation.

In conclusion we demonstrate through a multigenerational experiment with plants that have grown at elevated [CO_2] for potentially hundreds of generations, that both maternal and paternal effects can contribute to the plant response to elevated [CO_2]. Although we have not identified grandparental effects in this experiment we propose that these effects as well as parental effects may play an important role in the multigenerational response of plants to elevated [CO_2] under climate change and should be further investigated. In addition we find that plants originating from the natural CO_2 spring site have larger biomass and reallocate resources from reproductive to vegetative biomass, suggesting genetic or epigenetic adaptation to elevated [CO_2].

3.6 Acknowledgements

We thank F. Miglietta for assisting with access to the spring site and Kevin Staniland for help in the field. We acknowledge the help of Suzanne Milner, Mike Allwright and Billy Valdes with plant care and for assistance with harvests. This work was funded by the National Environmental Research Council (NERC), through a SPITFIRE DTP PhD studentship to JMS [grant number NE/L002531/1], and through funding to GT from UC Davis and The British Council.

3.7 Author Contributions

GT, MAC, THGE and JMS conceived the study together. GT, MAC and JS undertook field work. JMS undertook the experiment and phenotypic measurements, as well as data analysis with advice from THGE. JMS drafted the manuscript and all contributed to revision.

3.8 References

Andalo, C., Raquin, C., Machon, N., Godelle, B., & Mousseau, M. (1998) Direct and maternal effects of elevated CO 2 on early root growth of germinating *Arabidopsis thaliana* seedlings. *Annals of Botany*, **81**, 405-411.

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.

Bates, D.M., Maechler, M., Bolker, B., & Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**, 1-48.

Bayer, M., Nawy, T., Giglione, C., Galli, M., Meinnel, T., & Lukowitz, W. (2009) Paternal control of embryonic patterning in *Arabidopsis thaliana*. *Science*, **323**, 1485-1488.

Becklin, K. M., Walker, S. M., Way, D. A., & Ward, J. K. (2017) CO_2 studies remain key to understanding a future world. *New Phytologist*, **214**, 34-40.

Bernacchi, C. J., Coleman, J. S., Bazzaz, F. A., & McConnaughay, K. D. M. (2000) Biomass allocation in old-field annual species grown in elevated CO_2 environments: no evidence for optimal partitioning. *Global Change Biology*, **6**, 855-863.

Berner, R. A. (2006) GEOCARBSULF: a combined model for Phanerozoic atmospheric O_2 and CO_2. *Geochimica et Cosmochimica Acta*, **70**, 5653-5664.

Bettarini, I., Vaccari, F. P., & Miglietta, F. (1998) Elevated CO_2 concentrations and stomatal density: observations from 17 plant species growing in a CO_2 spring in central Italy. *Global Change Biology*, **4**, 17-22.

Bos, M., Harmens, H., & Vrieling, K. (1986) Gene flow in *Plantago* I. Gene flow and neighbourhood size in *P. lanceolata*. *Heredity*, **56**, 43.

Brooks, S P. and Gelman, A. (1998) General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, **7**, 434-455.

Case, A., Curtis, P., & Snow, A. (1998) Heritable variation in stomatal responses to elevated CO_2 in wild radish, *Raphanus raphanistrum* (Brassicaceae). *American Journal of Botany*, **85**, 253-253.

Case, A. L., Lacey, E. P., & Hopkins, R. G. (1996) Parental effects in *Plantago lanceolata* L. II. Manipulation of grandparental temperature and parental flowering time. *Heredity*, **76**, 287.

Curley, J. P., Mashoodh, R., & Champagne, F. A. (2011) Epigenetics and the origins of paternal effects. *Hormones and behavior*, **59**, 306-314.

De Costa, W. A. J. M., Weerakoon, W. M. W., Chinthaka, K. G. R., Herath, H. M. L. K., & Abeywardena, R. M. I. (2007) Genotypic variation in the response of rice (*Oryza sativa* L.) to increased atmospheric carbon dioxide and its physiological basis. *Journal of Agronomy and Crop Science*, **193**, 117-130.

de Freitas Lima, M., Eloy, N. B., de Siqueira, J. A. B., Inzé, D., Hemerly, A. S., & Ferreira, P. C. G. (2017) Molecular mechanisms of biomass increase in plants. *Biotechnology Research and Innovation*, **1**, 14-25.

Diggle, P. K., Abrahamson, N. J., Baker, R. L., Barnes, M. G., Koontz, T. L., Lay, C. R., Medeiros, J.S., Murgel, J.L., Shaner, M.G.M., Simpson, H.L., & Wu, C. C. (2010) Dynamics of maternal and paternal effects on embryo and seed development in wild radish (*Raphanus sativus*). *Annals of Botany*, **106**, 309-319.

Dippery, J. K., Tissue, D. T., Thomas, R. B., & Strain, B. R. (1995) Effects of low and elevated CO_2 on C_3 and C_4 annuals. *Oecologia*, **101**, 13-20.

English, S., Pen, I., Shea, N., & Uller, T. (2015) The information value of non-genetic inheritance in plants and animals. *PloS one*, **10**, e0116996.

Fajer, E. D., Bowers, M. D., & Bazzaz, F. A. (1991) Performance and allocation patterns of the perennial herb, *Plantago lanceolata*, in response to simulated herbivory and elevated CO_2 environments. *Oecologia*, **87**, 37-42.

Frenck, G., van der Linden, L., Mikkelsen, T. N., Brix, H., & Jørgensen, R. B. (2013) Response to multi-generational selection under elevated [CO_2] in two temperature regimes suggests enhanced carbon assimilation and increased reproductive output in *Brassica napus* L. *Ecology and evolution*, **3**, 1163-1172.

Galloway, L. F. (2001) The effect of maternal and paternal environments on seed characters in the herbaceous plant *Campanula americana* (Campanulaceae). *American Journal of Botany*, **88**, 832-840.

Galloway, L. F. (2005) Maternal effects provide phenotypic adaptation to local environmental conditions. *New Phytologist*, **166**, 93-100.

Galloway, L. F., & Etterson, J. R. (2007) Transgenerational plasticity is adaptive in the wild. *Science*, **318**, 1134-1136.

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*, https://doi.org/10.1111/nph.15487

Geber, M. A., & Griffen, L. R. (2003) Inheritance and natural selection on functional traits. *International Journal of Plant Sciences*, **164**, 21-42.

Gratani, L. (2014) Plant Phenotypic Plasticity in Response to Environmental Factors. *Advances in* Botany, 2014, http://dx.doi.org/10.1155/2014/208747.

Groot, M. P., Kooke, R., Knoben, N., Vergeer, P., Keurentjes, J. J., Ouborg, N. J., & Verhoeven, K. J. (2016) Effects of multi-generational stress exposure and offspring environment on the expression and persistence of transgenerational effects in *Arabidopsis thaliana*. *PloS one*, **11**, e0151566.

Groot, M. P., Kubisch, A., Ouborg, N. J., Pagel, J., Schmid, K. J., Vergeer, P., & Lampei, C. (2017) Transgenerational effects of mild heat in *Arabidopsis thaliana* show strong genotype specificity that is explained by climate at origin. *New Phytologist*, **215**, 1221-1234.

Grossman, J. D., & Rice, K. J. (2014) Contemporary evolution of an invasive grass in response to elevated atmospheric CO_2 at a Mojave Desert FACE site. *Ecology letters*, **17**, 710-716.

Grünzweig, J. M., & Dumbur, R. (2012) Seed traits, seed-reserve utilization and offspring performance across pre-industrial to future CO_2 concentrations in a Mediterranean community. *Oikos,* **121,** 579-588.

Hadfield, J.D. (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, **33**, 1-22.

He, J. S., Bazzaz, F. A., & Schmid, B. (2002) Interactive effects of diversity, nutrients and elevated CO_2 on experimental plant communities. *Oikos*, **97**, 337-348.

Heard, E., & Martienssen, R. A. (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell*, **157**, 95-109.

Latzel, V. (2015) Pitfalls in ecological research–transgenerational effects. *Folia geobotanica*, **50**, 75-85.

Heidelberger, P., & Welch, P. D. (1983) Simulation run length control in the presence of an initial transient. *Operations Research*, **31**, 1109-1144.

Henderson, C. R. (1976) A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics*, **32**, 69-83.

Herman, J. J., Sultan, S. E., Horgan-Kobelski, T., & Riggs, C. (2012) Adaptive transgenerational plasticity in an annual plant: grandparental and parental drought stress enhance performance of seedlings in dry soil. *Integrative and Comparative Biology,* **52**, 77-88.

Herman, J. J., Spencer, H. G., Donohue, K., Sultan, S. E. (2014) How stable 'should' epigenetic modifications be? Insights from adaptive plasticity and bet hedging. *Evolution,* **68**, 632-643.

Herschbach, C., Schulte, M., Von Ballmoos, P., Brunold, C., & Rennenberg, H. (2012) Sulfate and Nitrate Assimilation in Leaves of *Quercus ilex* and *Quercus pubescens* Grown Near Natural CO_2 Springs in Central Italy. In L.J. De Kok *et al*., (Eds.) *Sulfur Metabolism in Plants* (pp. 237-248). Springer, Dordrecht.

Holeski, L. M., Jander, G., & Agrawal, A. A. (2012) Transgenerational defense induction and epigenetic inheritance in plants. *Trends in ecology & evolution*, **27**, 618-626.

Hui, D., Sims, D. A., Johnson, D. W., Cheng, W., & Luo, Y. (2002) Effects of gradual versus step increases in carbon dioxide on *Plantago* photosynthesis and growth in a microcosm study. *Environmental and Experimental Botany*, **47**, 51-66.

Jablonski, L. M., Wang, X., & Curtis, P. S. (2002) Plant reproduction under elevated CO_2 conditions: a meta-analysis of reports on 79 crop and wild species. *New Phytologist*, **156**, 9-26.

Kassambara, A., (2016) ggcorrplot: Visualization of a Correlation Matrix using'ggplot2'. R package version 0.1.1. http://www.sthda.com/english/wiki/ggcorrplot

Klus, D. J., Kalisz, S., Curtis, P. S., Teeri, J. A., & Tonsor, S. J. (2001) Family-and population-level responses to atmospheric CO_2 concentration: gas exchange and the allocation of C, N, and biomass in *Plantago lanceolata* (*Plantaginaceae*). *American Journal of Botany*, 1080-1087.

Körner, C., & Miglietta, F. (1994) Long term effects of naturally elevated CO_2 on Mediterranean grassland and forest trees. *Oecologia*, **99**, 343-351.

Kou, H. P., Li, Y., Song, X. X., Ou, X. F., Xing, S. C., Ma, J., Liu, B., and Von Wettstein, D. (2011) Heritable alteration in DNA methylation induced by nitrogen-deficiency stress accompanies enhanced tolerance by progenies to the stress in rice (*Oryza sativa* L.). *J. Plant Physiol.* **168**, 1685–1693.

Kruuk, L.E.B., Slate, J., & Wilson, A.J. (2008) New answers for old questions: the evolutionary quantitative genetics of wild animal populations. *Annual Review of Ecology, Evolution, and Systematics,* **39**, 525–548.

Lacey EP. (1996) Parental effects in *Plantago lanceolata* L. I: A growth chamber experiment to examine pre- and postzygotic temperature effects. *Evolution,* **52**, 865–878.

Latzel, V., & Klimešová, J. (2009) Fitness of resprouters versus seeders in relation to nutrient availability in two *Plantago* species. *Acta Oecol,* **35**, 541–547.

Latzel, V., Hájek, T., Klimešová, J., & Gómez, S. (2009) Nutrients and disturbance history in two *Plantago* species: maternal effects as a clue for observed dichotomy between resprouting and seeding strategies. *Oikos*, **118**, 1669-1678.

Latzel, V., Klimešová, J., Hájek, T., Gómez, S., & Šmilauer, P. (2010) Maternal effects alter progeny's response to disturbance and nutrients in two *Plantago* species. *Oikos*, **119**, 1700-1710.

Latzel, V., Janeček, Š., Doležal, J., Klimešová, J., & Bossdorf, O. (2014) Adaptive transgenerational plasticity in the perennial *Plantago lanceolata*. *Oikos*, **123**, 41-46.

Lenth, R. 2018. R package emmeans: Estimated marginal means. R package version 1.3.0. https://cran.r-project.org/web/packages/emmeans/index.html

Leakey, A. D., Ainsworth, E. A., Bernacchi, C. J., Rogers, A., Long, S. P., & Ort, D. R. (2009) Elevated CO_2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. *Journal of experimental botany*, **60**, 2859-2876.

Leakey, A. D., & Lau, J. A. (2012) Evolutionary context for understanding and manipulating plant responses to past, present and future atmospheric [CO_2]. *Phil. Trans. R. Soc. B*, **367**, 613-629.

Maestre, F. T., & Reynolds, J. F. (2007) Biomass responses to elevated CO_2, soil heterogeneity and diversity: an experimental assessment with grassland assemblages. *Oecologia*, **151**, 512-520.

Marshall, D.J., & Uller, T. (2007) When is a maternal effect adaptive? *Oikos*, **116**, 1957-1963.

Marty, C., & BassiriRad, H. (2014) Seed germination and rising atmospheric CO_2 concentration: a meta-analysis of parental and direct effects. *New phytologist*, **202**, 401-414.

McFarlane, S. E., Gorrell, J. C., Coltman, D. W., Humphries, M. M., Boutin, S., & McAdam, A. G. (2014) Very low levels of direct additive genetic variance in fitness and fitness components in a red squirrel population. *Ecology and evolution*, **4**, 1729-1738.

Miglietta, F., Raschi, A., Bettarini, I., Resti, R., & Selvi, F. (1993) Natural CO_2 springs in Italy: a resource for examining long-term response of vegetation to rising atmospheric CO_2 concentrations. *Plant, Cell & Environment*, **16**, 873-878.

Miska, E. A., & Ferguson-Smith, A. C. (2016) Transgenerational inheritance: Models and mechanisms of non–DNA sequence–based inheritance. *Science*, **354**, 59-63.

Mousseau, T. A., & Fox, C. W. (1998) The adaptive significance of maternal effects. *Trends in ecology & evolution*, **13**, 403-407.

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., Carlisle, E., Dietterich, L. H., Fitzgerald, G., Hasegawa, T., & Holbrook, N. M. (2014) Increasing CO_2 threatens human nutrition. *Nature*, **510**, 139.

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 CO_2 concentration. *Oecologia*, **165**, 809-818.

Navarrete, S., Kemp, P. D., Pain, S. J., & Back, P. J. (2016) Bioactive compounds, aucubin and acteoside, in plantain (*Plantago lanceolata* L.) and their effect on in vitro rumen fermentation. *Animal feed science and technology*, **222**, 158-167.

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius, U., Poot, P., Purugganan, M.D., Richards, C.L., Valladares, F., & van Kleunen, M. (2010) Plant phenotypic plasticity in a changing climate. *Trends in plant science*, **15**, 684-692.

Norton, L. R., Firbank, L. G., Gray, A. J., & Watkinson, A. R. (1999) Responses to elevated temperature and CO_2 in the perennial grass *Agrostis curtisii* in relation to population origin. *Functional Ecology*, **13**, 29-37.

Paoletti, E., Pfanz, H., & Raschi, A. 2005. Pros and cons of CO_2 springs as experimental sites. In K. Omasa *et al*., (Eds.) *Plant responses to air pollution and global change*. pp. 195-202. Springer, Tokyo.

Poorter, H., & Navas, M. L. (2003) Plant growth and competition at elevated CO_2: on winners, losers and functional groups. *New Phytologist*, **157**, 175-198.

Potvin, C., & Tousignant, D. (1996) Evolutionary consequences of simulated global change: genetic adaptation or adaptive phenotypic plasticity. *Oecologia*, **108**, 683-693.

Prior, S. A., Runion, G. B., Marble, S. C., Rogers, H. H., Gilliam, C. H., & Torbert, H. A. (2011) A review of elevated atmospheric CO_2 effects on plant growth and water relations: implications for horticulture. *HortScience*, **46**, 158-162.

Rando, O. J. (2012) Daddy issues: paternal effects on phenotype. *Cell*, **151**, 702-708.

Reddy, A. R., Rasineni, G. K., & Raghavendra, A. S. (2010) The impact of global elevated CO_2 concentration on photosynthesis and plant productivity. *Current Science*, **99**, 46-57.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Richards, C. L., Alonso, C., Becker, C., Bossdorf, O., Bucher, E., Colomé-Tatché, M., Durka, W., Engelhardt, J., Gaspar, B., Gogol-Döring, A., & Grosse, I. (2017) Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology letters*, **20**, 1576-1590.

Scholefield, P. A., Doick, K. J., Herbert, B. M. J., Hewitt, C. N., Schnitzler, J. P., Pinelli, P., & Loreto, F. (2004) Impact of rising CO_2 on emissions of volatile organic compounds: isoprene emission from Phragmites australis growing at elevated CO_2 in a natural carbon dioxide spring. *Plant Cell and Environment,* **27,** 393-401.

Schulte, M., Von Ballmoos, P., Rennenberg, H., & Herschbach, C. (2002) Life-long growth of *Quercus ilex* L. at natural CO_2 springs acclimates sulphur, nitrogen and carbohydrate metabolism of the progeny to elevated pCO_2. *Plant Cell and Environment,* **25,** 1715-1727.

Schmid, B., and C. Dolt. (1994) Effects of paternal and maternal environment and genotype on offspring phenotype in *Solidago altissima* L. *Evolution*, **48**, 1525–1549.

Selvi. F. (1998) Flora of the mineral CO_2-spring Bossoleto (Rapolano Terme, Tuscany) and its relevance to ecological research. *Atti della Società toscana di scienze naturali, residente in Pisa, Memorie, Serie B*, **105**, 23-30.

Siefert, A., Violle, C., Chalmandrier, L., Albert, C. H., Taudiere, A., Fajardo, A., Aarssen, L. W, Baraloto, C., Carlucci, M. B., Cianciaruso, M.V., Dantas, V. L., Bello, F., Duarte, L. D. S., Fonesca, C. R., Freschet, G. T., Gaucherand, S., Gross, N., Hikosaka, K., Jackson, B., Jung, V., Kamiyama, C., Katabuchi, M., Kembel, S. W., Kichenin, E., Kraft, N. J. B., Lagerström, A., Le Bagousse-Pinguet, Y., Li, Y., Mason, N., Messier, J., Nakashizuka, T., McC. Overton, J., Peltzer, D. A., Pérez-Ramos, I. M., Pillar, V. D., Prentice, H. C., Richardson, S., Sasaki, T., Schamp, B. S., Schöb, C., Shipley, B., Sundqvist, M., Sykes, M. T., Vandewalle, M., & Wardle, D. A. (2015) A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology Letters,* **18***, 1406–1419.*

Springer, C. J., & Ward, J. K. (2007) Flowering time and elevated atmospheric CO_2. *New Phytologist*, **176**, 243-255.

Staddon, P. L., Fitter, A. H., & Robinson, D. (1999) Effects of mycorrhizal colonization and elevated atmospheric carbon dioxide on carbon fixation and below-ground carbon partitioning in *Plantago lanceolata*. *Journal of Experimental Botany*, **50**, 853-860.

Steiner, U., Tuljapurkar, S., & Roach, D. (2018) Quantifying genetic, environmental and individual stochastic variability in *Plantago lanceolata*. *bioRxiv*, 270603.

Steinger, T., Gall, R., & Schmid, B. (2000) Maternal and direct effects of elevated CO_2 on seed provisioning, germination and seedling growth in *Bromus erectus*. *Oecologia*, **123**, 475-480.

Steinger, T., Stephan, A., & Schmid, B. (2007) Predicting adaptive evolution under elevated atmospheric CO_2 in the perennial grass *Bromus erectus*. *Global Change Biology*, **13**, 1028-1039.

Stewart, A. V. (1996) Plantain (*Plantago lanceolata*)-a potential pasture species. *Proceedings of the Conference-New Zealand Grassland Association*, **58**, 77-86.

Tans, P., & Keeling, R. (2018) ESRL Global Monitoring Division Global Greenhouse Gas Reference Network. http://www.esrl.noaa.gov/gmd/ccgg/trends/ Accessed 18th October 2018.

Taylor, G., Tricker, P.J., Zhang, F.Z., Alston, V.J., Miglietta, F., & Kuzminsky, E., (2003) Spatial and Temporal Effects of Free-Air CO_2 Enrichment (POPFACE) on Leaf Growth, Cell Expansion, and Cell Production in a Closed Canopy of Poplar. *Plant Physiology*, **131**, 177-185.

Temme, A. A., Liu, J. C., Cornwell, W. K., Cornelissen, J. H., & Aerts, R. (2015) Winners always win: growth of a wide range of plant species from low to future high CO_2. *Ecology and Evolution*, **5**, 4949-4961.

Teng, N., Jin, B., Wang, Q., Hao, H., Ceulemans, R., Kuang, T., & Lin, J. (2009) No detectable maternal effects of elevated CO_2 on *Arabidopsis thaliana* over 15 generations. *PLoS One*, **4**, e6035.

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**, 833-843.

Uller, T. (2008) Developmental plasticity and the evolution of parental effects. *Trends in ecology & evolution*, **23**, 432-438.

Van Gardingen, P., Grace, J., Jeffree, C., Byari, S., Miglietta, F., Raschi, A., & Bettarini, I. (1997) Long-term effects of enhanced CO_2 concentrations on leaf gas exchange: research opportunities using CO_2 springs. In A. Raschi *et al*., (Eds.) *Plant Responses to Elevated CO_2 (*pp. 69-86). Cambrige University Press, Cambridge**,** UK.

Van Loon, M. P., Rietkerk, M., Dekker, S. C., Hikosaka, K., Ueda, M. U., & Anten, N. P. (2016) Plant–plant interactions mediate the plastic and genotypic response of *Plantago asiatica* to CO_2: an experiment with plant populations from naturally high CO_2 areas. *Annals of botany*, **117**, 1197-1207.

Varga, S., & Soulsbury, C. D. (2017) Paternal arbuscular mycorrhizal fungal status affects DNA methylation in seeds. *Biology letters*, **13**, 20170407.

Villemereuil, P., Gimenez, O., Doligez, B., and Freckleton, R. (2013) Comparing parent–offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations: a simulation study for Gaussian and binary traits. *Methods in Ecology and Evolution*, **4**, 260-275.

Wang, H. L. V., & Chekanova, J. A. (2016) Small RNAs: essential regulators of gene expression and defenses against environmental stresses in plants. *Wiley Interdisciplinary Reviews: RNA*, **7**, 356-381.

Ward, J. K., & Strain, B. R. 1997. Effects of low and elevated CO_2 partial pressure on growth and reproduction of Arabidopsis thaliana from different elevations. *Plant, Cell & Environment*, **20**, 254-260.

Ward, J. K., Antonovics, J., Thomas, R. B., & Strain, B. R. (2000) Is atmospheric CO_2 a selective agent on model C_3 annuals? *Oecologia*, **123**, 330-341.

Watson-Lazowski, A., Lin, Y., Miglietta, F., Edwards, R. J., Chapman, M. A., & Taylor, G. (2016) Plant adaptation or acclimation to rising CO_2? Insight from first multigenerational RNA-Seq transcriptome. *Global change biology*, **22**, 3760-3773.

Whittle, C., Otto, S., Johnston, M., and Krochko, J. (2009) Adaptive epigenetic memory of ancestral temperature regime in *Arabidopsis thaliana*. *Botany,* **87**, 650–657.

Wieneke, S., Prati, D., Brandl, R., Stöcklin, J., & Auge, H. (2004) Genetic variation in *Sanguisorba minor* after 6 years in situ selection under elevated CO_2. *Global Change Biology*, **10**, 1389-1401.

Wilson, A. J., Reale, D., Clements, M. N., Morrissey, M. M., Postma, E., Walling, C. A., Kruuk, L. E., & Nussey, D. H. (2010) An ecologist's guide to the animal model. *Journal of Animal Ecology*, *79*(1), 13-26.

Wolf, J. B., Brodie III, E. D., Cheverud, J. M., Moore, A. J., & Wade, M. J. (1998) Evolutionary consequences of indirect genetic effects. *Trends in ecology & evolution*, **13**, 64-69.

Wolf, J. B., & Wade, M. J. (2009) What are maternal effects (and what are they not)? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **364**, 1107-1115.

Wulff, R. D., & Alexander, H. M. (1985) Intraspecific variation in the response to CO_2 enrichment in seeds and seedlings of *Plantago lanceolata* L. *Oecologia*, **66**, 458-460.

Wulff, R. D., Causin, H. F., Benitez, O., & Bacalini, P. A. (1999) Intraspecific variability and maternal effects in the response to nutrient addition in *Chenopodium album*. *Canadian Journal of Botany*, **77**, 1150-1158.

Yamori, W., Hikosaka, K., & Way, D. A. (2014) Temperature response of photosynthesis in C_3, C_4, and CAM plants: temperature acclimation and temperature adaptation. *Photosynthesis research*, **119**, 101-117.

Zheng, Y., Li, F., Hao, L., Shedayi, A. A., Guo, L., Ma, C., Huang, B., & Xu, M. (2018) The optimal CO_2 concentrations for the growth of three perennial grass species. *BMC plant biology*, **18**, 27.

Ziska, L. H., & Bunce, J. A. (1993) The influence of elevated CO_2 and temperature on seed germination and emergence from soil. *Field Crops Research*, **34**, 147-157.

Chapter 4 The role of methylation in the single and multigenerational plant response to elevated [CO_2] at a natural CO_2 spring

J. M. Saban[a], A. J. Watson-Lazowski[a, b], Y. Lin[a], M. A. Chapman[a], G. Taylor[a,c]

[a] Biological Sciences, University of Southampton, Life Sciences, Building 85, Highfield Campus, Southampton, SO17 1BJ, UK

[b] Hawkesbury Institute for the Environment, University of Western Sydney, Penrith, NSW 2753, Australia

[c] Department of Plant Sciences, University of California, Davis, CA 95616, USA

In preparation for submission

4.1 Abstract (250 words)

Elevated atmospheric CO_2 concentrations ($[CO_2]$) predicted under climate change will significantly affect plant functioning, and understanding these responses is key to predicting and managing the impact of climate change. Fundamental questions about the mechanisms and processes that coordinate the multigenerational response of plants to environmental change remain. The multigenerational elevated $[CO_2]$ response is a particular challenge to study because of the high cost of $[CO_2]$ enrichment facilities. In this study we utilise *Plantago lanceolata* L. growing for potentially hundreds of generations at elevated and ambient $[CO_2]$ at a natural CO_2 spring to analyse the mechanistic basis of the single and multigenerational response to elevated $[CO_2]$. We show that single generation exposure to elevated $[CO_2]$ induces plastic remodelling of the plant methylome, and that some of this may be stably transgenerationally inherited. However there is considerably more differentiation between the methylation profiles of plants that have grown for multiple generations at elevated relative to ambient $[CO_2]$ than is induced by growth at elevated $[CO_2]$ for a single generation. A small component of the differences in methylation between these two populations can be attributed to a correlation between genetic and methylation profile variation, but we hypothesise that a larger component is the result of spontaneous epimutation and subsequent selection or drift.

4.2 Introduction

The study of the mechanistic basis of plant response to the environment is becoming increasingly tractable in non-model species with the increasing affordability of High Throughput Sequencing (HTS). Arguably, nowhere is the application of this technology more important than in predicting how plants will respond to climate change. Understanding the molecular mechanisms by which plants coordinate a phenotypic response to environmental change, and how evolutionary processes act on these mechanisms over many generations, is a key challenge to predicting plant responses under anthropogenic climate change. Increased atmospheric [CO_2] will be a significant feature of future climates, with [CO_2] upwards of 430 ppm (IPCC, 2014) for the first time in millions of years of evolutionary time (Pearson and Palmer, 2000). Our ability to predict the response of plants to this increased carbon availability has critical importance to prediction of ecosystem change (Barnaby and Ziska, 2012), biosphere interactions (Forkel *et al*., 2016) and food security (Wheeler and Von Braun, 2013; Myers *et al*., 2014).

Plant plastic phenotypic responses to elevated [CO_2] within a single generation have been extensively quantified in diverse plant species and using various experimental designs (Ainsworth and Long 2005; Ainsworth and Rogers, 2007; Wang *et al*., 2012). However, our understanding of the multigenerational response of plants to elevated [CO_2] has been limited by the challenge of growing multiple generations of plants in experimentally elevated [CO_2]. Evidence for elevated [CO_2] acting as a selection pressure on standing genetic variation over short timescales is limited and controversial in selection experiments (Ward *et al*., 2000; Wieneke *et al*., 2004; Frenck *et al*., 2013) and rarer still in realistic field conditions (Leakey and Lau, 2012; Grossman and Rice, 2014).

Natural CO_2 springs provide a resource for investigating plant response to elevated [CO_2] over multiple generations without the extensive labour, time and financial costs associated with other systems (Paoletti *et al*., 2005). Plant response to multiple generations of elevated [CO_2] at natural CO_2 springs are generally consistent in direction and magnitude to those observed in single generation Free Air CO_2 Enrichment (FACE) experiments (Saban *et al*., 2018) but additionally offer the possibility of quantifying the relative contribution of indirect and direct genetic effects. Growing plants from populations experiencing many generations of growth at elevated or ambient [CO_2] in

crossed factored experiments of single generation ambient and elevated [CO_2], have generally indicated that these responses are not solely plastic and that there is adaptation to elevated [CO_2] in populations at natural CO_2 springs (Onoda *et al*., 2009; Nakamura *et al*., 2011; Watson-Lazowski *et al*., 2016). The mechanism of this adaptation remains unclear, with one study identifying adaptation to elevated [CO_2] in the gene expression profiles of plants at a CO_2 spring but very little genetic divergence between that population and one growing at a nearby ambient control site (Watson-Lazowski *et al*., 2016).

Global DNA methylation patterns have been shown to be responsive to abiotic environmental conditions, including salinity (Karan *et al*., 2012; Yaish *et al*., 2018), temperature (Ma, *et al*., 2015) and drought (Neves *et al*., 2017), and may coordinate adaptive phenotypes (Xia *et al*., 2016), but the role of methylation has never been explored in plant plastic or adaptive responses to elevated [CO_2]. Methylation of cytosine in DNA occurs more extensively in plants than animals, and with large variation in patterns between species (Niederhuth *et al*., 2016; Yi, 2017). In plants methylation occurs in three cytosine contexts, CG, CHH and CHG (where H is any base except G) (Henderson and 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 and Cao, 2017). Broadly, DNA methylation appears to function to silence the mobility of transposable elements, contribute to genome stability and integrity, and may play a role in gene expression regulation (Zhang *et al*., 2010; Law and Jacobsen 2010; Saze *et al*., 2012; Eichten *et al*., 2014; Zilberman, 2017).

Given the potential role of DNA methylation in modulating gene expression as part of a plastic response to environmental cues (Garg *et al*., 2015), coordination of some element of plant response to elevated [CO_2] by reprogramming of global methylation is an attractive hypothesis (Watson-Lazowski *et al*., 2016). Further, the observation that in plants methylation can be maintained through mitotic and meiotic cell division (Law and Jacobsen *et al*., 2010; Verhoeven *et al*., 2010; Quadrana and Colot, 2016), has led to the hypothesis that methylation could provide transgenerational memory of ancestral environment, contributing to phenotype expression in offspring (Heard and Martienssen, 2014; Quadrana and Colot, 2016). However experimental evidence of environmentally

induced methylation patterns that both influence phenotype and are inherited into the next generation is rare (Quadrana and Colot, 2016; Crisp *et al*., 2016).

We sought to elucidate the mechanistic basis for the plastic and adaptive response of a herbaceous rosette-forming plant *Plantago lanceolata* L. exposed to elevated [CO_2] at a natural CO_2 spring in Italy. The Bossoleto natural CO_2 spring is anecdotally more than 100 years old and the population of *P. lanceolata* was first documented in 1998 (Bettarini et al., 1998). We combine gene expression, DNA methylation and DNA sequence datasets to explore the plastic and adaptive response of *P. lanceolata* to elevated [CO_2] with the aim of answering the following questions i) Does DNA methylation play a role in the plastic response of plants to elevated [CO_2]? ii) Do plants growing for many generations at elevated [CO_2] adapt to these conditions, and if so does this occur solely by direct genetic mechanisms? and iii) If methylation plays a role in the adaptation of plants to elevated [CO_2], does this occur via environmentally induced methylation changes that are transgenerationally stable?

4.3 Methods

4.3.1 Plant material and sampling site

The study system for this work was *Plantago lanceolata* L. (Plantaginaceae), a herbaceous perennial with widespread geographical distribution. The experimental design has been described previously in Watson-Lazowski *et al*., 2016 and is summarised briefly here. Seeds were collected from nine maternal plants growing in naturally elevated [CO_2] near to the CO_2 spring at Bossoleto, Italy (Lat. 43°17', Long. 11°35') on the 12th May 2008. At this site the average daytime [CO_2] is around 1000 µmol mol^{-1}, with a range of 400-1200 µmol mol^{-1} (Scholefield *et al*., 2004). For comparison, seed was also obtained on this date from nine *P. lanceolata* individuals growing at a nearby (ca. 200 m apart) ambient [CO_2] control site.

In order to reduce the potential for indirect genetic (especially maternal) effects to confound the interpretation of genetic or epigenetic adaptation in this experiment, seeds were grown for one generation in the same environment in the glasshouse at the University of Southampton. *P. lanceolata* is an obligate outcrosser, therefore plants were crossed within families by isolation of maternal siblings in a muslin tent. Seed was

collected every month and stored at 5 °C. In 2009 seeds from within family crosses were sown on compost in 20 cm pots. Three weeks after germination they were moved to eight environmentally controlled chambers, with four each set to ambient and elevated [CO_2] (410.63 ± 33.74 ppm and 718 ± 46.81 ppm respectively). Flow rate was measured at 3.4 m s^{-1} and with photosynthetically active radiation (PAR) at 104-134 µmol m^{-2} s^{-1}, temperature 22 °C /17 °C and a 16 hour day length. A rigorous randomisation schedule was employed to reduce chamber effects, by randomising pots within and between chambers every two days, and swapping chamber treatments every week. On the 58th day after establishment, the second or third young expanding leaf was harvested into liquid nitrogen and stored at -80 °C for sequencing.

4.3.2 Phenotypic measurements

Above-ground biomass and whole plant leaf area was measured by harvesting plants on the 124th day of the experiment. Stomatal density, stomatal index, epidermal cell area and guard cell length were calculated from epidermal leaf imprints of a mature leaf at final harvest (Potvin and Tousignant, 1996). These measurements and statistical analyses are described in detail in Watson-Lazowski *et al*. (2016).

4.3.3 Transcriptome assembly and gene expression analysis

The RNA-Seq dataset previously analysed and reported in Watson-Lazowski *et al.,* 2016 was re-analysed here using increased computational resources, from raw RNA-Seq reads through to gene ontology analysis. Briefly, RNA was extracted from frozen leaf material of six plants from each of four treatment groups in the crossed factored chamber experiment (Spring-Ambient, Spring-Elevated, Control-Ambient and Control-Elevated). RNA was extracted using a modified CTAB protocol (Taylor *et al*., 2005) and sequenced at the Instituto di Genomica Apilicata (IGA) as previously described.

Sequencing produced a total of 302.2 million raw reads which were quality checked using FastQC (Andrews, 2010). Trimmomatic v0.32 (Bolger *et al*., 2014) was used to trim adapters and other Illumina specific reads, bases at the beginning and end of a sequence that had a quality threshold below 5, and where the average quality per base of a sliding window dropped below 15. After trimming, reads less than 40 bp long were removed. The reference transcriptome was assembled *de novo* from three samples per treatment group

(total twelve) using the Trinity package and pipeline for subsequent gene expression analysis (Haas *et al*., 2013). The reference transcriptome was assembled from reads that were normalised both individually and upon concatenation using the script *insilico_read_normalisation.pl,* with maximum coverage 30, in order to increase efficiency while maintaining complexity (Haas *et al*., 2013). The assembled transcriptome comprised 160,279 transcripts in 100,890 components where transcripts represent different isoforms that form a component and components are loosely comparable to genes. The sequences of components were extracted from the assembled transcriptome and aligned to The Arabidposis Information Resource (TAIR10) (Berardini *et al*., 2015) using BLASTX (E<10) (Stephen *et al*., 1997).

For analysis of mRNA expression, RNA sequencing libraries for each individual were mapped back to the *de novo* assembled transcriptome using *run_RSEM_align_n_estimate.pl* and converted to coordinate sorted bam files. The script *abundance_estimates_to_matrix.pl* in the Trinity package was used to convert estimated component counts to TMM-normalised FPKM in order to normalise counts by total number of transcripts and transcript length. To visualise similarity of expression among samples, principal component analysis was conducted in R v3.5.1 (R-Core team, 2015) using the FactoMineR (Lê et al., 2008) package on all TMM normalised and log10+1 transformed transcript counts.

Transcripts that were differentially expressed between plants originating from different populations, under different growth [CO_2] or with a population x growth [CO_2] effect on expression were identified by implementing a generalized linear model (glm) with negative binomial error distribution in edgeR (McCarthy *et al*., 2012) in R v 3.5.1 (R-Core team, 2015). Transcripts were considered differentially expressed when FDR<0.05. Differentially expressed transcripts were functionally annotated and assigned to Gene Ontology (GO) categories using DAVID (Dennis *et al*., 2003; Huang, *et al*., 2007) and these were visualised using functions within the GOplot package in R (Walter *et al*. 2015).

4.3.4 Genome assembly

DNA was extracted from a *P. lanceolata* individual that originated as seed from the Bossoleto spring site and was grown in the glasshouse at the University of Southampton. DNA was extracted using a modified version of Doyle and Doyle (1987) CTAB extraction

protocol. DNA concentration was measured using Qubit DNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies, CA, USA). Fragment distribution of the DNA library was measured using the DNA Nano 6000 Assay Kit of Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Novogene Bioinformatics Institute, Beijing, China, performed paired end sequencing (with 350 base insert size) of the sample using an Illumina 2500 platform (Illumina, USA).

Sequencing produced 106 Gb of raw sequencing data which was quality checked using FastQC (Andrews, 2010) and filtered for contaminants with FastQ Screen v0.11.3 (http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/). Raw sequence reads were trimmed with Trimmomatic v0.36 (Bolger *et al*., 2014), which removed Illumina adapters and leading and trailing N bases with quality below 3, trimmed sequences with a low quality score (Phred +33) and removed reads <36 bases in length. KmerGenie (Chikhi and Medvedev, 2013) identified optimal kmer length for *de novo* assembly as k=105 (Chikhi and Medvedev, 2013). Further error correction was performed using SOAPec using *KmerFreq_HA* and *Corrector_HA* with maximum kmer size (Luo *et al*., 2012). SOAPec corrects errors based on kmer frequency spectrum such that low frequency kmers are removed. Error corrected reads were assembled using ABySS (Simpson *et al*., 2009). Basic statistics on the assembly were called using the script stats.sh in the BBMap package (Bushnell, 2015). Genome completeness was assessed using BUSCO v3 (Simão *et al*., 2015). The assembled genome consisted of 4,071,019 contigs with a maximum length of 1.3 Mb. We filtered the genome to those larger than 2 kb for further analyses using seqtk (Li, 2012) and assessed genome completeness of this filtered version using BUSCO v3 (Simão *et al*., 2015)

Gene prediction

Genes were predicted *ab initio* from the 2 kb filtered basic *P. lanceolata* genome contig 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 with AUGUSTUS (Stanke and 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 via Maker-derived AED scores, and by

protein domain conservation using InterProScan v5.30 (Jones *et al*., 2014) with PANTHER v12.0 (Mi *et al*., 2012) and PFAM v31.0 (Finn *et al*., 2015) to query predicted genes against known protein domains.

Transposable element prediction

Preliminary identification of repeat regions including transposable elements was conducted on the 2kb size filtered genome assembly using RepeatMasker v4.0.7 (http://www.repeatmasker.org/) and the RepeatMasker libraries (http://www.girinst.org/server/RepBase/).

Chloroplast assembly

In order to assemble the chloroplast genome the raw whole genome sequences were subsampled to 30 million reads and converted to fasta format. The chloroplast genome was assembled from these sequences using Novoplasty v2.6.3 (Dierckxsens, *et al*., 2017) with the chloroplast gene *rbcL* from *P. lanceolata* as a seed sequence (insert range=1.3; coverage cutoff =10,000) (Olmstead and Reeves, 1995). Two assemblies were produced differing in the orientation of an inverted repeat, however comparison to chloroplast sequences of closely related *Plantago media* (Zhu *et al*., 2016) suggests one is a more likely orientation. The chloroplast genomes were annotated by comparison to *Mimulus guttatus* and visualised using the GESeq (Tillich *et al*., 2017) and OGdraw (Lohse *et al*., 2013) components of CHLOROBOX (https://www.mpimp-golm.mpg.de/chlorobox).

4.3.5 Whole genome (WGS) and whole genome bisulfite sequencing (WGBS)

DNA for WGS and WGBS was extracted from the same individuals that were used to extract RNA for transcriptome assembly using a modified version of the Doyle and Doyle (1987) CTAB protocol. Extracted DNA was cleaned using carboxylate magnetic beads (Sera-Mag Magnetic Beads, GE Healthcare Life Sciences, Chicago, IL, USA) as per the manufacturer's instruction.

All library preparation and sequencing was carried out by the DNA Technologies and Expression Analysis Cores at the UC Davis Genome Center (CA, USA). DNA quality and quantity was analysed using a Bioanalyser (Agilent Technologies, Santa Clara, CA, USA). For whole genome sequencing, DNA was sheared by sonication on the Covaris E220

system (Covaris, MA, USA) and libraries were prepared using the KAPA library preparation kit (Kapa Biosystems, MA, USA). Bisulfite conversion was carried out using the Zymo EZ DNA Methylation-Lightning kit (Zymo Research, CA, USA) and bisulfite converted reads were then prepared using the TruSeq DNA Methylation kit (Illumina, CA, USA). Whole genome sequencing libraries were sequenced as 150 bp paired end reads, while whole genome bisulfite sequencing libraries were sequenced as 150 bp single end reads with a 20 % PhiX spike-in to balance nucleotide composition for sequencing. Both libraries were sequenced on an Illumina HiSeq4000 sequencing platform (Illumina, CA, USA) and resulted in 216 Gb of WGBS data and 245 Gb of WGS data.

Reads mapping to *Homo sapien*, *Escherichia coli*, *Saccharomyces cerevisiae* and *Enterobacteria phage phiX* were removed from the WGBS and WGS data set through FastQ Screen (http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/).

4.3.6 Conversion efficiency analysis

In order to calculate the conversion efficiency of the bisulfite conversion we utilised the fact that the chloroplast genome should be unmethylated to identify the number of unmethylated cytosines that were not converted to thymine (Fojtová, *et al*., 2001). Reads for each sample were subsampled to 800,000 reads and mapped to the chloroplast genome using Bismark with parameters –L 20, --score_min L,0,-0.4 (Krueger and Andrews, 2011). Conversion efficiency was calculated as the percentage of cytosines in reads mapping to the chloroplast that were called as methylated.

4.3.7 WGS analysis and variant calling

Clean reads were aligned to a 2 kb size filtered version of the genome using Bowtie2 (Langmead and Salzberg, 2012). These contigs were stitched together into a 'superscaffold' with N's greater than the length of the longest contig using ScaffoldStitcher (Haj *et al*., 2016). This decreased computational resources for processing of the alignment maps with the Genome Analysis Tool Kit (GATK) pipeline (Van der Auwera *et al*., 2013). We removed PCR duplicates and added read group (sequencing run) information to alignment files using Picard Tools v2.8.3 and then used GATK v3.7 to realign around indels. We generated Genomic VCF (GVCF) files for each sample using the GenotypeGVCFs module of GATK v3.7. We extracted SNPs and called high quality variants

(both homozygote and heterozygote sites) by filtering out SNPs that met any of the following criteria; mapping quality <20, variant confidence divided by the unfiltered depth of reference samples <3, Phred-scaled p-value of Fisher's test to detect strand bias <60, Strand Odds Ratio>3, Read Position Rank Sum test <-5 or >10 and Mapping Quality Rank Sum test <-10 (see Appendix Figure C. 1 and Appendix Figure C. 2 for justification of these parameters). Singleton SNPs were removed using vcftools with minor allele frequency of 0.03, and SNPs were filtered to include only those with identity information in at least eight of twelve individuals per population to facilitate downstream analysis of sequence divergence. We visualised variation as a PCA using smartPCA within the Eigensoft package (Price *et al*., 2006). Vcftools was used to calculate basic population genetic diversity statistics, nucleotide diversity, Tajima's D and genome-wide Fst, while both absolute (D_{xy}) and relative (Fst) measures of sequence divergence in 10 kb complete sliding windows with a step size of 1 kb were calculated using a set of python scripts (https://github.com/simonhmartin/genomics_general) (Martin, 2017).

Identification of outlier FST using Bayescan

Bayescan v2.1 (Foll and Gaggiotti, 2008) was used to identify regions with high differentiation between spring and control populations. Putatively divergent SNPs were identified using a Bayescan q-value threshold of 0.05 and Bayescan was run discarding the first 50,000 iterations, and outputting every 10^{th} iteration to a total of 5,000 outputted iterations. Chain convergence was confirmed by examining trace plots, using Heidelberg and Welch's convergence diagnostic, and by checking for autocorrelation in the R package CODA (Plummer *et al*., 2006).

4.3.8 **WGBS processing**

Due to the high levels of fragmentation in the assembled *P. lanceolata* genome we aligned all whole genome bisulfite sequencing reads to this single reference rather than making a reference for each individual. This approach has the advantage of being computationally tractable and is widespread in the literature (Schönberger *et al*., 2016; 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 when these are bisulfite converted. Using whole genome sequencing to identify SNPs (described earlier) we found that 12.4 % of allegedly

differentially methylated sites (DMSs) were incorrectly called as a result of the presence of a C to T SNP. Methylation contexts were affected differently, with just 2.8 % of CHH DMSs called erroneously, 14.7 % for CG and 12.5 % for CHG contexts. This exemplifies that alignment of bisulfite sequencing reads to a reference genome other than the genome of the individual can erroneously inflate the number of (allegedly) differentially methylated sites called. In our case the miscalled sites were also clustered across the genome so that they significantly affected the calling of differentially methylated regions (DMRs). We acknowledge that there may be sites at which there are both C-T polymorphisms and differential methylation that are not identified by this approach, however, exclusion of sites at which there are polymorphisms at least gives a conservative estimate of differential methylation.

To analyse methylation patterns in *P. lanceolata* the 2 kb filtered genome was *in silico* bisulfite converted to forward and reverse read versions using the *bismark_genome_preparation* module in Bismark v0.19.0 (Krueger and Andrews, 2011) and utilising Bowtie2 v2.3.1 alignment (Langmead and Salzberg, 2012). Trimmed and filtered WGBS reads were similarly converted *in silico* and mapped to the converted 2 kb filtered genomes with Bismark, with a maximum of 20 mismatches and specifying the minimum alignment score as a function of read length with parameter --score_min L,0,-0.4 (Krueger and Andrews, 2011). The *bismark_methylation_extractor* component of Bismark was used to extract methylation calls for each cytosine and write out context dependent methylation calls into separate files. Following visual inspection of M-bias plots of average methylation across the length of the read we identified that the ten bases at the 3’ end of the read and the 3 bases at the 5’ end of the read were erratically methylated, and since this is likely a sequencing artefact they were ignored in the methylation call. Genome-wide weighted methylation was calculated as described by Neiderhuth *et al*., (2016).

4.3.9 Identifying differential methylation

Differential methylation was analysed through two separate approaches, firstly as tiles to identify genome-wide patterns of differential methylation, and secondly individual sites were analysed to identify treatment associated differentially methylated sites and regions (Eichten *et al*., 2016; Ganguly *et al*., 2017).

A. 1000 bp window analysis in methylKit

Methylation call files were converted to coverage and bed files and sorted by chromosome with bismark2bedgraph (Krueger and Andrews, 2011). Coverage files were converted to report files using the *coverage2cytosine* component of Bismark (Krueger and Andrews, 2011), and these report files then read into Methylkit (Akalin, A., 2016) using the *readBismarkFiles* function in R v 3.4.1 (R Core Team, 2013). 1000 bp tiles present across all samples were identified using *filterByCoverage*, *tileMethylCounts* and *unite* functions in methylKit. Using the *calculateDiffMeth* function, a chi-squared teat was implemented to identify differential methylation of 1000 bp tiles between any pairwise treatment groups, using Benjamini-Hochberg correction with maximum FDR corrected p-value <0.00001 and a minimum methylation difference of 5 % (Akalin *et al*., 2012, Benjamini and Hochberg, 1995). Percentage methylation of these differentially methylated tiles were visualised as heatmaps in the R package pheatmap (Kolde, 2013).

B. DMS identification in DSS

For a more stringent identification of DMS and DMRs, the R package DSS (Wu *et al*., 2015, Park and Wu, 2016) was employed to analyse the methylation calls produced by the *bismark2bedgraph* script in Bismark. Given an experimental design matrix and methylation calls for each sample in the form of counts, DSS implements algorithms for the identification of differential methylation using the dispersion shrinkage method and Wald tests assuming a beta-binomial error distribution. Downstream analyses were conducted using DMSs filtered by FDR<0.05 and DMRs identified by the *callDMRs* function and filtered by FDR<0.05.

4.3.10 SNP, DMR and DE component comparison

For analysis of the co-location of differentially expressed transcripts to regions of the genome with DMSs and SNPs we mapped transcripts to the 2 kb filtered genome with BLAST and retained the match with the highest e-value, where that evalue was < 1e-04 and the transcript had >95 % similarity to the genome. Of the 160,729 transcripts assembled, 44 % mapped to the 2 kb filtered genome under these parameters.

For analysis of exon methylation, we extracted the start and end positions of transcripts that mapped to the 2 kb filtered genome and used R packages methylKit (Akalin *et al*.,

2012) and GenomicRanges (Lawrence *et al*., 2013) to extract methylation differences across these regions. The average methylation difference between spring and control populations across these regions was plotted against the log fold change in expression of the transcript in spring versus control plants. To further examine the relationship between gene methylation and gene expression we extracted the positions of gene features predicted by maker and plotted methylation of those features against the expression of all transcripts mapping to the gene.

The number of DMS in complete 10 kb windows with 1 kb step size was plotted against both D_{xy} and Fst calculated for that window to explore the correlation between sequence divergence and differential methylation.

Finally, we cross-referenced the location to which differentially expressed transcripts map with the location of DMSs and F_{ST} outlier SNPs. This enabled us to identify candidate genes involved in the single and multigenerational response to elevated [CO_2] that could be regulated by changes in methylation or that could be the target of selection.

4.4 Results

4.4.1 Genome assembly and feature prediction

Short read genome assembly produced a genome of 1,425,357 kb of sequence across 4,075,744 contigs, with N50 of 1.8 kb (Appendix Table C. 1). In order to analyse trends across the genome, the basic genome contig assembly was filtered to contigs greater than 2 kb which included 13.7 % of the sequence across 29,103 contigs (Appendix Table C. 2 and Appendix Figure C. 3). BUSCO (Simão *et al*., 2015) analysis of genome completeness identified 72.9 % of 1440 single copy orthologues from *Arabidopsis thaliana* in the full basic assembly and this dropped to 63.2 % when filtered to contigs larger than 2 kb (Appendix Figure C. 4). 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 *et al*., 2014) and 53 % identified as having a recognisable protein domain in InterProScan 5 (Jones *et al*., 2014). We considered predicted genes with AED<0.5 as sufficiently supported to allow for analysis of general trends of methylation across gene features, but acknowledge that gene prediction was generally poor, likely as a result of the high fragmentation of the contig assembly and a lack of protein sequence evidence from *Plantago* or closely related species for training. RepeatMasker identified 895 transposons, and the chloroplast was assembled as two alternative sequences of 149.5 Kb, differing in the orientation of an inverted repeat sequence (Appendix Figure C. 5).

4.4.2 Low overall genetic differentiation between spring and control derived populations

Population genomic analysis of SNP's in whole genome sequences supported previous findings of very little genetic differentiation between spring and control site populations (Watson-Lazowski *et al*., 2016) with F_{ST} = 0.050 ± 0.125 (mean ± standard deviation) across the genome and similar in exon regions (Table 4.1). Nucleotide diversity was significantly higher in the spring population across the whole genome (π_C =0. 272 ± 0.180 vs π_S=0.275 ± 0.179; t-test; t=-16.3, p<0.0001) and across exon regions (Table 4.1). The exon regions of the spring population had a negative Tajima's D, which indicates some purifying selection (D_C =0.053 ± 0.925 vs D_S=-0.073 ± 0.952). Tajima's D was positive in both populations across the whole genome indicating low levels of extreme frequency

polymorphisms characteristic of balancing selection or a decrease in population size. Comparing the two populations, Tajima's D was higher in the control population than the spring population both across the whole genome and in exon regions (Table 4.1). The most parsimonious explanation for this pattern of differentiation is that the spring population originated from the control population and subsequently underwent weak purifying selection in exon regions.

				T-test	
Population statistic	**Region**	**Population**	**Estimate**	**t-statistic**	**P-value**
F_{ST} across whole genome	Whole genome		0.050 ± 0.125		
F_{ST} across exon regions	Exon regions		0.048 ± 0.112		
Nucleotide diversity (π)	Whole genome	Control	0.280 ± 0.141		
		Spring	0.264 ± 0.144		
				4.26	<0.0001
	Exon regions	Control	0.280 ± 0.141		
		Spring	0.264 ± 0.144		
				11.14	<0.0001
Tajima's D	Whole genome	Control	0.200 ± 0.948		
		Spring	0.174 ± 0.989		
				14.7	<0.0001
	Exon regions	Control	0.053 ± 0.925		
		Spring	-0.073 ± 0.952		
				10.24	<0.0001

Table 4.1: Population genomic analysis statistics calculated from SNPs for whole genome and exon-only regions.

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.

4.4.3 Divergence between the response of spring and control derived plants to elevated [CO_2]

Phenotypic differences in plants originating from the spring versus control site populations, and growing in ambient versus elevated [CO_2] in this experiment have previously been discussed in Watson-Lazowski *et al*., 2016. Of the eight phenotypic traits measured in this experiment (Watson-Lazowski *et al*., 2016), two (epidermal cell number and above ground biomass) responded plastically to elevated [CO_2] but showed no evidence of adaptation. However the expression of three traits (stomatal index, single leaf dry mass and area) was significantly affected by the population site of origin, and these fixed differences may be the result of adaptation.

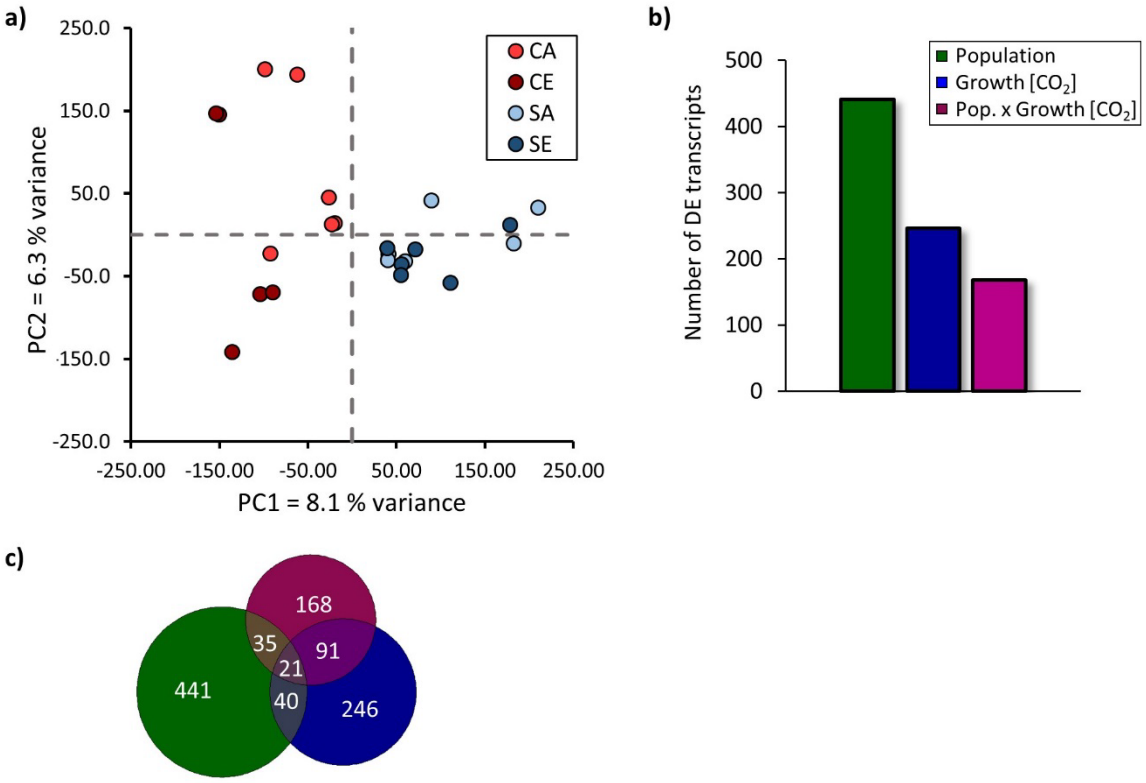


Figure 4.1 Differential expression of transcripts in individuals from four treatment groups.

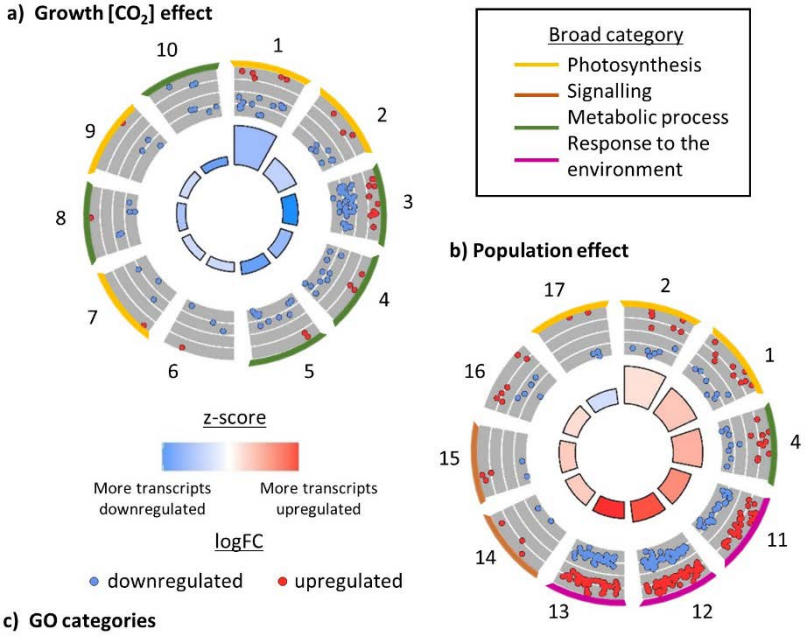
a) Number of differentially expressed (DE) transcripts associated with population site of origin (spring vs control), growth [CO_2] (ambient vs elevated) and the interaction between the two **b)** Expression profiles (PCoA) of the 24 plants, CA=Control Ambient, CE=Control Elevated, SA=Spring Ambient, SE=Spring Elevated. **c)** Overlap between the number of differentially expressed genes associated with population site of origin, growth [CO_2] or an interaction between the two.

A total of 160,279 transcripts were assembled in 100,890 components in Trinity, and 89 % of transcripts had a BLASTN match with e-value <1e-04 to the transcriptome assembled in Watson-Lazowski *et al*., (2016). BUSCO (Simão *et al*., 2015) analysis of transcriptome completeness identified 85.5 % of 1440 single copy orthologues from *Arabidopsis thaliana* (Appendix Figure C. 6). Principal coordinate analysis (PCoA) indicated that transcriptome expression profiles differed more between population site of origin (Spring or Control) than between growth [CO_2] exposure (Ambient of Elevated) (Figure 4.1a). This was also seen when we identified differential expression; there were more differentially expressed (DE) transcripts between populations than there were between growth [CO_2]

treamtents (Figure 4.1b). Further, the expression profiles of plants originating from the spring site were more similar across growth [CO_2] treatments than those from the control site, suggesting less diversity in transcript expression response to changes in growth [CO_2] in the spring population than the control population (Figure 4.1). Some overlap between multigenerational and single generational response to elevated [CO_2] was evident, with 14 % of genes that were DE in elevated [CO_2] vs ambient [CO_2] also DE in spring vs control populations, significantly more than would be expected by chance alone (chi-squared test, X^2=36, df=1, p<0.0001).

Transcripts were more frequently downregulated rather than upregulated by growth at elevated [CO_2] with all of the top ten enriched GO categories having more downregulated than upregulated transcripts. In contrast, transcripts DE by population site of origin were more likely to be upregulated in the spring plants relative to those from the control site (Figure 4.2). The same two GO categories; Photosynthesis and Photosynthesis light reaction, were the most highly represented in both the growth [CO_2] and population expression response. Additionally, 50 % and 55 % of transcripts that were identified as DE in these two categories were DE both by population and by growth [CO_2]. Across all GO categories there were 40 transcripts that were identified as DE both by growth [CO_2] and by population, and all of these showed the same directional response in expression in elevated versus ambient [CO_2] and in spring versus control populations. Of the top ten GO categories associated with population site of origin and growth [CO_2], those involved in response to the environment were only found associated with the population.

Only 20 % of the transcripts identified as DE in this analysis mapped with BLASTN (e-value <1e-04) 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 are supported in both studies, including evidence of more differential expression between populations relative to growth [CO_2] and enriched GO categories. Further, transcripts described in more detail later were differentially expressed in both analyses.



Number	ID	GO Category	Adjusted P value a)	Adjusted P value b)
1	Photosynthesis	GO:0015979	1.5e-10	8.4e-06
2	Photosynthesis, light reaction	GO:0019684	3.1e-06	8.2e-07
3	Organonitrogen compound metabolic process	GO:1901564	4.0e-05	
4	Generation of precursor metabolites and energy	GO:0006091	1.4e-04	1.6e-05
5	Cofactor metabolic process	GO:0051186	1.5e-04	
6	Reductive pentose-phosphate cycle	GO:0019253	7.5e-04	
7	Photosynthesis, dark reaction	GO:0019685	8.7e-04	
8	Chlorophyll metabolic process	GO:0015994	9.6e-04	
9	Photosynthesis, light harvesting photosystem I	GO:0009768	10.0e-04	
10	Pyridine-containing compound metabolic process	GO:0072524	1.3e-03	
11	Response to abiotic stimulus	GO:0009628		3.4e-04
12	Response to stimulus	GO:0050896		4.2e-04
13	Response to stress	GO:0006950		1.9e-03
14	Carbohydrate mediated signalling	GO:0009756		2.4e-03
15	Sugar mediated signalling pathway	GO:0010182		3.0e-03
16	Cofactor biosynthetic process	GO:0051188		3.2e-03
17	Photosynthesis, light harvesting	GO:0009765		4.4e-03

Figure 4.2: Gene ontology (GO) analysis of differentially expressed transcripts identifies key biological processes that show differential expression under single generation or multigenerational growth under elevated [CO_2].
a) Circular plot highlighting differential transcript expression in the ten GO categories that encompass the most differentially expressed genes between growth [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 upregulated in elevated [CO_2] and blue down regulated. 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 are up or downregulated. The colour of the outermost lines indicate broad categorisations of the GO terms **b)** Differentially expressed GO ontology categories in plants originating from spring versus control sites **c)** GO term identifiers for those categories visualised in a) and b).

4.4.4 Methylome remodelling under single generation exposure to elevated [CO_2] but many more differences in methylation profiles between populations

WGBS produced a total of 2,712 million single end reads across the 24 samples, averaging 113 million reads per sample (30X coverage). Conversion efficiency was > 98.5 % for all samples (Appendix Table C. 3). Methylation of cytosines in all contexts was generally high with 83 %, 70 % and 15 % of cytosines methylated in the CG, CHG and CHH contexts respectively. Methylation was depleted at the transcription start site of predicted genes and CHG and CHH methylation were also depleted across the gene body (Figure 4.3). Methylation in the CG context increased to near non-genic levels across the gene body and was depleted again at the transcription end site. We did not see an increase in methylation over predicted transposons but whether this was because of poor transposon prediction is unclear (Figure 4.3).

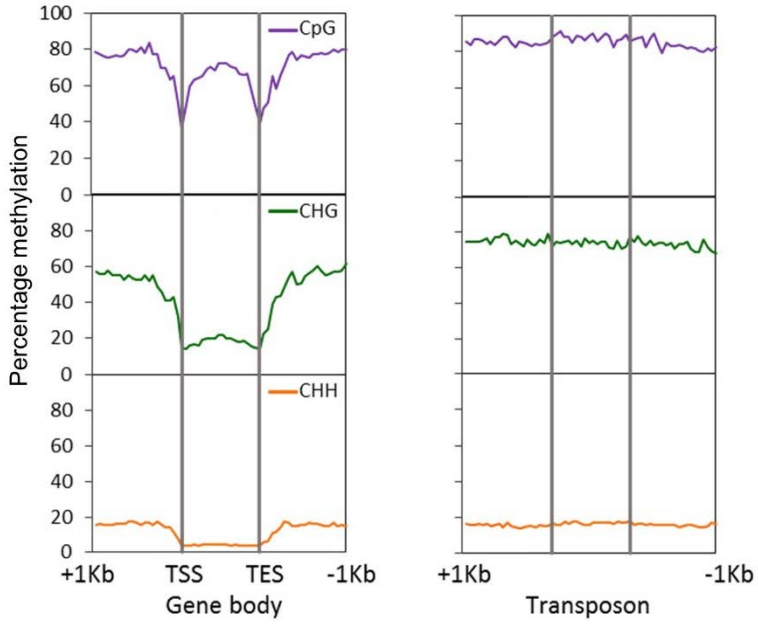


Figure 4.3: Mean methylation of predicted gene features in each methylation context.
Percentage methylation 1 kb upstream and 1 kb downstream of the gene features are calculated in 50 bp intervals, and across the gene body and transposon in 5 % intervals. TSS=transcription start site, TES=transcription end site.

The largest variation in methylation across the genome was seen between populations rather than associated with growth [CO_2] treatment (Figure 4.4a), with the exception of one spring plant individual which had a methylation profile anomalously similar to the control derived plants. There was also evidence of remodelling of the methylome in response to elevated [CO_2] in control plants (Figure 4.4). In an overview of whole genome methylation in 1 kb tiles, when clustering the 24 individuals based on similarity of CG

methylation, there was clustering by growth [CO_2] in the control population but not the spring. To a lesser extent this was also seen for CHH methylation profile clustering, but not seen for CHG. In contrast the methylation profiles of plants from the spring population did not cluster by similarity according to their growth [CO_2] conditions and were more likely to cluster by relatedness in all three sequence contexts. This could indicate reduced responsivity of especially the CG methylome to elevated [CO_2] in plants originating from the spring site population.

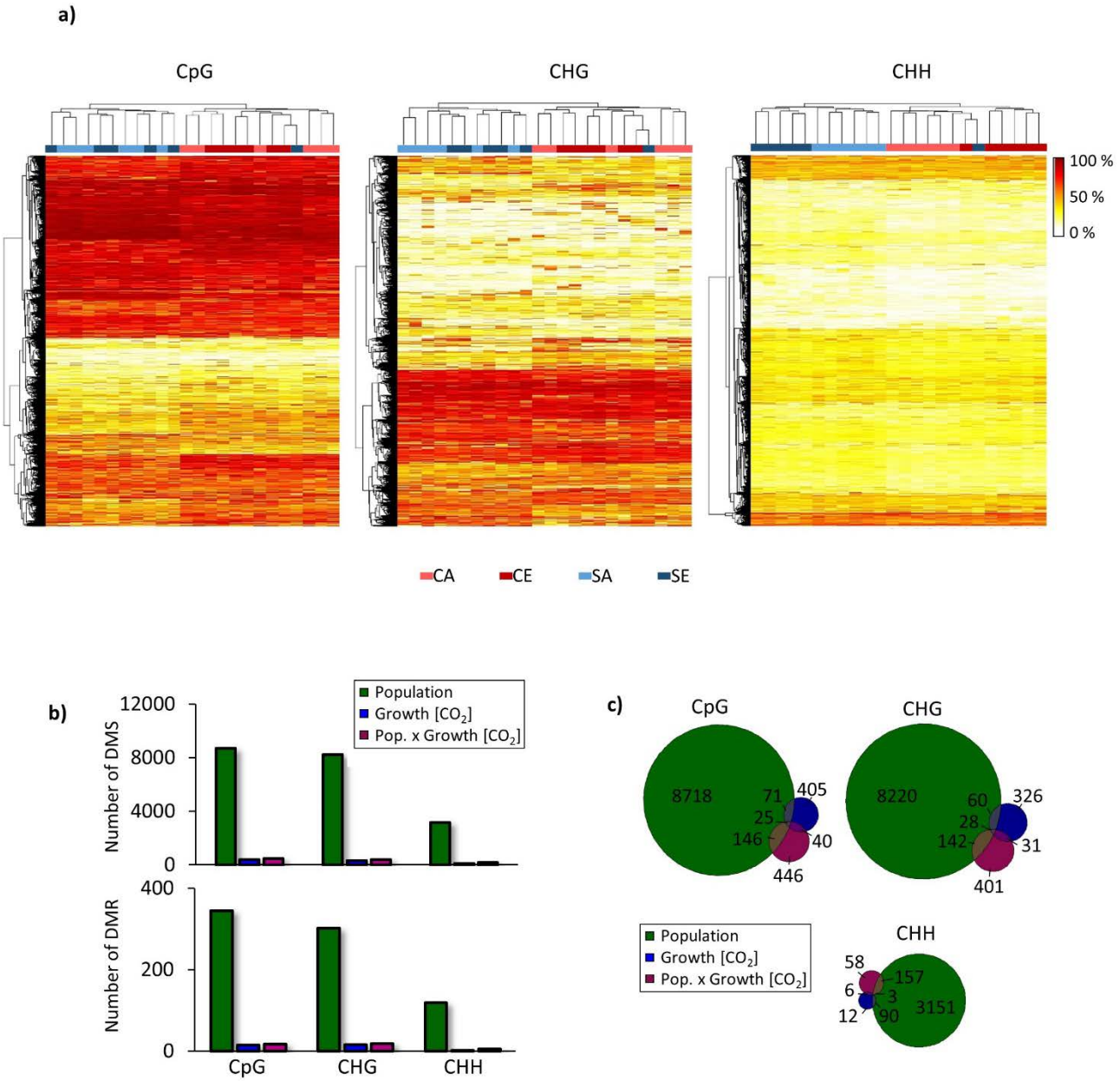


Figure 4.4: There are more differentially methylated sites and regions associated with individuals that originate from different populations (S or C) than there are associated with growth [CO_2] treatment (A or E).

a) Heatmaps of average methylation of differentially methylated pre-specified 1000 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 differentially methylated sites (DMS) and regions (DMR) as identified by R package DSS **c)** The overlap of DMS that were significantly differentially methylated in more than one category.

We identified a total of 21,914 differentially methylated sites (DMSs) associated with growth [CO_2], population or an interaction between these effects. Of DMSs identified in all contexts 92 % were associated with differences between the two populations and were disproportionately represented among cytosine contexts (chi-squared test, population; X^2=30636, df=2, p<0.001, growth [CO_2]; X^2=1574.5, df=2, p<0.001, population x growth [CO_2]; X^2=1564.3, df=2, p<0.001) with most DMSs being in the CG or CHG context (Figure 4.4). In support of the conclusions from analysis of methylation variation in 1 kb tiles, there were around 20 times more DMSs associated with population site of origin than there were associated with growth [CO_2], and 18 % of the DMSs associated with growth [CO_2] were also associated with population site of origin (Figure 4.4) and this was significantly more than would be expected by chance (chi-squared test; X^2=196, df=1, p<0.0001).

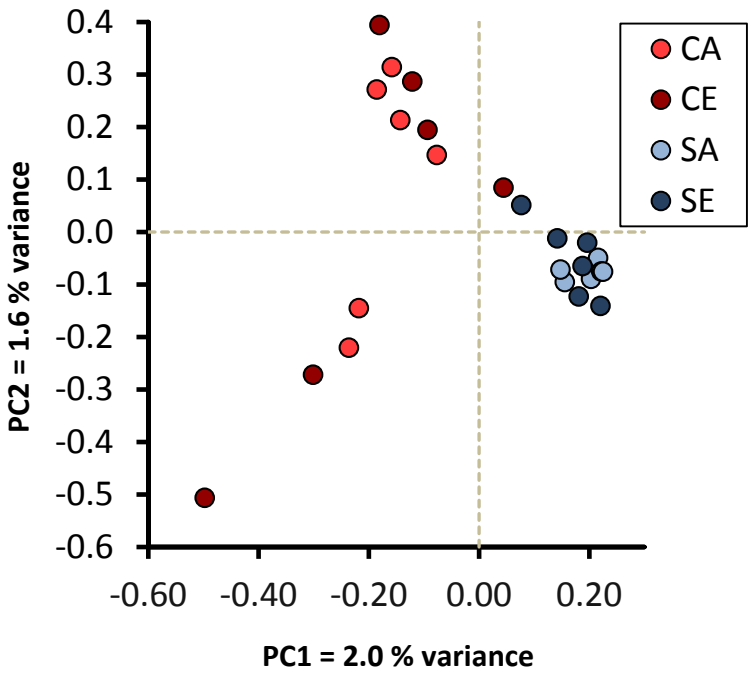


Figure 4.5: Principal component analysis of SNP variation

PCA for all 1.8 million single nucleotide polymorphisms between plants in this experiment (n=24, 6 reps per treatment group) with singleton SNPs removed and SNPs filtered to those that contain information for at least 8/12 individuals per population. 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.

We note here that 12.4% of identified DMS would have been called erroneously if the genome for each individual in this analysis had not been sequenced to identify C->T SNPs. Further, because the C->T SNPs between individuals were clustered in the genome this also would have resulted in the erroneous calling of DMRs. This serves to caution the interpretation of DMSs in the absence of WGS for each individual.

4.4.5 Correlations between genetic and methylome variation account for a small fraction of differences in methylation between populations

WGS produced 1,365 million paired end reads with an average of 57 million per sample (15X coverage). A total of 1.8 million SNPs were identified across the 24 individuals analysed after excluding SNPs for which less than eight individuals had a base call. Of these, 974 (0.05%) were identified as putative targets of divergent selection between the spring and control populations as F_{ST} outliers (Q<0.05). In PCoA the first PC explained 2.0 % of the variance, separating individuals by population and indicating less variation within the spring population (Figure 4.5). The second PC explained 1.6 % of the variation whereby related individuals were found close together, and again there was less variation within the spring population.

When the genome was fragmented into complete 10 kb tiles with a 1 kb step size (and excluding those <10 kb), there were significant positive correlations between the number of DMSs associated with population of origin and both the absolute (D_{xy}) and relative (F_{ST}) measure of sequence divergence across the tile (Figure 4.6). Correlation with sequence divergence explained 6-12 % of the number of DMSs in a 10 kb window across the three sequence contexts and this suggests that to some extent variation in methylation between the two populations aligns with genetic variation, whereby regions of the genome with high sequence divergence (that are potentially undergoing selection) are also more likely to harbour more DMSs.

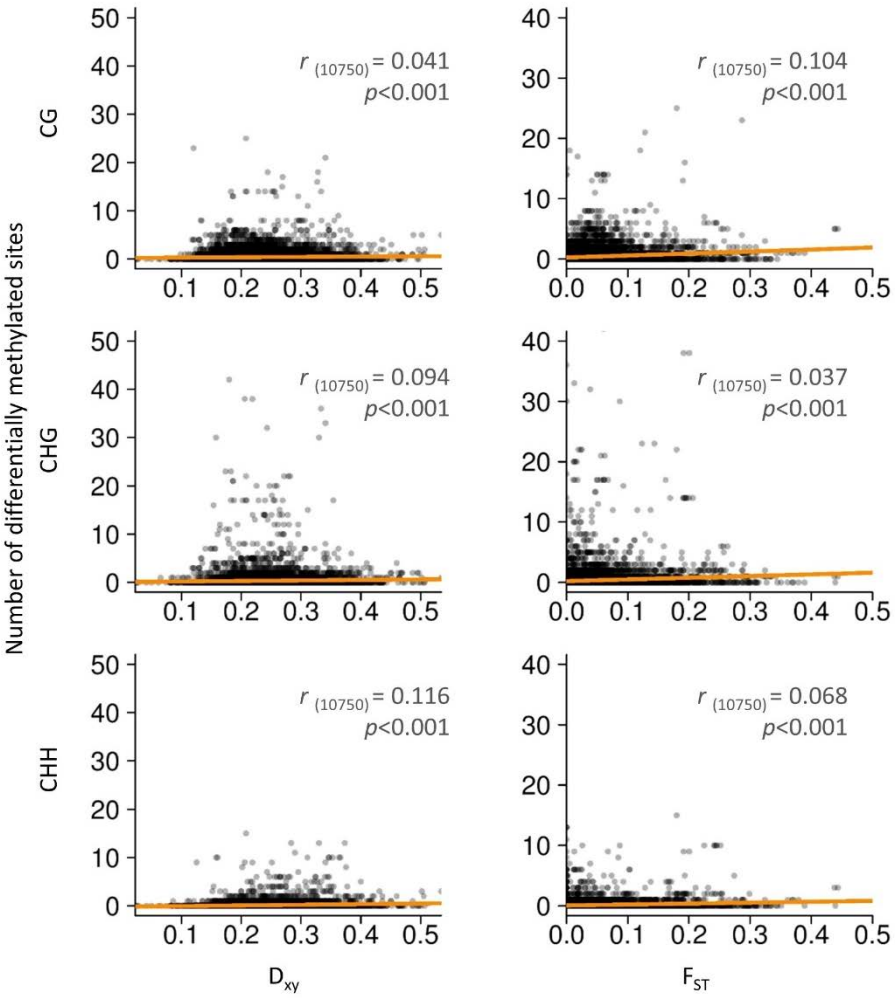


Figure 4.6: The relationship between the number of DMSs between spring- and control-originating individuals in complete 10 kb windows, and an absolute (D_{xy}; left) or a relative (F_{ST}; right) measure of sequence divergence for that window.

4.4.6 Co-location of differentially expressed transcripts to DMSs and SNPs highlight potential targets for differential methylation and selection

In total, 3871 transcripts mapped to within 1 kb of a DMSs (85 % of these transcripts) or an outlier SNP (20 %). Twenty-six DE transcripts mapping to twenty components (loosely comparable to 'genes') were identified as co-locating to within 1 kb of at least one DMSs (17 components) or SNPs (four components), with one DE component within 1 kb of both DMSs and outlier SNP (Appendix Table C. 4). All 20 of these components were DE both in this analysis of gene expression (using a glm to identify DE) and in the Watson-Lazowski *et al.,* 2016 analysis of this data (using pairwise t-tests between treatment groups). Components mapping to within 1 kb of a DMS were not enriched for those that were DE (chi-squared test, X^2=4.3, df=2, NS) and neither were genes mapping to within 1 kb of a SNP (chi-squared test, X^2=1.5, df=2, NS). We found no DMSs/outlier SNPs co-locating to within the exon region of a gene but a weak association between CHG methylation across the exon region and the expression of that region (Appendix Figure C. 7 and Appendix Table C. 5).

Two of the 20 components identified as co-locating to within 1 kb of a DMS or SNP were involved in the abscisic acid (ABA) pathway, which has been shown to be involved in plant response to elevated [CO_2] (Chater *et al*., 2015) (Figure 4.7). These were also the two DE components that were found within 1 kb of the most DMSs/outlier SNPs. One of these was annotated as matching an FLDH sequence (an NAD^+-dependent dehydrogenase that likely acts as a negative regulator of ABA signalling) and the other matched an RD22 sequence (a protein which confers increased resistance to drought and salt stress, is responsive to light, and is induced by ABA) (Yamaguchi-Shinozaki *et al*., 1993; Iwasaki *et al*., 1995; Goh *et al*., 2003). Details of the expression and DMSs/outlier SNPs within 1 kb for selected components that were DE are shown in Appendix Figure C. 8 and Appendix Figure C. 9 with the full list provided in Appendix Table C. 6.

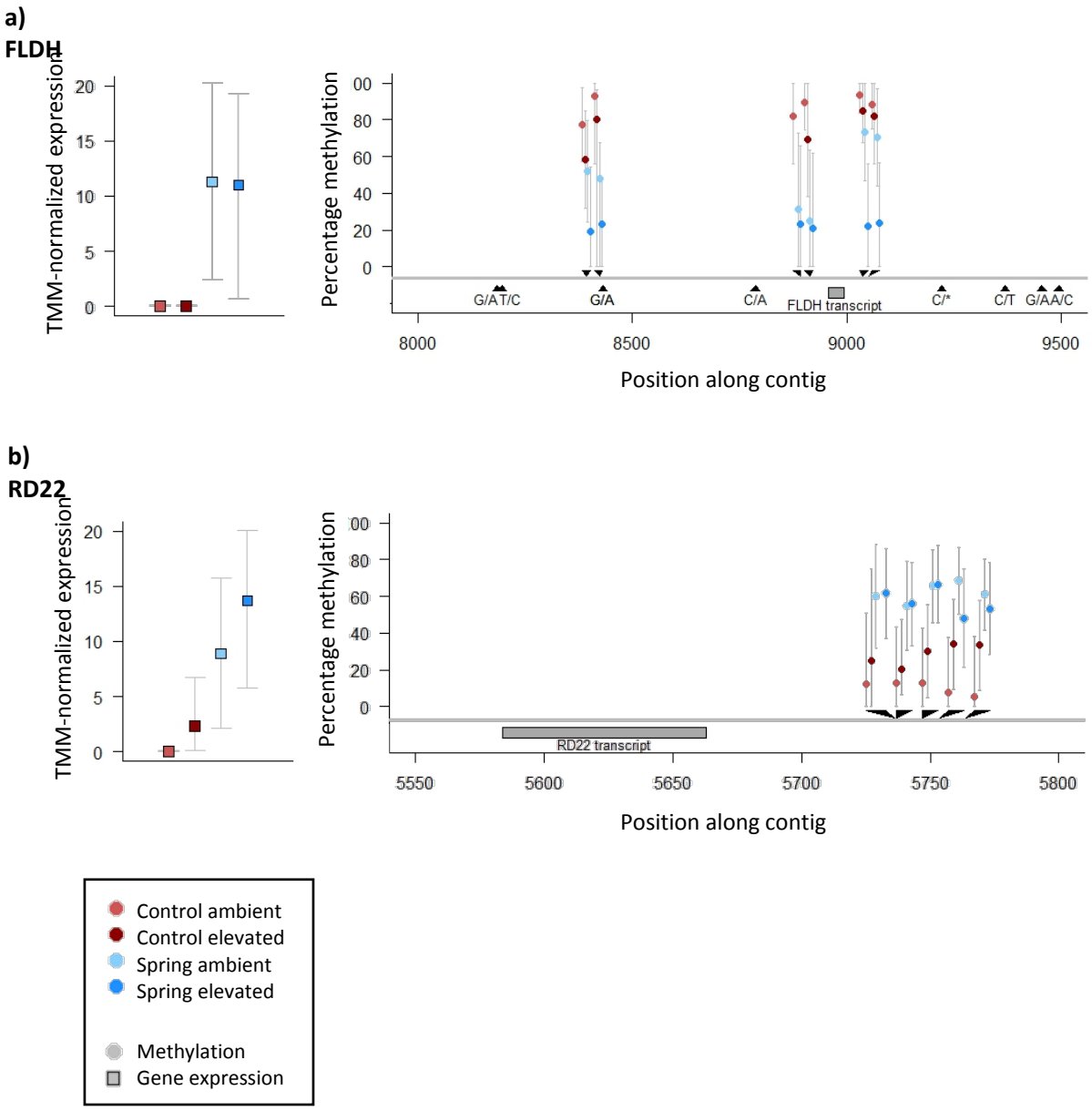


Figure 4.7: Differentially expressed transcripts that map to within 1 kb of a differentially methylated site or SNP, and are annotated as being involved in the abscisic acid pathway. TMM-normalized expression of the four treatment groups based on population site of origin (spring or control) and growth [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 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)** NAD(P)-binding Rossmann-fold superfamily protein (FLDH) **b)** BURP domain-containing protein (RD22).

4.5 Discussion

Understanding the multigenerational response of plants to rising [CO_2] is of critical importance to predict and manage the impact of climate change. However the challenge of growing plants for multiple generations in experimental conditions has limited research into if and how plants can adapt to elevated [CO_2]. To address this we used *Plantago lanceolata* plants growing at a natural CO_2 spring to elucidate the mechanistic basis for the single and multigenerational responses of a wild population of plants exposed to elevated [CO_2]. Previously it was noted that there was extensive transcriptome divergence between the control and spring derived populations, but very little genetic divergence (Watson-Lazowski *et al*., 2016). We build on the RNA-Seq data analysed in Watson-Lazowski *et al*., (2016) by reanalysing gene expression and combining it with whole genome and methylation sequencing datasets. Although we acknowledge that our reference genome is fragmented, we were still able to analyse methylation and sequence divergence in nearly half of the gene space, providing the first insight into the role of methylation in plant response to elevated [CO_2]. From this experiment we highlight a role for methylome remodelling in the plastic response of *P. lanceolata* to elevated [CO_2] and in the potential adaptation of the spring site population to naturally elevated [CO_2].

4.5.1 Low overall genetic 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 the transcriptome (Watson-Lazowski *et al*., 2016) and by calculations for divergence between populations of *P. lanceolata* in other studies (Bos *et al*., 1986; Tonsor *et al*., 1993; Van Dijk *et al*., 1988). Limited pollen-dispersal distance (1.5m; Bos *et al*., 1986; Tonsor, *et al*., 1993), passive seed dispersal (0.08m; Bos *et al*., 1986) and obligate outcrossing likely results in high within population genetic diversity and high genomic heterogeneity (Gáspár *et al*., 2018), such that interpopulation divergence is generally low. The identification of lower genetic diversity in the spring population suggested that the spring population likely originated from the control population and had undergone weak purifying selection in exon regions. Further, similar patterns of variation are seen when we compare the PCAs of SNP identity and of

gene expression. This implies that genetic variation does play a role in gene expression differences between spring and control populations.

4.5.2 Divergence between the response of spring and control derived plants to elevated [CO_2]: Evidence for adaptation

Analysis of phenotypes in this crossed factored experiment evidenced a plastic response to elevated [CO_2] with no evidence for adaptation in two of eight traits (Watson-Lazowski *et al*., 2016). Three of eight traits were significantly affected by population site of origin and this could be the result of adaptation to elevated [CO_2] in the natural CO_2 spring population. Adaptation to elevated [CO_2] has been inferred in growth, biomass and gas exchange traits in a number of other plant species at natural CO_2 spring sites (Barnes *et al*., 1997; Polle *et al*., 2001; Onoda *et al*., 2009; Nakamura *et al*., 2011). Reaction norms of biomass traits in spring relative to control plants additionally highlight a loss of plasticity in traits that are initially responsive to elevated [CO_2]. Loss of plasticity can occur through genetic accommodation, and this is thought to arise if plasticity is costly or selection is high, or through relaxed selection and genetic drift (Schlichting and Wund, 2014). Genetic assimilation has previously been employed to explain reduced plasticity in leaf stomatal conductance of *Bromus madritensis* L. following just seven years growth at elevated [CO_2] (Pigliucci *et al*., 2006; Grossman and Rice, 2014). However the causal genetic changes responsible were not identified in this study, and it is possible that epigenetic inheritance could also facilitate a loss of trait plasticity.

Fewer gene expression differences associated with the spring population response to growth [CO_2] relative to the control population also suggests that the spring population could be adapted to elevated [CO_2]. The overlap of genes that are DE between [CO_2] treatments and between populations (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 is a result of adaptation to the elevated [CO_2] conditions at the spring site. Further since the direction of expression differences between control versus spring populations and between ambient vs elevated [CO_2] is consistent for all transcripts that were DE in both categories, the regulation of these transcripts could have been a target for selection. We did not extensively measure photosynthetic rate in this experiment and therefore cannot directly link differences in photosynthesis with

differential regulation of genes involved in this pathway. Nevertheless, differences in photosynthetic rate would be expected to have downstream consequences on plant functioning, and we observed potentially adaptive fixed differences in biomass traits in spring versus control plant populations.

Transcripts that were DE by population but not growth [CO_2] could be related to local adaptation conditioned by other environmental differences between sites, for example, pH differences of the soil (Körner and Miglietta, 1994) or altered microbial activity (Krüger *et al*., 2011; Šibanc *et al*., 2014; Šibanc *et al*., 2018). Alternatively these fixed differences in gene expression may have arisen as spontaneous genetic or epigenetic mutations to gene regulatory elements that were fixed in the population through selection or drift.

4.5.3 Methylome remodelling under single generation exposure to elevated [CO_2] but many more differences in methylation profiles between populations

A total of 821 differentially methylated sites (DMSs) and 33 differentially methylated regions (DMRs) were associated with growth at elevated [CO_2] in this analysis. Environmentally induced methylation changes in wild non-model plants have previously been associated with drought (Rico *et al*., 2014; Liang *et al*., 2014; Alsdurf *et al*., 2016) and light environment (Baker *et al*., 2018), but this is the first time the methylome has been shown to be responsive to elevated [CO_2]. These differences could play a significant role in coordinating plant response to elevated [CO_2] if they impact the regulation of genes (Chinnusamy and Zhu, 2009). Whole genome DNA methylation analyses combined with gene expression data have identified differential methylation associated with environmental change that correlate to the expression of critical response genes (Garg *et al*., 2015; Yong-Villalobos *et al*., 2016). In this analysis we failed to detect genes in proximity (1 kb) to DMS associated with growth [CO_2]. This may in part be due to the limitations of this analysis to detect co-location of DMS/outlier SNP and DE transcripts. The genome assembly used for this analysis only facilitated the analysis of a proportion of the genome (13 % of the entire genome, with 44 % of the transcriptome mapping under selected parameters) where co-location of DMSs/outlier SNPs and DE genes will be missed if they map to different contigs. Additionally, this co-location approach can only detect where methylation could impact gene regulation through a *cis* acting mechanism and not those that act in *trans* (Neiderhuth and Schmitz, 2017). Finally identification of

co-locating genes and DMSs/outlier SNPs is limited by a lack of robust estimates for linkage disequilibrium (LD) in *P. lanceolata*, where a 1 kb estimate of average LD decay is likely to be too high considering that *P. lanceolata* is an outcrossing species, but an average of 500 bases may be conservative. Further LD is known to be highly heterogenous across the genome (Gupta *et al*., 2005).

Broad scale trends in methylation associated with growth [CO_2] suggested that there could be reduced responsivity of the methylome to elevated [CO_2] in the plants originating from the spring site population. Within the control plant populations methylation profiles were more similar between individuals exposed to the same growth [CO_2] than between families for CG (and to a lesser extent CHH) while in the spring plant population methylation profiles clustered to a greater extent by family.

We identified substantial variation in methylation patterns between populations and limited genetic variation, and DE transcripts more frequently co-located to DMSs than to SNPs. The large number of DMSs associated with population (and the low number of SNPs) suggests that there is greater divergence in DNA methylation profiles than in sequence polymorphisms between the two populations.

This highlights a role for indirect genetic effects in adaptation to elevated [CO_2], although this analysis explicitly considers only the potential role of epigenetic inheritance of DNA methylation changes. Other indirect genetic effects, such as maternal effects on seed provisioning, may also facilitate adaptation to elevated [CO_2]. These were not considered here due to their potential to confound interpretation of adaptation as phenotypic differences resulting from direct genetic effects and epigenetic inheritance. Further, this also means that parental effects mediated by epigenetic inheritance of DNA methylation states are likely to be be underestimated here, where a proportion are reversed during a generation of growth at ambient [CO_2].

4.5.4 Correlations between SNP and methylome variation only account for a small fraction of the large differences in methylation between populations

We found that 6-12 % of the variation in number of DMSs in a 10 kb window was explained by a positive correlation with sequence divergence of that window (as calculated from SNPs). 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 do not consider the relationship between insertions, deletions or larger structural variants with variation in methylation profiles, which may provide further insight (Schmitz *et al*., 2013; Kawakatsu *et al*., 2016, Eichten *et al.,* 2016). This may be particularly important considering the role of methylation in silencing transposons (Saze *et al*., 2012) and the link between transposon activity and mutation rates (Wicker et al., 2016). Further, the short read assembly of a highly heterozygous individual will miss repeat rich regions (Whiteford *et al*., 2005; Alkan *et al*., 2010). Analysis of the correlation between methylation profiles and indel or structural variation could be achieved with the existing dataset, but will require a more complete genome assembly from long read sequencing. However this may still only be expected to explain a relatively small proportion of the variation in methylation profiles between the two populations since SNP variation and indel variation are shown to co-vary (Zhang et al., 2008). Thus we hypothesise that even taking into account indels and larger structural variation, genetic variation will be expected to explain only a small proportion of the differences in methylation associated with population, with spontaneously occurring epimutations contributing to the variation to a greater extent.

The larger number of methylation differences (~20x DMSs) associated with population site of origin relative to growth [CO_2] may be unexpected given the large within population genetic diversity and comparatively smaller between population genetic divergence. The identification of DMSs associated with population site of origin, even after two generations of growth in a common ambient [CO_2] environment suggests that much of the variation in methylation profiles is stable. The overlap between DMSs that are differentially methylated according to both growth [CO_2] and population additionally suggests that some of the stably transmitted variation in methylation is related to response to environmental [CO_2]. This supports a role for epigenetic inheritance as a mechanism of transgenerational indirect genetic effects, which are predicted to evolve in plant response to elevated [CO_2] because of the predictability of seasonal cycles, and the generally directional and spatially homogenous annual changes in [CO_2] (Tans and Keeling, 2018). Although significantly more sites are differentially methylated in both the population site of origin and growth [CO_2] response than would be expected by chance,

they still comprise just 1-3 % of the DMSs associated with population. This analysis suggests that although plastic remodelling of the methylome is induced by elevated [CO_2] and may result in transgenerationally stable methylation, DMSs between spring and control populations are more likely to arise independently of this remodelling. Here we evidence that a component of the methylation differences between the populations are related to genetic variation (whether or not this variation is adaptive) but a large portion must arise from the accumulation of spontaneous epimutations that are either subject to selection or increase in the population through epigenetic drift.

In this case, plants growing at natural CO_2 springs for hundreds of generations can provide critical insight into the evolution of the plant methylome under elevated [CO_2]. The availability of genomic resources are likely to have limited our ability to detect DMSs co-locating to genes or features that may impact gene regulation in this study. However, as genomic resources are further developed for the 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.

Although DNA methylation is the best studied mechanism, it remains unclear whether other chromain modifications could also potentially facilitate epigenetic inheritance (Eichten et al., 2014). Evidence largely suggests that histone modifications are erased and re-established through replication (Johnson et al., 2007), but histone H3 lysine 9 modifications have been show to be transgenerationally stable through meiosis when active removal fails (Audergon et al., 2015). Additionally, advances in the detections of other DNA base modifications (such as 5-hydroxymethylcytosine and N^6-methyladenine) are beginning to expand the scope of research into mechanisms that could facilitate epigenetic inheritance of an environmental response (Kumar et al., 2018).

4.5.5 Co-location of differentially expressed transcripts to DMSs and SNPs highlight potential targets for differential methylation and selection

This analysis identified a handful of candidate genes that could potentially play a role in adaptation to elevated [CO_2] via changes in methylation patterns. Among them, two genes that are involved in the abscissic acid (ABA) pathway, which has been identified as providing a pathway to coordinate stomatal responses to elevated [CO_2] (Chater *et al*., 2015; Youshi *et al*., 2015). The FLDH component mapped to a region near

hypomethylated DMSs in plants originating from the spring population, and were DE with increased expression in plants originating from the spring population. There were also 8 outlier SNPs in the 1 kb up and downstream of the transcript, suggesting that this region could be under selection.

The RD22 gene was significantly upregulated both in plants exposed to elevated [CO_2] and in plants originating from the spring, which makes it a good candidate for playing a role in plant adaptation to elevated [CO_2]. It was also located in the vicinity of a small region of significantly increased methylation in plants originating from the spring site. Contrary to the central dogma that increased methylation results in compaction of chromatin that represses gene transcription, it is increasingly recognised that methylation can act to both activate and repress gene expression as well as play a role in splicing and transcription factor recruitment (Tirado-Magallanes *et al*., 2017). What this makes clear is that more detailed characterisation 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 characterisation of candidate genes for their role in elevated [CO_2] response. Accordingly 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.

4.6 Conclusion

We found that the *P. lanceolata* methylome is responsive to elevated [CO_2], but there were significantly more methylation differences associated with multiple generations of growth at the naturally elevated [CO_2] spring site versus the ambient [CO_2] control site. Genetic differentiation between the two populations is generally low but there is substantial divergence in methylation profiles. A correlation between genetic variation and variation in methylation profiles accounts for some of the variation in methylation and a small proportion may be attributed to environmentally induced and transgenerational stable methylation differences. However it is likely that the majority results from spontaneous epimutation and subsequent selection or drift. We identified a handful of candidate genes that are DE and in proximity to differentially methylated sites but these require further analysis.

4.7 Acknowledgements

We acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work and further thank F. Miglietta for assisting with access to the spring site. This work was funded by the National Environmental Research Council (NERC), through a SPITFIRE DTP PhD studentship to JMS [grant number NE/L002531/1], and through funding to GT from UC Davis and The British Council.

4.8 Author Contributions

YL conducted the experiment and extracted RNA; AJWL undertook laboratory research associated with the RNA-Seq; JMS undertook the laboratory work and analysed data associated with the genome assembly, RNA-Seq, WGS and WGBS and drafted the manuscript; MAC contributed to supervision and writing; GT conceived the study, undertook supervision and contributed to writing.

4.9 References

Ainsworth, E. A., & Long, S. P. (2005) What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytologist*, **165**, 351-372.

Ainsworth, E. A., & Rogers, A. (2007) The response of photosynthesis and stomatal conductance to rising [CO_2]: mechanisms and environmental interactions. *Plant, cell & environment*, **30**, 258-270.

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**, R87.

Akalin, A., (2016) CallDMRs() function. GitHub repository, https://gist.github.com/al2na/4839e615e2401d73fe51

Alkan, C., Sajjadian, S., & Eichler, E. E. (2010) Limitations of next-generation genome sequence assembly. *Nature methods*, **8**, 61-65.

Alsdurf, J., Anderson, C., & Siemens, D. H. (2016) Epigenetics of drought-induced trans-generational plasticity: consequences for range limit development. AoB Plants, **8**, https://doi.org/10.1093/aobpla/plv146.

Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. Babraham Institute, Cambridge, UK. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Audergon, P. N., Catania, S., Kagansky, A., Tong, P., Shukla, M., Pidoux, A. L., & Allshire, R. C. (2015) Restricted epigenetic inheritance of H3K9 methylation. *Science*, **348,** 132-135.

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**, doi: 10.3389/fpls.2018.01251.

Barnaby, J. Y., & Ziska, L. H. (2012) Plant Responses to Elevated CO_2. *eLS*, doi:10.1002/9780470015902.a0023718.

Barnes, J. D., Bettarini, I., Polle, A., Slee, N., Raines, C., Miglietta, F., Raschi, A., & Fordham, M. (1997) The impact of elevated CO_2 on growth and photosynthesis in *Agrostis canina* L. ssp. *monteluccii* adapted to contrasting atmospheric CO_2 concentrations. *Oecologia*, ***110,*** 169-178.

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**, 289-300.

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**, 474-485.

Bettarini, I., Vaccari, F. P., & Miglietta, F. (1998) Elevated CO_2 concentrations and stomatal density: observations from 17 plant species growing in a CO_2 spring in central Italy. *Global Change Biology*, **4**, 17-22.

Bolger, A. M., Lohse, M., & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114-2120.

Bos, M., Harmens, H., & Vrieling, K. (1986) Gene flow in *Plantago* I. Gene flow and neighbourhood size in *P. lanceolata*. *Heredity*, **56**, 43.

Bushnell, B. (2016) BBMap short read aligner. University of California, Berkeley, California. URL http://sourceforge. net/projects/bbmap.

Cantarel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., Holt, C., Alvarado, A.S., & Yandell, M. (2008) MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome research*, **18**, 188-196.

Campbell, M. S., Holt, C., Moore, B., & Yandell, M. (2014) Genome annotation and curation using MAKER and MAKER-P. *Current Protocols in Bioinformatics*, **48**, 4-11.

Chater, C., Peng, K., Movahedi, M., Dunn, J. A., Walker, H. J., Liang, Y.K., McLachlan, D. H., Casson, S., Isner, J. C., Wilson, I., & Hetherington, A. M. (2015). Elevated CO_2-induced responses in stomata require ABA and ABA signalling. *Current Biology*, **25**, 2709–2716.

Chikhi, R., & Medvedev, P. (2013) Informed and automated k-mer size selection for genome assembly. *Bioinformatics*, **30**, 31-37.

Chinnusamy, V., & Zhu, J. K. (2009) Epigenetic regulation of stress responses in plants. *Current opinion in plant biology*, **12**, 133-139.

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**, e1501340.

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**, R60.

Dierckxsens, N., Mardulyn, P., & Smits, G. (2017) NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, **45**, e18.

Doyle, J. J., & Doyle, J. L. (1987). CTAB DNA extraction in plants. *Phytochemical Bulletin*, **19**, 11-15.

Eichten, S. R., Schmitz, R. J., & Springer, N. M. (2014) Epigenetics: Beyond Chromatin Modifications and Complex Genetic Regulation. *Plant Physiology*, **165**, 933–947.

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**, 1520-1531.

Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., & Salazar, G. A. (2015) The Pfam protein families database: towards a more sustainable future. *Nucleic acids research*, **44**, 279-285.

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**, 977–993

Fojtová, M., Kovařík, A., & Matyášek, R. (2001) Cytosine methylation of plastid genome in higher plants. Fact or artefact? *Plant Science*, **160**, 585-593.

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R., Heimann, M., Thonicke, K., Zaehle, S., & Reichstein, M. (2016) Enhanced seasonal CO_2 exchange caused by amplified plant productivity in northern ecosystems. *Science*, **351**, 696-699.

Frenck, G., Linden, L., Mikkelsen, T. N., Brix, H., & Jørgensen, R. B. (2013) Response to multi-generational selection under elevated [CO_2] in two temperature regimes suggests enhanced carbon assimilation and increased reproductive output in *Brassica napus* L. *Ecology and evolution*, **3**, 1163-1172.

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**, doi: 10.1038/srep14922.

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**, 1893-2017.

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*, https://doi.org/10.1111/nph.15487.

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**, 240-255.

Grossman, J. D., & Rice, K. J. (2014) Contemporary evolution of an invasive grass in response to elevated atmospheric CO_2 at a Mojave Desert FACE site. *Ecology letters*, **17**, 710-716.

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**, 461-485.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., ... & MacManes, M. D. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols*, **8**, 1494-1512.

Haj, A., & O'Connor, D.H., (2016) Scaffold Stitcher. Bitbucket repository, https://bitbucket.org/dholab/scaffoldstitcher.git

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, **4**, 44-57.

Heard, E., & Martienssen, R. A. (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell*, **157**, 95-109.

Henderson, I. R., & Jacobsen, S. E. (2007) Epigenetic inheritance in plants. *Nature*, **447**, 418–424.

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**, 391–398.

IPCC. (2014). Summary for Policymakers. In: C. B. Field *et al.,* (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. Cambridge, United Kingdom, and New York, NY, USA: Cambridge University Press. pg 29.

Johnson, L. M., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J., & Jacobsen, S. E. (2007) The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Current Biology*, **17**, 379–384.

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., & Pesseat, S. (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics*, **30**, 1236-1240.

Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., & Walichiewicz, J. (2005) RepBase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research*, **110**, 462–467.

Karan, R., Leon, T. D., Biradar, H., & Subudhi, P. K. (2012) Salt stress induced variation in DNA methylation pattern and its influence on gene expression in contrasting rice genotypes. *PLoS One*, **7**, e40203.

Kawakatsu, T., Stuart, T., Valdes, M., Breakfield, N., Schmitz, R. J., Nery, J. R., Urich, M. A., Han, X., Lister, R., Benfey, P. N., & Ecker, J. R. (2016) Unique cell-type-specific patterns of DNA methylation in the root meristem. *Nature plants*, **2**, 16058.

Kolde, R. (2013) pheatmap: Pretty Heatmaps, http://CRAN.R-project.org/package=pheatmap.

Körner, C., & Miglietta, F. (1994) Long term effects of naturally elevated CO_2 on Mediterranean grassland and forest trees. *Oecologia*, **99**, 343-351.

Korf, I. (2004) Gene finding in novel genomes. *BMC Bioinformatics*, **14**, https://doi.org/10.1186/1471-2105-5-59.

Krueger, F., & Andrews, S. R. (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*, **27**, 1571-1572.

Krüger, M., Jones, D., Frerichs, J., Oppermann, B. I., West, J., Coombs, P., Green, K., Barlow, T., Lister, R., Shaw, R., & Strutt, M. (2011) Effects of elevated CO_2 concentrations on the vegetation and microbial populations at a terrestrial CO_2 vent at Laacher See, Germany. *International Journal of Greenhouse Gas Control*, **5**, 1093-1098.

Kumar, S., Chinnusamy, V., & Mohapatra, T. (2018) Epigenetics of Modified DNA Bases: 5-Methylcytosine and Beyond. *Frontiers in genetics*, **9**, 640.

Lê, S., Josse, J. & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*. **25,** 1-18.

Langmead, B., & Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357–359.

Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., & Carey, V. J. (2013). Software for computing and annotating genomic ranges. *PLoS computational biology*, **9**, e1003118.

Law, J. A., & Jacobsen, S. E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics*, **11**, 204-220.

Leakey, A. D., & Lau, J. A. (2012) Evolutionary context for understanding and manipulating plant responses to past, present and future atmospheric [CO_2]. *Philosophical Transactions of the Royal Society B*, **367**, 613-629.

Li, H. (2016) seqtk: A fast and lightweight tool for processing sequences. GitHub repository, https://github.com/lh3/seqtk.

Liang, D., Zhang, Z., Wu, H., Huang, C., Shuai, P., Ye, C. Y., Tang, S., Wang, Y., Yang, L., Wang, J., & Yin, W. (2014) Single-base-resolution methylomes of *Populus trichocarpa* reveal the association between DNA methylation and drought stress. *BMC genetics*, **15**, S9.

Lohse, M., Drechsel, O., Kahlau, S., & Bock, R. (2013) OrganellarGenomeDRAW - a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research,* **41**, 575-581.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., & Tang, J. (2012) SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience*, **1**, doi:10.1186/2047-217X-1-18.

Lu, X., Wang, X., Chen, X., Shu, N., Wang, J., Wang, D., Wang, S., Fan, W., Guo, L., Guo, X., & Ye, W. (2017) Single-base resolution methylomes of upland cotton (*Gossypium hirsutum* L.) reveal epigenome modifications in response to drought stress. *BMC genomics*, **18**, 1-14.

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**, doi:10.1186/s12870-015-0623-1.

Martin, S. H. (2017) VCF_processing, GitHub repository, 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**, 4288-4297.

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**, 377-386.

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., Carlisle, E., Dietterich, L. H., Fitzgerald, G., Hasegawa, T., & Holbrook, N. M. (2014) Increasing CO_2 threatens human nutrition. *Nature*, **510**, 139

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 CO_2 concentration. *Oecologia*, **165**, 809-818.

Neves, D. M., da Hora Almeida, L. A., Santana-Vieira, D. D. S., Freschi, L., Ferreira, C. F., dos Santos Soares Filho, W., Costa, M.G.C., Micheli, F., Coelho Filho, M.A., & da Silva Gesteira, A. (2017) Recurrent water deficit causes epigenetic and hormonal changes in citrus plants. *Scientific reports*, **7**, 13684.

Niederhuth, C. E., Bewick, A. J., Ji, L., Alabady, M. S., Do Kim, K., Li, Q., Rohr, N.A., Rambani, A., Burke, J.M., Udall, J.A. & Egesi, C. (2016) Widespread natural variation of DNA methylation within angiosperms. *Genome biology*, **17**, doi: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**, 149-156.

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**, 176-193.

Onoda, Y., Hirose, T., & Hikosaka, K. (2009) Does leaf photosynthesis adapt to CO_2-enriched environments? An experiment on plants originating from three natural CO_2 springs. *New Phytologist*, **182**, 698-709.

Paoletti, E., Pfanz, H., & Raschi, A. (2005) Pros and cons of CO_2 springs as experimental sites. In Plant responses to air pollution and global change. pp. 195-202. Springer, Tokyo.

Park, Y., & Wu, H. (2016) Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics*, **32**, 1446-1453.

Pearson, P. N., & Palmer, M. R. (2000) Atmospheric carbon dioxide concentrations over the past 60 million years. *Nature*, **406**, 695-699.

Pigliucci, M., Murren, C. J., & Schlichting, C. D. (2006) Phenotypic plasticity and evolution by genetic assimilation. *Journal of Experimental Biology*, **209**, 2362–2367.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006) CODA: Convergence diagnosis and output analysis for MCMC. *R News*, **6**, 7–11.

Polle, A., McKee, I., & Blaschke, L. (2001) Altered physiological and growth responses to elevated [CO_2] in offspring from holm oak (*Quercus ilex* L.) mother trees with lifetime exposure to naturally elevated [CO_2]. *Plant, Cell & Environment*, **24**, 1075-1083.

Potvin, C., & Tousignant, D. (1996) Evolutionary consequences of simulated global change: genetic adaptation or adaptive phenotypic plasticity. *Oecologia*, **108**, 683-693.

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**, 904.

Quadrana, L., & Colot, V. (2016) Plant transgenerational epigenetics. *Annual review of genetics*, **50**, 467-491.

R-Core-Team. (2015) R: A language and environment for statistical computing [Online]. R Foundation for Statistical Computing, Vienna, Austria. Available: https://www.R-project.org/.

Rico, L., Ogaya, R., Barbeta, A., & Penuelas, J. (2014) Changes in DNA methylation fingerprint of *Quercus ilex* trees in response to experimental field drought simulating projected climate change. *Plant Biology*, **16**, 419-427.

Saban, J.M. Chapman, M.A., & Taylor, G. (2018) Face facts hold for multiple generations; Evidence from natural CO_2 springs. *Global Change Biology*, doi:10.111/gcb.14437.

Saze, H., Tsugane, K., Kanno, T., & Nishimura, T. (2012) DNA methylation in plants: relationship to small RNAs and histone modifications, and functions in transposon inactivation. *Plant and Cell Physiology*, **53**, 766-784.

Schlichting, C. D., & Wund, M. A. (2014) Phenotypic plasticity and epigenetic marking: an assessment of evidence for genetic accommodation. *Evolution*, **68**, 656-672.

Schönberger, B., Chen, X., Mager, S., & Ludewig, U. (2016) Site-dependent differences in DNA methylation and their impact on plant establishment and phosphorus nutrition in *Populus trichocarpa*. *PloS One*, **11**, e0168623.

Schmitz, R. J., Schultz, M. D., Urich, M. A., Nery, J. R., Pelizzola, M., Libiger, O., Alix, A., McCosh, R. B., Chen, H., Schork, N. J., & Ecker, J. R. (2013) Patterns of population epigenomic diversity. *Nature*, **495**, 193-198.

Šibanc, N., Dumbrell, A. J., Mandić-Mulec, I., & Maček, I. (2014) Impacts of naturally elevated soil CO_2 concentrations on communities of soil archaea and bacteria. *Soil Biology and Biochemistry*, **68**, 348-356.

Šibanc, N., Zalar, P., Schroers, H. J., Zajc, J., Pontes, A., Sampaio, J. P., & Macek, I. (2018*) Occultifur mephitis* fa, sp. nov. and other yeast species from hypoxic and elevated CO_2. *International Journal of Systematic and Evolutionary Microbiology*, **68**, 2285-2298.

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**, 3210-3212.

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**, 1117-1123.

Song, X., & Cao, X. (2017) Context and Complexity: Analyzing Methylation in Trinucleotide Sequences. *Trends in plant science*, **22**, 351-353.

Stanke, M., & Waack, S. (2003) Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, **19**, 215–225.

Tans, P., & Keeling, R. (2018) ESRL Global Monitoring Division Global Greenhouse Gas Reference Network. http://www.esrl.noaa.gov/gmd/ccgg/trends/ Accessed 18th October 2018.

Taylor, G., Street, N. R., Tricker, P. J., Sjödin, A., Graham, L., Skogström, O., Calfapietra, C., Scarascia-Mugnozza, G., & Jansson, S. (2005) The transcriptome of *Populus* in elevated CO_2. *New Phytologist*, **167**, 143-154.

Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E. S., Fischer, A., Bock, R., and Greiner, S., (2017) GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Research*, **45**, 6-11.

Tirado-Magallanes, R., Rebbani, K., Lim, R., Pradhan, S., & Benoukraf, T. (2017) Whole genome DNA methylation: beyond genes silencing. *Oncotarget*, **8**, 5629.

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**, 833-843.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J. & Banks, E. (2013) From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Current protocols in bioinformatics*, **43**, 11-10.

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**, 518-528.

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**, 1108-1118.

Walter, W., Sánchez-Cabo, F., & Ricote, M. (2015) GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics*, **31**, 2912-2914.

Ward, J. K., Antonovics, J., Thomas, R. B., & Strain, B. R. (2000) Is atmospheric CO_2 a selective agent on model C 3 annuals? *Oecologia*, **123**, 330-341.

Wang, D., Heckathorn, S. A., Wang, X., & Philpott, S. M. (2012) A meta-analysis of plant physiological and growth responses to temperature and elevated CO_2. Oecologia, **169**, 1-13.

Watson-Lazowski, A., Lin, Y., Miglietta, F., Edwards, R. J., Chapman, M. A., & Taylor, G. (2016) Plant adaptation or acclimation to rising CO_2? Insight from first multigenerational RNA-Seq transcriptome. *Global change biology*, **22**, 3760-3773.

Wieneke, S., Prati, D., Brandl, R., Stöcklin, J., & Auge, H. (2004) Genetic variation in *Sanguisorba minor* after 6 years in situ selection under elevated CO_2. *Global Change Biology*, **10**, 1389-1401.

Wheeler, T., & Von Braun, J. (2013) Climate change impacts on global food security. *Science*, **341**, 508-513.

Whiteford, N., Haslam, N., Weber, G., Prügel-Bennett, A., Essex, J. W., Roach, P. L., Bradley, M., & Neylon, C. (2005) An analysis of the feasibility of short read sequencing. *Nucleic acids research*, **33**, doi:10.1093/nar/gni170.

Wu, H., Xu, T., Feng, H., Chen, L., Li, B., Yao, B., Qin, Z., Jin, P., & Conneely, K. N. (2015) Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. *Nucleic acids research*, **4**, https://doi.org/10.1093/nar/gkv715.

Xia, H., Huang, W., Xiong, J., Tao, T., Zheng, X., Wei, H., Yue, Y., Chen, L., & Luo, L. (2016) Adaptive epigenetic differentiation between upland and lowland rice ecotypes revealed by methylation-sensitive amplified polymorphism. *PloS One*, **11**, e0157810.

Yamaguchi-Shinozaki, K., & Shinozaki, K. (1993) The plant hormone abscisic acid mediates the drought-induced expression but not the seed-specific expression of *rd22*, a gene responsive to dehydration stress in *Arabidopsis thaliana*. *Molecular and General Genetics MGG*, **238**, 17-25.

Yong-Villalobos, L., Cervantes-Pérez, S. A., Gutiérrez-Alanis, D., Gonzáles-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 signaling & behavior*, **11**, e1173300.

Youshi, T., and Santrucek, J. (2015) Superimposed behaviour of gm under ABA-induced stomata closing and low CO_2. *Plant Cell Environment*, **38**, 385–387.

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 caliph medic (*Medicago truncatula*). *BMC genomics*, **19**, 78.

Yi, S. V. (2017). Insights into Epigenome Evolution from Animal and Plant Methylomes. *Genome biology and evolution*, **9**, 3189-3201.

Zhang, M., Kimatu, J. N., Xu, K., & Liu, B. (2010) DNA cytosine methylation in plant development. *Journal of Genetics and Genomics*, **37**, 1-12.

Zhu, A., Guo, W., Gupta, S., Fan, W., & Mower, J. P. (2016) Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *New Phytologist*, **209**, 1747-1756.

Zilberman, D. (2017). An evolutionary case for functional gene body methylation in plants and animals. *Genome biology*, **18**, 87-91.

Chapter 5 General Discussion

Atmospheric CO_2 concentrations ($[CO_2]$) have increased rapidly since the industrial revolution and are predicted to continue to rise to beyond that seen in the last 23 million years (IPCC, 2014; Pearson and Palmer, 2003). The process of photosynthesis is fundamental to human existence, providing the oxygen we breathe, the food we eat, as well as fuel and fibre for shelter and energy. As a key and often rate limiting substrate for photosynthesis (Ward and Strain, 1997; Zheng *et al*., 2018), changes in atmospheric $[CO_2]$ will therefore impact plant functioning within ecosystems and predicting these changes is critical to managing the impact of climate change (Becklin *et al*., 2017). However, the vast majority of research in this area has focused on the single generation plastic response of plants (mostly crops) to elevated CO_2 ($[CO_2]$) (Ainsworth *et al*., 2007; Leakey *et al*., 2009), and the challenges associated with multigenerational experiments has meant that research into the multigenerational response of plants to elevated $[CO_2]$ has been comparatively neglected. Here we have aimed to address key gaps in current understanding of the multigenerational plant response to elevated $[CO_2]$ and argue for the use of plants growing around natural CO_2 springs (in combination with other approaches) to study the phenotypes produced and the mechanisms and processes that produce them.

Natural CO_2 springs have been used to study the long-term (trees) and multigenerational (herbs) responses of plants to elevated $[CO_2]$ since the early 1990's (Miglietta *et al*., 1993; Raschi *et al*., 1997). In chapter 2 the available research was synthesised through a systematic search of the literature and subsequent meta-analysis to identify trends in the morphology of plants growing at natural CO_2 springs, and compare them to meta-analyses of the plastic response to elevated $[CO_2]$, in order to draw conclusions about the potential validity of single generation analyses in predicting multigenerational response. We found that there was remarkable consistency in the magnitude and direction of the single generation plastic response to elevated $[CO_2]$ and those observed in plants at natural CO_2 springs that have been growing at elevated $[CO_2]$ for multiple generations. However, prediction of plant responses to multigenerational elevated $[CO_2]$ requires more than just an overview of the trends in plant responses, it also requires an understanding of how those phenotypes are coordinated. Accordingly, we sought to

understand the relative contributions of plasticity, indirect and direct genetic effects to the multigenerational plant elevated [CO_2] response. Chapter 3 and 4 combine an experimental approach with statistical models and HTS technologies to study the components of phenotypic responses of *Plantago lanceolata* to single and multigenerational elevated [CO_2] exposure.

5.1 The use of natural CO_2 springs to study elevated [CO_2] response

Natural CO_2 springs sites have many advantages over studies which utilise experimentally elevated [CO_2], including the reduced cost, natural environmental conditions, and the flexibility for varied experiments involving gradients, reciprocal transplants or the collection of material for controlled environment experiments (Paoletti *et al*., 2005). The most important advantage is that these sites have already created populations of plants that have grown in an enriched [CO_2] environment for many generations (and that would take years to produce from a [CO_2] enrichment experiment) and this is becoming an increasingly significant consideration as the need to predict and prepare for the effects of climate change become increasingly urgent (Stocker *et al*., 2013; Myers *et al*., 2014). Although natural CO_2 springs were initially popular as a system to study plant response to elevated [CO_2] (Figure 1.6), this research has largely been abandoned and there have only been two publications resulting from experiments with plants at these sites in the last five years (Van Loon *et al*., 2016; Watson-Lazowski *et al*., 2016). Many of the criticisms of the use of natural CO_2 springs are justified and require acknowledgement and management (Paoletti *et al*., 2005). The vent emissions of natural CO_2 springs can contain contaminants (especially sulphur), are often variable in quantity over time and the concentrations experienced by plants are dependent on weather conditions (Raschi *et al*., 1997). Additionally the release of CO_2 from the ground can result in vertical gradients (which particularly need to be considered for tall plants) and changes to soil properties including pH (Vodnik *et al*., 2002; Onoda *et al*., 2009; Ueda *et al*., 2017). This critically necessitates the thorough characterisation of natural CO_2 spring sites and the evaluation of the suitability of the site as a model. In our meta-analysis of plant responses to elevated [CO_2] at natural CO_2 springs we included selection criteria such as recorded concentrations of sulphur contamination and [CO_2] to filter published research, and recommended that [CO_2] over time and across gradients, sulphur contamination, soil pH and soil nitrogen and

carbon characteristics were evaluated at sites as standard practice. For experimental work we used plants growing at the natural CO_2 spring Bossoleto in Italy, which to date is arguably the most well-characterised study site available (Körner and Miglietta, 1994; Scholefield *et al*., 2004; Kies *et al*., 2015; Saban *et al*., 2015 *unpublished data*).

Where natural CO_2 springs are extensively characterised and selected based on their suitability as a model for plant response to elevated [CO_2] under climate change, they can still provide crucial insight into long-term and multigenerational plant response to elevated [CO_2], when findings are considered in the context of other experimental approaches. This is in much the same way that experiments at the other end of the scale, with plants in pots in controlled chambers must interpret results within the constraints enforced by using this model (Poorter *et al*., 2012).

5.2 Plant responses to elevated [CO_2] over multigenerations of exposure at natural CO_2 springs are consistent with those predicted from single generation studies

Hundreds if not thousands of studies that record plant plastic response to elevated [CO_2] have been published to date (Jablonski *et al*., 2002; Ainsworth and Long, 2005; Ainsworth and Rogers, 2007; Wang *et al*., 2015; Kimball *et al*., 2016). Meta-analyses of plant plastic responses to elevated [CO_2] recorded in these studies highlight that there are trends in the plastic response to elevated [CO_2] across functional groups, but physiological responses vary between and even within species. These trends have been used to inform and validate predictive models of for example plant primary production, crop quality and canopy transpiration (Vanuytrecht and Thorburn, 2017).

In chapter two of this thesis we demonstrate that trends seen in plant plastic response to elevated [CO_2] are consistent with those observed in plants at natural CO_2 springs. Increased CO_2 availability increases the rate of photosynthesis, and over days to weeks of exposure plants may acclimate photosynthetic capacity through down regulation of components of the photosynthetic pathway (Leakey *et al*., 2009). However photosynthetic rate remains enhanced despite acclimation (Nowak *et al*., 2004; Ainsworth and Rogers, 2007) and this enhancement is seen even after many generations of exposure to elevated [CO_2].

Enhanced photosynthesis leads to increased stomatal closure which reduces water lost through transpiration both in the single and multigenerational response (Leakey *et al*., 2009). As well as increased stomatal closure, in the plastic response to elevated [CO_2], signalling from mature leaves to developing leaves (Lake *et al*., 2001; Miyazawa *et al*., 2005) can result in changes in stomatal density that increase water use efficiency (Woodward, 1987; Kürschner *et al*., 1997; Bettarini *et al*., 1998; Franks and Beerling, 2009; Hirano *et al*., 2012). However, meta-analysis of the multigenerational (spring) response and the plastic (FACE) response did not detect significant effects of single or multigenerational elevated [CO_2] exposure on stomatal density and there was significant variation in the response between species (Ainsworth and Rogers, 2007). The absence of a trends towards reduced stomatal density in these plants is difficult to rationalise with the observation of a negative correlation between SD and [CO_2] in the geological record and from herbarium specimens over the last century (Woodward, 1987; Franks and Beerling, 2009). Functional analyses of the molecular basis of stomatal patterning under elevated [CO_2] is ongoing and should provide insight into this response in the near future (Engineer *et al*., 2016; Xu *et al*., 2016; He *et al*., 2018). Similarly, understanding the molecular basis of the decline in leaf nitrogen observed in this meta-analysis of CO_2 spring population responses and in FACE site studies is undergoing active research (Seneweera *et al*., 2011; Bloom *et al*., 2012; Ruiz-Vera *et al*., 2017; Terrer *et al*., 2018). In both FACE and natural CO_2 sites, more studies that characterise soil nitrogen dynamics in addition to measuring plant nitrogen content are required to interpret the effect of elevated [CO_2] on plant nitrogen dynamics (Ross *et al*., 2000; Ueda *et al*., 2017; Gamage *et al*., 2018).

Overall, the tight correlation between FACE estimates of trait response to elevated [CO_2] and natural CO_2 spring estimates is reassuring that the multigenerational and long-term plant response to elevated [CO_2] is similar in magnitude and direction to that predicted by FACE. By association the positive photosynthetic response of plants to elevated [CO_2] is predicted to continue even when incorporating potential adaptation over many generations (IPCC, 2014). Further this correlation provides some validation that plant responses to elevated [CO_2] at natural CO_2 sites are relevant to the prediction of future elevated [CO_2] plant responses.

5.3 Indirect and direct genetic effects in *P. lanceolata* multigenerational response to elevated [CO_2]

Experiments in both chapter three and chapter four demonstrated that phenotypic plasticity was the predominant component of the multigenerational response to elevated [CO_2]. However there was also evidence that plants originating from the natural CO_2 spring site had adapted to life at elevated [CO_2]. In chapter two we demonstrate that transgenerational indirect genetic effects, both maternal and paternal contribute to the multigenerational response. A significant paternal effect that was responsive to elevated [CO_2] meant offspring that had paternal spring site ancestry were delayed in flowering. A maternal effect on flowering in response to elevated [CO_2] has been widely documented, but never a paternal effect, and the mechanism by which this could occur is elusive and intriguing (Jablonski *et al*., 2002). Delayed flowering time may result in altered reproductive opportunities, including delayed seed dispersal and the adaptive value of this is unclear. Reproductive strategy also appears to be altered under elevated [CO_2] with offspring with maternal spring site ancestry producing significantly more seeds per spike, but less reproductive biomass overall when compared to those with control site ancestry. Phenotypes related to biomass production had a fixed (potentially adaptive) component to their multigenerational response to elevated [CO_2] in both chapter three and in the phenotypes presented in Watson-Lazowski *et al*., 2016 that inform chapter four. In chapter four, above ground biomass was significantly affected by growth [CO_2] while in chapter three above ground biomass was significantly affected by site of origin but not growth [CO_2], although the reaction norms were similar. This discrepancy may be associated with differences in the experimental design, statistical approach and/or statistical power between the two experiments.

Population genetic analyses revealed that there was little sequence divergence between the populations at the spring and control sites, but there was high within population diversity, likely as a result of obligate outcrossing and limited pollen and seed dispersal. Just four differentially expressed genes were mapped to within 1 kb of a SNP between the two populations. This seemed to suggest there is only a limited role for direct genetic adaptation in the multigenerational response of *P. lanceolata* to elevated [CO_2], where indirect genetic effects may play a more significant role.

5.4 Changes in methylation form part of both the plastic and adaptive response to elevated [CO_2]

Chapter four is the first study to analyse global methylation pattern remodelling in response to elevated [CO_2] and it demonstrates that methylation profiles are remodelled by growth at elevated [CO_2]. Some of these changes can be maintained through at least two generations of growth at ambient [CO_2] as environmentally induced and heritably stable chromatin modifications. However, there were many more methylation differences between the spring and control population that were not induced by elevated [CO_2], likely resulting from spontaneous epimutations that are stably inherited and fixed in the population by selection or drift. Although other environmental differences between the spring and the control site could be partially responsible for the large difference in methylation profiles between spring and control plant populations (Körner and Miglietta, 1994), we considered this was likely to be small in comparison to the effect of epimutation because [CO_2] is the key environmental difference that differentiates these two population sites. Although gene flow is sufficiently high for low sequence divergence between the populations, potentially a higher epimutation rate or stronger selection has facilitated the accumulation of methylation differences.

5.5 Broad implications of the thesis

This thesis contributes to a body of research aimed at predicting plant responses to elevated atmospheric [CO_2] under future CO_2 emissions scenarios. Chapter two suggests that despite intrinsic limitations of single generation FACE studies, the conclusions drawn from them (including those used in policy making) may well be robust to the additional impact of direct and indirect genetic effects in the multigenerational response of plants to elevated [CO_2]. Chapter three provides evidences that plants growing for multiple generations at elevated [CO_2] have a 'fixed' increase in biomass, which could be further investigated to direct breeding efforts of commercially farmed *Plantago lanceolata* and related species. Finally, chapters three and four investigate the mechanisms that coordinate the multigenerational response to elevated [CO_2], which although preliminary, because it considers only a single species, could (in combination with other studies) contribute to greater predictive power to model plant responses to elevated [CO_2].

5.6 Study limitations

There were several limitations to the interpretation of the results presented here. For analysis of general trends in plant response at natural CO_2 springs the relatively low number of publications limited the analysis of functional groups and site effects. We also found that many of the CO_2 spring sites did not characterise parameters such as plant nitrogen source, which would facilitate interpretation of responses. In the transgenerational experiment of chapter three one key limitation was the availability of experimental facilities, where ideally we would have grown the third generation at the same temperature as the two previous generations but were unable to because the environmentally controlled chambers with [CO_2] control did not facilitate a lower temperature range. To mitigate against this we included the 'Generation' factor with Monte Carlo Markov Chain (MCMC) generalised linear mixed models (GLMMs) but this would have reduced the ability to detect parental and grandparental effects. This study could further have been improved by larger sample sizes and more generations but these were restricted by the practicalities of managing this experiment. In chapter four the identification of several genes associated with photosynthesis and differentially expressed both in elevated relative to ambient [CO_2] and in spring vs control derived plants, meant that it would have been particularly informative to have measured photosynthetic rate in this experiment, so that the phenotype could be directly related to gene expression changes. Lastly a major limitation to the identification of candidate genes that could be affected by differential methylation or genetic variation was the quality of the reference genome assembled. Unfortunately this kind of fragmented assembly was the best that could be expected from the assembly of the genome of a heterozygous individual with a large genome that likely contains large repeat regions with short reads and a single insert size. However a fully annotated genome, using multiple insert sizes and with both short and long read sequencing was beyond the financial constraints of this project. The genome assembled, although basic and extremely fragmented, provided sufficient scaffolding to analyse genetic variation and global trends in methylation in this system to a far greater resolution than has ever previously been achieved (or attempted).

One further major limitation of this study is the isolation of the effect of elevated [CO_2] from other environmental factors in experimental manipulations, especially factors such as increased temperature and drought, which are also likely to be key features of

predicted climates (IPCC, 2014). Although single factor experiments facilitate a dissection of the mechanisms coordinating a response to a specific environmental factor, it is also important to study predicted environmental changes in combination with others. This allows the detection of interaction effects that are relevant to predicting plant responses to climate change (Albert et al., 2011). One strength of the use of natural CO_2 springs is that plant response to elevated [CO_2] is determined within the context of an ecosystem, with 'natural' environmental conditions such as limited nutrient or water supply. However, experimental manipulations involving multiple factors are also needed to provide a controlled approach to identifying interactions between factors and the mechanisms that coordinate them (Mikkelsen *et al.,* 2007). Plant response to elevated [CO_2] and increased temperature are the most well studied combination of environmental factors involving plant response to elevated [CO_2] (Dusenge *et al.,* 2019). Elevated [CO_2] is expected to increase photosynthetic rate across a range of temperatures (Wang *et al.,* 2012) and increase the heat tolerance of photosynthesis (Rodrigues *et al.,* 2016) but may also exacerbate heat stress by reducing stomatal conductance (Kimball and Bernacchi, 2006). On a global scale this interaction predicts that CO_2 fertilisation may be stonger in tropical than temperate regions (Baig *et al.,* 2015). However, since temperature also varies on smaller spatial and temporal scales, plant response to elevated [CO_2] will be variable on these scales, through moderation by temperature and other factors. Thus it is clear that a deeper understanding of how these factors interact will be required to predict the impact of climate change on ecosystems.

5.7 Conclusions

This thesis aimed to address gaps in the current knowledge of plant response to elevated [CO_2] over multiple generations. Over three data chapters we have shown that i) Single generation plastic responses to elevated [CO_2] are consistent with the multigenerational (herbs) and long-term (trees) response of plants to elevated [CO_2] at natural CO_2 springs ii) In an experimental system using *P. lanceolata* from the Bossoleto CO_2 spring, the phenotypic response of plants to elevated [CO_2] over multiple generations is predominantly plastic but with both direct and indirect genetic effects contributing to some traits iii) in this system methylation is remodelled under the plastic response to elevated [CO_2] and may also contribute to adaptation.

5.8 Future work

In order to take advantage of plants growing at natural CO_2 springs as a precious resource for studying plant adaptation to elevated [CO_2], future research at these sites will need to focus on two key areas; firstly the full characterisation of properties of the environment experienced by plants at CO_2 spring sites and secondly the development of genomic resources for non-model species that would facilitate further genomic analyses of these plants, taking advantage of increasingly affordable HTS technologies. Additionally in order to corroborate and extend the findings of this study, methylation profiles of plants growing under elevated [CO_2] in other species and other experimental designs are needed.

5.9 References

Ainsworth, E. A., & Long, S. P. (2005) What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytologist*, **165**, 351-372.

Ainsworth, E. A., & Rogers, A. (2007) The response of photosynthesis and stomatal conductance to rising [CO_2]: mechanisms and environmental interactions. *Plant, Cell & Environment*, **30**, 258-270.

Albert, K. R., Ro-Poulsen, H., Mikkelsen, T. N., Michelsen, A., Van der Linden, L., & Beier, C. (2011) Interactive effects of elevated CO_2, warming, and drought on photosynthesis of *Deschampsia flexuosa* in a temperate heath ecosystem. *Journal of Experimental Botany*, **62**, 4253-4266.

Baig, S., Medlyn, B. E., Mercado, L. M., & Zaehle, S. (2015) Does the growth response of woody plants to elevated CO_2 increase with temperature? A model-oriented meta-analysis. *Global Change Biology*, **21**, 4303–4319.

Becklin, K. M., Walker, S. M., Way, D. A., & Ward, J. K. (2017) CO_2 studies remain key to understanding a future world. *New Phytologist*, **214**, 34-40.

Bettarini, I., Vaccari, F. P., & Miglietta, F. (1998) Elevated CO_2 concentrations and stomatal density: observations from 17 plant species growing in a CO_2 spring in central Italy. *Global Change Biology*, **4**, 17-22.

Dusenge, M. E., Duarte, A. G., & Way, D. A. (2019) Plant carbon metabolism and climate change: elevated CO_2 and temperature impacts on photosynthesis, photorespiration and respiration. *New Phytologist*, **221**, 32-49.

Engineer, C. B., Hashimoto-Sugimoto, M., Negi, J., Israelsson-Nordstrom, M., Azoulay-Shemer, T., Rappel, W. J., Iba, K., & Schroeder, J. I. (2016) CO_2 sensing and CO_2 regulation of stomatal conductance: advances and open questions. *Trends in plant science*, **21**, 16-30.

Franks, P. J., & Beerling, D. J. (2009) Maximum leaf conductance driven by CO_2 effects on stomatal size and density over geologic time. *Proceedings of the National Academy of Sciences*, **106**, 10343-10347.

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. *Plant, Cell & Environment,* **41**, 1233-1246.

Hirano, A., Hongo, I., & Koike, T. (2012) Morphological and physiological responses of Siebold's beech (*Fagus crenata*) seedlings grown under CO_2 concentrations ranging from pre-industrial to expected future levels. *Landscape and ecological engineering*, **8**, 59-67.

He, J., Zhang, R. X., Peng, K., **Tagliavia**, C., Li, S., Xue, S., Liu, A., Hu, H., Zhang, J., Hubbard, K. E., & Held, K. (2018) The BIG protein distinguishes the process of CO_2-induced stomatal closure from the inhibition of stomatal opening by CO_2. *New Phytologist*, **218**, 232-241.

IPCC. (2014) Summary for Policymakers. In: C. B. Field *et al.,* (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. Cambridge, United Kingdom, and New York, NY, USA: Cambridge University Press. pg 29.

Lake, J. A., Woodward, F. I., & Quick, W. P. (2002) Long-distance CO_2 signalling in plants. *Journal of Experimental Botany*, **53**, 183-193.

Leakey, A. D., Ainsworth, E. A., Bernacchi, C. J., Rogers, A., Long, S. P., & Ort, D. R. (2009) Elevated CO_2 effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. *Journal of experimental botany*, **60**, 2859-2876.

Jablonski, L. M., Wang, X. Z., & Curtis, P. S. (2002) Plant reproduction under elevated CO_2 conditions: a meta-analysis of reports on 79 crop and wild species. *New Phytologist*, **156**, 9–26.

Kies, A., Hengesch, O., Tosheva, Z., Raschi, A., & Pfanz, H. (2015). Diurnal CO_2-cycles and temperature regimes in a natural CO_2 gas lake. *International Journal of Greenhouse Gas Control*, **37**, 142-145.

Kimball, B. A., & Bernacchi, C. J. (2006) Evapotranspiration, canopy temperature, and plant water relations. In: H. Blum *et al.,* (eds.) Managed ecosystems and CO_2. Berlin, Germany: Springer-Verlag. pg 311–324.

Kimball, B. A. (2016) Crop responses to elevated CO_2 and interactions with H_2O, N, and temperature. *Current opinion in plant biology*, **31**, 36-43.

Körner, C., & Miglietta, F. (1994) Long term effects of naturally elevated CO_2 on Mediterranean grassland and forest trees. *Oecologia*, **99**, 343-351.

Kürschner, W. M., Wagner, F., Visscher, E. H., & Visscher, H. (1997) Predicting the response of leaf stomatal frequency to a future CO_2-enriched atmosphere: constraints from historical observations. *Geologische Rundschau*, **86**, 512-517.

Miglietta, F., Raschi, A., Bettarini, I., Resti, R., & Selvi, F. (1993) Natural CO_2 springs in Italy: a resource for examining long-term response of vegetation to rising atmospheric CO_2 concentrations. *Plant, Cell & Environment*, **16**, 873-878.

Mikkelsen, T. N., Beier, C., Jonasson, S., Holmstrup, M., Schmidt, I. K., Ambus, P., Pilegaard, K., Michelsen, A., Albert, K., Andresen, L. C., & Arndal, M. F. (2008) Experimental design of multifactor climate change experiments with elevated CO_2, warming and drought: the CLIMAITE project. *Functional Ecology*, **22**, 185-195.

Miyazawa, S. I., Livingston, N. J., & Turpin, D. H. (2005) Stomatal development in new leaves is related to the stomatal conductance of mature leaves in poplar (*Populus trichocarpa× P. deltoides*). *Journal of Experimental Botany*, **57**, 373-380.

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., Carlisle, E., Dietterich, L. H., Fitzgerald, G., Hasegawa, T., & Holbrook, N. M. (2014) Increasing CO_2 threatens human nutrition. *Nature*, **510**, 139.

Nowak, R. S., Ellsworth, D. S., & Smith, S. D. (2004) Functional responses of plants to elevated atmospheric CO_2–do photosynthetic and productivity data from FACE experiments support early predictions? *New Phytologist*, **162**, 253-280.

Onoda, Y., Hirose, T., & Hikosaka, K. (2009) Does leaf photosynthesis adapt to CO_2-enriched environments? An experiment on plants originating from three natural CO_2 springs. *New Phytologist*, **182**, 698-709.

Paoletti, E., Pfanz, H., & Raschi, A. (2005) Pros and cons of CO_2 springs as experimental sites. In Plant responses to air pollution and global change. pp. 195-202. Springer, Tokyo.

Pearson, P. N., & Palmer, M. R. (2000) Atmospheric carbon dioxide concentrations over the past 60 million years. *Nature*, **406**, 695.

Poorter, H., Bühler, J., van Dusschoten, D., Climent, J., & Postma, J. A. (2012) Pot size matters: a meta-analysis of the effects of rooting volume on plant growth. *Functional Plant Biology*, **39**, 839-850.

Raschi, A., Miglietta, F., Tognetti, R., & van Gardingen, P. (Eds.) (1997) Plant responses to elevated CO_2: evidence from natural springs. Cambridge University Press, Cambridge, UK.

Rodrigues, W. P., Martins, M. Q., Fortunato, A. S., Rodrigues, A. P., Semedo, J. N., Simões-Costa, M. C., Leitão, A.E., Colwell, F., Goulao, L., & Máguas, C. (2016) Long-term elevated air [CO_2] strengthens photosynthetic functioning and mitigates the impact of supra-optimal temperatures in tropical *Coffea arabica* and *C. canephora* species. *Global Change Biology*, **22**, 415-431.

Ross, D. J., Tate, K. R., Newton, P. C. D., Wilde, R. H., & Clark, H. (2000) Carbon and nitrogen pools and mineralization in a grassland gley soil under elevated carbon dioxide at a natural CO_2 spring. *Global Change Biology*, **6**, 779-790.

Ruiz-Vera, U. M., De Souza, A. P., Long, S. P., & Ort, D. R. (2017) The role of sink strength and nitrogen availability in the down-regulation of photosynthetic capacity in field-grown *Nicotiana tabacum* L. at elevated CO_2 concentration. *Frontiers in Plant Science*, **8**, 998.

Scholefield, P. A., Doick, K. J., Herbert, B. M. J., Hewitt, C. S., Schnitzler, J. P., Pinelli, P., & Loreto, F. (2004) Impact of rising CO_2 on emissions of volatile organic compounds: isoprene emission from *Phragmites australis* growing at elevated CO_2 in a natural carbon dioxide spring. *Plant, Cell & Environment*, **27**, 393-401.

Stocker, T. F. (2013) The closing door of climate targets. *Science*, **339**, 280-282.

Terrer, C., Vicca, S., Stocker, B. D., Hungate, B. A., Phillips, R. P., Reich, P. B., Finzi, A. C., & Prentice, I. C. (2018) Ecosystem responses to elevated CO_2 governed by plant–soil interactions and the cost of nitrogen acquisition. *New Phytologist*, **217**, 507-522.

Ueda, M. U., Onoda, Y., Kamiyama, C., & Hikosaka, K. (2017) Decades-long effects of high CO_2 concentration on soil nitrogen dynamics at a natural CO_2 spring. *Ecological research*, **32**, 215-225.

Van Loon, M. P., Rietkerk, M., Dekker, S. C., Hikosaka, K., Ueda, M. U., & Anten, N. P. (2016) Plant–plant interactions mediate the plastic and genotypic response of *Plantago asiatica* to CO_2: an experiment with plant populations from naturally high CO_2 areas. *Annals of botany*, **117**, 1197-1207.

Vanuytrecht, E., & Thorburn, P. J. (2017) Responses to atmospheric CO_2 concentrations in crop simulation models: a review of current simple and semicomplex representations and options for model development. *Global Change Biology*, **23**, 1806-1820.

Vodnik, D., Pfanz, H., Wittmann, C., Macek, I., Kastelec, D., Turk, B., & Batic, F. (2002) Photosynthetic acclimation in plants growing near a carbon dioxide spring. *Phyton*, **42**, 239-244.

Wang, D., Heckathorn, S. A., Wang, X., & Philpott, S. M. (2012) A meta-analysis of plant physiological and growth responses to temperature and elevated CO_2. *Oecologia*, **169**, 1-13.

Wang, X., Taub, D. R., & Jablonski, L. M. (2015) Reproductive allocation in plants as affected by elevated carbon dioxide and other environmental changes: a synthesis using meta-analysis and graphical vector analysis. *Oecologia*, **177**, 1075-1087.

Ward, J., & Strain, B. (1997) Effects of low and elevated CO_2 partial pressure on growth and reproduction of *Arabidopsis thaliana* from different elevations. *Plant, Cell & Environment*, **20**, 254-260.

Watson-Lazowski, A., Lin, Y., Miglietta, F., Edwards, R. J., Chapman, M. A., & Taylor, G. (2016) Plant adaptation or acclimation to rising CO_2? Insight from first multigenerational RNA-Seq transcriptome. *Global Change Biology*, **22**, 3760-3773.

Wicker, T., Yu, Y., Haberer, G., Mayer, K. F., Marri, P. R., Rounsley, S., Chen, M., Zuccolo, A., Panaud, O., Wing, R. A., & Roffler, S. (2016) DNA transposon activity is associated with increased mutation rates in genes of rice and other grasses. *Nature communications,* **7**, 12790.

Woodward, F. I. (1987) Stomatal numbers are sensitive to increases in CO_2 from pre-industrial levels. *Nature*, **327**, 617.

Xu, Z., Jiang, Y., Jia, B., & Zhou, G. (2016) Elevated-CO_2 response of stomata and its dependence on environmental factors. *Frontiers in Plant Science*, **7**, 657.

Zhang, W., Sun, X., Yuan, H., Araki, H., Wang, J., & Tian, D. (2008) The pattern of insertion/deletion polymorphism in *Arabidopsis thaliana*. *Mol. Genet. Genomics,* **280**, 351-361.

Zheng, Y., Li, F., Hao, L., Shedayi, A. A., Guo, L., Ma, C., Huang, B., & Xu, M. (2018) The optimal CO_2 concentrations for the growth of three perennial grass species. *BMC plant biology*, **18**, 27.

Appendix A **Chapter 2 Supplementary Information**

Appendix A.1 | Extended materials and methods information

Appendix A.2 | Supplementary references

Appendix A.3 | Supplementary Tables and Figures

Appendix Table A.1 | Table of spring sites used to study plant responses to elevated [CO_2]

Appendix Table A.2 | Tests for heterogeneity in data collected for each trait measured across natural CO_2 spring sites

Appendix Table A.3 | Publication bias statistics for traits where publication bias was detected in CO_2 spring meta-analysis

Appendix Figure A.1 | Photosynthetic rate percentage difference between plants at elevated and ambient [CO_2] at naturally occurring CO_2 springs from individual studies included in this meta-analysis

Appendix Figure A.2 | Stomatal conductance (g_s) percentage difference between plants at elevated and ambient [CO_2] at naturally occurring CO_2 springs from individual studies included in this meta-analysis

A.1 Extended materials and methods information

Search strings used for systematic search on 3rd July 2017 were as follows:
CO_2 OR carbon dioxide) AND (spring OR natural OR enrichment/enriched OR mofette OR outgassing OR vent) AND (plant/s)

Systematic searches of the literature returned 3,294 studies which were screened for relevance by title, and relevant titles were then assessed by abstract for their potential to meet the following stringent inclusion criteria:

1) Publications must study naturally growing plant species at both a (terrestrial) spring site and a local control site with similar environmental conditions
2) There must be a minimum difference of 100 ppm in average daily CO_2 concentration between designated spring and control sites
3) Control sites must have an average daily CO_2 concentration below 435 ppm and spring sites must have an average daily CO_2 concentration above 465 ppm
4) Traits measured must be quantitative for inclusion in the meta-analysis (for example studies only reporting presence/absence of species were not included)
5) At least three individuals must be sampled from each site per species, and at least two measurements must be made per plant (where the type of measurement taken allowed for this).
6) As required for effect size calculation, traits are only included if mean trait value, a measure of variance (standard error or standard deviation) and sample size are given in the study
7) Measurements of plants taken from springs with contamination by $[H_2S] > 0.02$ ppm or $[SO_2] > 0.015$ppm are not included in this analysis

A.2 Supplementary references

Data provided to meta-analyses

1. Bettarini, I., Vaccari, F. P., & Miglietta, F. (1998). Elevated CO2 concentrations and stomatal density: observations from 17 plant species growing in a CO2 spring in central Italy. Global Change Biology, 4, 17-22.

2. Blaschke, L., Schulte, M., Raschi, A., Slee, N., Rennenberg, H., & Polle, A. (2001). Photosynthesis, soluble and structural carbon compounds in two Mediterranean oak species (Quercus pubescens and Q. ilex) after lifetime growth at naturally elevated CO2 concentrations. Plant Biology, 3, 288-298.

3. Chaves, M. M., Pereira, J. S., Cerasoli, S., Clifton-Brown, J., Miglietta, F., & Raschi, A. (1995). Leaf metabolism during summer drought in Quercus ilex trees with lifetime exposure to elevated CO2. Journal of Biogeography, 255-259.

4. Körner, C., & Miglietta, F. (1994). Long term effects of naturally elevated CO 2 on mediterranean grassland and forest trees. Oecologia, 99, 343-351.

5. Miglietta, F., Raschi, A., Körner, C., & Vaccari, F. P. (1998). Isotope discrimination and photosynthesis of vegetation growing in the Bossoleto CO2 spring. Chemosphere, 36, 771-776.

6. Rapparini, F., Baraldi, R., Miglietta, F., & Loreto, F. (2004). Isoprenoid emission in trees of Quercus pubescens and Quercus ilex with lifetime exposure to naturally high CO2 environment. Plant, Cell & Environment, 27, 381-391.

7. 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 & Environment, 27, 393-401.

8. Schwanz, P., & Polle, A. (1998). Antioxidative systems, pigment and protein contents in leaves of adult Mediterranean oak species (Quercus pubescens and Q. ilex) with lifetime exposure to elevated CO 2. The New Phytologist, 140, 411-423.

9. Stylinski, C. D., Oechel, W. C., Gamon, J. A., Tissue, D. T., Miglietta, F., & Raschi, A. (2000). Effects of lifelong [CO2] enrichment on carboxylation and light utilization of Quercus pubescens Willd. examined with gas exchange, biochemistry and optical techniques. Plant, Cell & Environment, 23, 1353-1362.

10. Tognetti, R., Longobucco, A., Miglietta, F., & Raschi, A. (1998). Transpiration and stomatal behaviour of Quercus ilex plants during the summer in a Mediterranean carbon dioxide spring. Plant, Cell & Environment, 21, 613-622.

11. Tognetti, R., Longobucco, A., Miglietta, F., & Raschi, A. (1999). Water relations, stomatal response and transpiration of Quercus pubescens trees during summer in a Mediterranean carbon dioxide spring. Tree Physiology, 19, 261-270.

12. Tognetti, R., Johnson, J. D., Michelozzi, M., & Raschi, A. (1998). Response of foliar metabolism in mature trees of Quercus pubescens and Quercus ilex to long-term elevated CO2. Environmental and Experimental Botany, 39, 233-245.

13. Tognetti, R., Minnocci, A., Peñuelas, J., Raschi, A., & Jones, M. B. (2000). Comparative field water relations of three Mediterranean shrub species co-occurring at a natural CO2 vent. Journal of Experimental Botany, 51, 1135-1146.

14. Jones, M. B., Brown, J. C., Raschi, A., & Miglietta, F. (1995). The effects on Arbutus unedo L. of long-term exposure to elevated CO2. Global Change Biology, 1, 295-302.

15. Peñuelas, J., Castells, E., Joffre, R., & Tognetti, R. (2002). Carbon-based secondary and structural compounds in Mediterranean shrubs growing near a natural CO2 spring. Global Change Biology, 8, 281-288.

16. Peñuelas, J., Filella, I., & Tognetti, R. (2001). Leaf mineral concentrations of Erica arborea, Juniperus communis and Myrtus communis growing in the proximity of a natural CO2 spring. Global Change Biology, 7, 291-301.

17. Tognetti, R., & Peñuelas, J. (2003). Nitrogen and carbon concentrations, and stable isotope ratios in Mediterranean shrubs growing in the proximity of a CO 2 spring. Biologia Plantarum, 46, 411-418.

18. Cook, A. C., Tissue, D. T., Roberts, S. W., & Oechel, W. C. (1998). Effects of long-term elevated [CO2] from natural CO2 springs on Nardus stricta: photosynthesis, biochemistry, growth and phenology. Plant, Cell & Environment, 21, 417-425.

19. Marchi, S., Tognetti, R., Vaccari, F. P., Lanini, M., Kaligarič, M., Miglietta, F., & Raschi, A. (2004). Physiological and morphological responses of grassland species to elevated atmospheric CO2 concentrations in FACE-systems and natural CO2 springs. Functional plant biology, 31, 181-194

20. Pfanz, H., Vodnik, D., Wittmann, C., Aschan, G., Batic, F., Turk, B., & Macek, I. (2007). Photosynthetic performance (CO2-compensation point, carboxylation efficiency, and net photosynthesis) of timothy grass (Phleum pratense L.) is affected by elevated carbon dioxide in post-volcanic mofette areas. Environmental and Experimental Botany, 61, 41-48.

21. Vodnik, D., Pfanz, H., Wittmann, C., Macek, I., Kastelec, D., Turk, B., & Batic, F. (2002). Photosynthetic acclimation in plants growing near a carbon dioxide spring. PHYTON-HORN-, 42, 239-244.

22. Onoda, Y., Hirose, T., & Hikosaka, K. (2007). Effect of elevated CO 2 levels on leaf starch, nitrogen and photosynthesis of plants growing at three natural CO 2 springs in Japan. Ecological research, 22, 475-484.

23. Stock, W. D., Ludwig, F., Morrow, C., Midgley, G. F., Wand, S. J., Allsopp, N., & Bell, T. L. (2005). Long-term effects of elevated atmospheric CO 2 on species composition and productivity of a southern African C 4 dominated grassland in the vicinity of a CO 2 exhalation. Plant Ecology, 178, 211-224.

24. Fernandez, M. D., Pieters, A., Donoso, C., Tezara, W., Azkue, M., Herrera, C., ... & Herrera, A. (1998). Effects of a natural source of very high CO 2 concentration on the leaf gas exchange, xylem water potential and stomatal characteristics of plants of Spatiphylum cannifolium and Bauhinia multinervia. The New Phytologist, 138, 689-697.

25. Marín, O., Rengifo, E., Herrera, A., & Tezara, W. (2005). Seasonal changes in water relations, photosynthesis and leaf anatomy of two species growing along a natural CO2 gradient. Interciencia, 30, 33-38.

26. Hättenschwiler S, Miglietta F, Raschi A, Körner C (1997) Morphological adjustments of mature Quercus ilex trees to elevated CO2. Acta Oecologica, 18, 361-365

Other sites that have been used to study plant response to elevated [CO_2]

Vejpustková, M., Thomalla, A., Cihak, T., Lomsky, B., & Pfanz, H. (2016) Growth of Populus tremula on CO_2-enriched soil at a natural mofette site. *Dendrobiology*, *75*.

Krüger, M., Jones, D., Frerichs, J., Oppermann, B. I., West, J., Coombs, P., Green, K., Barlow, T., Lister, R., Shaw, R., Strutt, M., & Moller, I. (2011) Effects of elevated CO_2 concentrations on the vegetation and microbial populations at a terrestrial CO_2 vent at Laacher See, Germany. *International Journal of Greenhouse Gas Control*, **5**, 1093-1098

Bartak, M., Raschi, A., & Tognetti, R. (1999) Photosynthetic characteristics of sun and shade leaves in the canopy of *Arbutus unedo* L. trees exposed to in situ long-term elevated CO_2. *Photosynthetica*, **37**, 1-16.

Miglietta, F., & Raschi, A. (1993) Studying the effect of elevated CO_2 in the open in a naturally enriched environment in Central Italy. *Vegetation*, **104**, 391-400.

Onoda, Y., Hikosaka, K., & Hirose, T. (2005) Natural CO_2 springs in Japan: A case study of vegetation dynamics. *Phyton*, **45**, 389-394.

Van Loon, M. P., Rietkerk, M., Dekker, S. C., Hikosaka, K., Ueda, M. U., & Anten, N. P. (2016). Plant–plant interactions mediate the plastic and genotypic response of *Plantago asiatica* to CO_2: an experiment with plant populations from naturally high CO_2 areas. *Annals of botany*, ***117***, 1197-1207.

Newton, P. C. D., Bell, C. C., & Clark, H. (1996) Carbon dioxide emissions from mineral springs in Northland and the potential of these sites for studying the effects of elevated carbon dioxide on pastures. *New Zealand Journal of Agricultural Research*, **39**, 33-40.

Ehleringer, J. R., Sandquist, D. R., & Philips, S. L. (1997) Burning coal seams in southern Utah: a natural system for studies of plant responses to elevated CO_2. In *Plant Responses to Elevated CO_2: Evidence from Natural Springs* (pp. 56-68). Cambridge University Press.

Woodward, F. I., Thompson, G. B., & McKee, I. F. (1991).The effects of elevated concentrations of carbon dioxide on individual plants, populations, communities and ecosystems. *Annals of Botany*, 23-38.

Sharma, S., & Williams, D. G. (2009) Carbon and oxygen isotope analysis of leaf biomass reveals contrasting photosynthetic responses to elevated CO_2 near geologic vents in Yellowstone National Park. *Biogeosciences*, **6**, 25-31.

Sajna, N., Meister, M., Bolhà r-Nordenkampf, H. R., & Kaligaric, M. (2013) Response of Semi-natural Wet Meadow to Natural Geogenic CO_2 Enrichment. *International Journal of Agriculture & Biology*, **15**.

A.3 Supplementary Tables and Figures

Appendix Table A. 1: Table of spring sites used to study plant responses to elevated [CO_2]

Country	Site	Latitude	Longitude	Gas composition of at the spring study site			Soil pH		Koepper climate classification
				[CO_2]	[H_2S]	[SO_2]	Spring site	Control site	
Iceland	Ólafsvík	64.9	-23.7	519-1179 ppmv	< 0.5 ppmv No smell (< 0.025 ppmv)	-	-	-	Polar tundra (ET)
Italy	Armaiolo	43.3	11.6	~500-2300 ppmmol	-	-	-	-	Temperate with dry hot summer (Csa)
	Bossoleto	43.3	11.6	400-1200 ppmmol	0.022 ppmv	0.012 ppmv	5.6-6.8	7.5-7.7	Temperate with dry hot summer (Csa)
	I Borboi	43.4	10.7	500-1500 ppmmol	0.060 ppmmol	0.004 ppmmol		6-7	Temperate with dry hot summer (Csa)
	Laiatico	43.4	10.8	400-1500 ppmmol	0.022 ppmv	0.004 ppmv	6-7	6-7	Temperate with dry hot summer (Csa)
Slovenia	Strmec	46.7	16	500-1000 ppmmol	-	-	5.0-5.2	5.0-5.2	Temperate without dry season, warm summer (Cfb)
Japan	Nibu	38.5	140.0	~450-850 ppmmol	< 0.1 ppmv No smell (< 0.025 ppmv)	-	3.6-4.2	4.1-4.3	Temperate without dry season, hot summer (Cfa)
	Ryuzin-numa	40.7	141.0	~550-890 ppmmol	< 0.1 ppmv No smell (< 0.025 ppmv)	-	3.5-3.7	2.8-3.4	Temperate without dry season, hot summer (Cfa)
	Yuno kawa	40.7	141.0	~460-630 ppmmol	< 0.1 ppmv No smell (< 0.025 ppmv)	-	3.7-4.5	3.7-4.5	Temperate without dry season, hot summer (Cfa)
South Africa	Pleasant View	-30.7	30.02	~480-600 ppmmol	-	-	4.2-4.4	4.3-4.5	Temperate without dry season, warm summer (Cfb)
Venezuela	Sta. Ana	10.6	-63.13	(S1) ~34200-35800 ppmmol at the vent (S2) ~26800-27200 ppmmol at the vent	< 0.1 ppmmol	-	-	-	Tropical savannah (Aw)

				CO_2 concentrations approx. 1000					
Czech Republic	Plesná stream	50.1	12.46	>600 ppm at 50 cm vertical	-	-	-	-	Temperate without dry season, warm summer (Cfb)
Germany	Laacher See	50.4	7.25	Gradient 100-0% explored	-	-	4.0-6.0	5.5-6.3	Temperate oceanic climate (Cfb)
Italy	Orciatico	43.4	10.67	Avg. 465 ppmmol	-	-	-	-	Temperate with dry hot summer (Csa)
	Solfatara	42.5	12.13	450-850 ppmmol	0.245 ppmv	0.018ppmv	3.3-2.1	4.5-4.1	Temperate with dry hot summer (Csa)
Japan	Tashiro	40.7	140.92	400-1000 ppmmol	< 0.03 ppmmol	<0.03ppmmol	-	-	Temperate without dry season, hot summer (Cfa)
	Asahi	38.2	140	2123-2509 ppm	-	-	-	-	Temperate without dry season, hot summer (Cfa)
	Kosaka	40.4	140.8	503-7019 ppm	-	-	-	-	Temperate without dry season, hot summer (Cfa)
New Zealand	Hakanoa springs	-35.7	174.27	480-725 ppmv	< 0.18 ppmv	-	5.2-5.7	5.2-5.7	Temperate without dry season, warm summer (Cfb)
USA	Burning hills	37.3	-111.37	400-1000 ppm	-	-	-	-	Cold-desert climate (Bwk)
	Ichetucknee river springs	30.0	-82.76	450-500 ppmmol	-	-	-	-	Temperate without dry season, hot summer (Cfa)
	Ochre Springs	44.6	-110.4	419-482 ppmv	-	-	-	-	Dry summer subarctic (Dsc)
	Mammoth Upper Terrace	45.0	-110.7	401-607 ppmv	-	-	-	-	Warm-summer continental (Dfb)
Slovenia	Rihtarovci	46.6	16.1	Gradient 400 to >2500 ppmmol	-	-	-	-	Temperate without dry season, warm summer (Cfb)

Appendix Table A. 2: Tests for heterogeneity in data collected for each trait measured across natural CO_2 spring sites

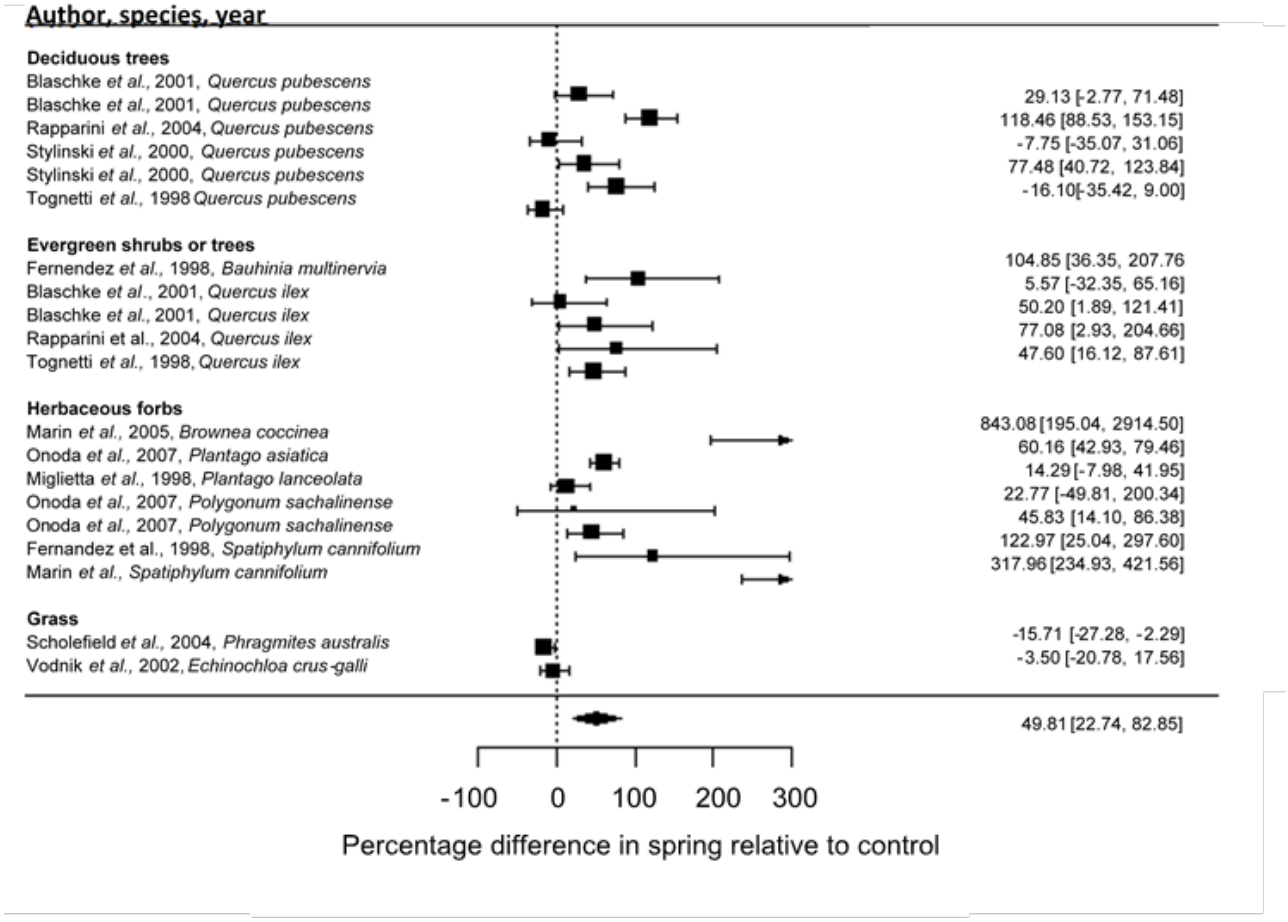
Statistic	Q stat	Df	p-val	I^2
Description	Is the variability in the observed effect size larger than would be expected based on sampling variability alone?			Percentage of total variation across studies that is due to heterogeneity rather than chance
Stomatal conductance	360.6	31	<0.0001	92.28%
Abaxial stomatal index	214.3	22	<0.0001	88.75%
Adaxial stomatal index	109.5	7	<0.0001	99.13%
Abaxial stomatal density	279.8	24	<0.0001	93.93%
Adaxial stomatal density	100.8	8	<0.0001	98.38%
Photosynthetic rate	180.7	8	<0.0001	96.47%
V_{cmax}	13.9	6	0.0312	56.42%
J_{max}	21.1	6	0.017	71.17%
Leaf chlorophyll content	34.8	7	<0.0001	80.88%
Leaf carbon content	63.3	14	<0.0001	78.69%
Leaf sugar content	172.5	4	<0.0001	95.94%
Leaf starch content	236.3	12	<0.0001	96.04%
Leaf total non-structural carbohydrate content	311.8	10	<0.0001	97.00%
Leaf nitrogen content	232.5	22	<0.0001	89.79%
Leaf carbon:nitrogen ratio	8.1	4	0.0893	49.87%
Specific leaf area	33.5	4	<0.0001	84.58%

Appendix Table A. 3: Publication bias statistics for traits where publication bias was detected in CO_2 spring meta-analysis

	Egger's test for funnel plot asymmetry		Rosenthal's Fail-safe number	
Trait	**T test**	**P value**	**Fail-safe number**	**5N +10**
Abaxial stomatal index	2.7938	0.0109	92	125
Adaxial stomatal density	-3.1970	0.0151	95	55
Leaf chlorophyll content	2.5425	0.0439	0?	50
Leaf carbon content	-2.3206	0.0372	27	85

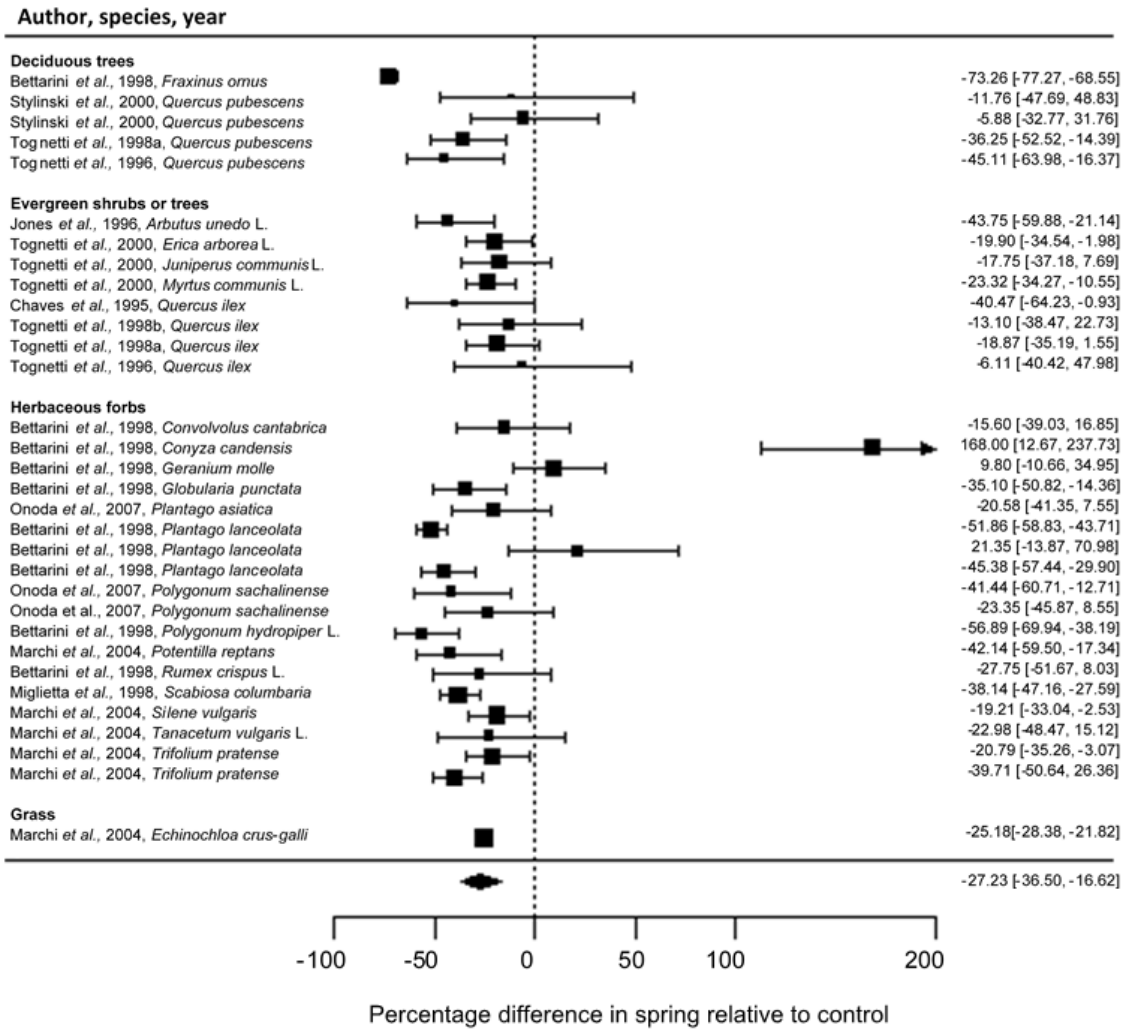
Appendix Figure A. 1: Photosynthetic rate percentage difference between plants at elevated and ambient [CO_2] at naturally occurring CO_2 springs from individual studies included in this

Author(s) and species appear on the left hand side, numbers on the right hand side are effect size, with 95% confidence intervals in parentheses. Squares indicating mean effect size are drawn proportionally to the precision of the estimate. The summary polygon at the bottom of the plot indicates the mean effects size when all 20 estimates are analysed together using a random effects model. Note that in subgroup analysis plants were categorised as trees, which included both 'deciduous trees' and 'evergreen shrubs or trees', or herbs, which included 'herbaceous forbs' and grasses, rather than the functional groups used for visualisation here.



Appendix Figure A. 2: Stomatal conductance (g_s) percentage difference between plants at elevated and ambient [CO_2] at naturally occurring CO_2 springs from individual studies included in this meta-analysis.

Author(s) and species appear on the left hand side, numbers on the right hand side are effect size, with 95% confidence intervals in parentheses. Squares indicating mean effect size are drawn proportionally to the precision of the estimate. The summary polygon at the bottom of the plot indicates the mean effects size when all 32 estimates are analysed together using a random effects model. Note that in subgroup analysis plants were categorised as trees, which included both 'deciduous trees' and 'evergreen shrubs or trees' or herbs, which included 'herbaceous forbs' and grasses, or just herbaceous forbs, rather than the functional groups used for visualisation here.



Appendix B **Chapter 3 Supplementary Information**

Appendix B.1 | Supplementary Tables and Figures

Appendix Table B.1 | Estimations of heritability and transgenerational indirect genetic effects when the fixed term 'Generation' was not included in the model.

Appendix Table B.2 | Estimations of fixed effects when the fixed effect term 'Generation' was not included in the model.

Appendix Table B.3 | Posterior mode of variance estimates for each fixed effect and interaction effect models of reproductive and fitness traits where interactions were not significant.

Appendix Table B.4 | Posterior mode of variance estimates for each fixed effect and interaction effect for models of abaxial leaf traits where interaction effects were not significant.

Appendix Table B.5 | The effect of the specification of different priors for random effects in the G structure for reproductive and fitness traits modelled using a Gaussian distribution.

Appendix Table B.6 | The effect of the specification of different priors for random effects in the G structure for abaxial leaf traits modelled using a Gaussian distribution.

Appendix Figure B.1 | The impact of discretization on the distribution of trait values for traits categorised and modelled using an ordinal distribution.

Appendix Table B.7| Posterior mode estimates of heritability and transgenerational indirect genetic effects for number of seeds per spike and seed weight when alternative categorisation is used.

Appendix Table B.8 | Posterior mode estimates of fixed effects for number of seeds per spike and seed weight when alternative categorisation is used.

Appendix Table B.9 | The effect of the specification of different priors for random effects for biomass traits modelled using an ordinal distribution.

Appendix Figure B.2 | Q-Q plots used to evaluate the assumption of normality for biomass traits.

Appendix Table B.10 | Effect of [CO_2] treatment and ancestry of plants in generation 3 on whether or not plants reached the floral to vegetative transition before harvest.

Appendix Figure B.3 | Pairwise correlation between traits in a three generation experiment with *P. lanceolata.*

Appendix Table B.11 | Effect estimates of [CO_2] treatment, ancestral location (Cross), generational environment (Generation) and interaction effects on *Plantago lanceolata* reproductive and fitness traits.

Appendix Figure B.4 | Raw multigenerational reproductive and fitness traits of *Plantago lanceolata* in response to ambient and elevated growth [CO_2].

Appendix Figure B.5 | Raw multigenerational abaxial stomatal patterning traits of *Plantago lanceolata* in response to ambient and elevated growth [CO_2].

Appendix Figure B.6 | Mean biomass components of *P. lanceolata* originating from spring and control populations (or hybrids of these ancestries) when grown in ambient or elevated [CO_2] following two generations of growth at ambient [CO_2].

B.1 Supplementary Tables and Figures

Appendix Table B. 1: Estimations of heritability and transgenerational indirect genetic effects when the fixed term 'Generation' was not included in the model.

Estimates are given as the posterior mode with 95% Credible Interval lower (95 % CI Lb) and upper bounds (95 % CI Ub). The removal of this term resulted in a calculated paternal effect for abaxial stomatal density but did not significantly affect the calculation of heritability or parental effects for any other leaf morphology trait or any reproductive traits.

	Posterior mode	95% Lb CI	95% Ub CI
Log days to germination			
Heritability	0.16	0.00	0.40
Maternal effects	0.34	0.00	0.56
Paternal effects	0.09	0.00	0.34
Log days to first spike			
Heritability	0.01	0.00	0.03
Maternal effects	0.01	0.00	0.03
Paternal effects	0.91	0.87	0.96
Number of rosettes			
Heritability	0.11	0.00	0.39
Maternal effects	0.06	0.00	0.23
Paternal effects	0.16	0.00	0.38
Number of seeds per spike			
Heritability	0.00	0.00	0.02
Maternal effects	0.64	0.46	0.84
Paternal effects	0.04	0.00	0.14
Maternal grandmaternal effects	0.25	0.05	0.46
Paternal grandpaternal effects	0.04	0.00	0.13
Seed weight			
Heritability	0.00	0.00	0.02
Maternal effects	0.78	0.57	0.96
Paternal effects	0.11	0.00	0.29
Maternal grandmaternal effects	0.03	0.00	0.10
Paternal grandpaternal effects	0.06	0.00	0.20
Log single leaf area			
Heritability	0.07	0.00	0.22
Maternal effects	0.07	0.00	0.20
Paternal effects	0.06	0.00	0.23
Log single leaf dry weight			
Heritability	0.24	0.00	0.45
Maternal effects	0.06	0.00	0.16
Paternal effects	0.04	0.00	0.15

Log abaxial stomatal conductance			
Heritability	0.03	0.00	0.11
Maternal effects	0.04	0.00	0.16
Paternal effects	0.04	0.00	0.14
Log abaxial stomatal density			
Heritability	0.02	0.00	0.07
Maternal effects	0.02	0.00	0.09
Paternal effects	0.73	0.58	0.85
Log abaxial stomatal index			
Heritability	0.13	0.00	0.33
Maternal effects	0.06	0.00	0.21
Paternal effects	0.09	0.00	0.28
Log abaxial epidermal cell area			
Heritability	0.03	0.00	0.10
Maternal effects	0.04	0.00	0.17
Paternal effects	0.40	0.07	0.69
Log abaxial guard cell length			
Heritability	0.04	0.00	0.14
Maternal effects	0.05	0.00	0.20
Paternal effects	0.35	0.07	0.61

Appendix Table B. 2: Estimations of fixed effects when the fixed effect term 'Generation' was not included in the model.

Estimates are given as the posterior mode with 95 % Credible Interval lower 95 % CI Lb) and upper bounds (95 % CI Ub) and p-value (pMCMC) significance indicated as *** p<0.001, ** p<0.01, *p<0.01, † p<0.1.

	Posterior mode	95% Lb CI	95% Ub CI	pMCMC	
Log days to germination					
Intercept	2.73	2.37	3.12	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$	-0.21	-0.35	-0.08	0.00	**
Population $_{CONTROL - SPRING}$	0.33	-0.18	0.83	0.20	
Population $_{CONTROL - HYBRID}$	0.07	-0.37	0.53	0.75	
Log days to first spike					
Intercept	4.40	4.230	4.49	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$	-0.09	-0.18	0.00	0.057	†
Population $_{CONTROL - SPRING}$	0.08	-0.05	0.21	0.212	
Population $_{CONTROL - HYBRID}$	-0.73	-1.04	-0.36	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL - SPRING}$	0.17	0.03	0.32	0.027	*
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL - HYBRID}$	-0.03	-0.15	0.11	0.696	
Number of rosettes					
Intercept	1.49	0.91	2.08	<3e-04	***
$[CO_2]$ treatment $_{AMB- ELEV}$	0.74	0.15	1.42	0.02	*
Population $_{CONTROL - SPRING}$	0.03	-0.70	0.73	0.90	
Population $_{CONTROL - HYBRID}$	-0.21	-1.04	0.65	0.64	
$[CO_2]$ treatment $_{AMB- ELEV}$ x Population $_{CONTROL - SPRING}$	-0.48	-1.46	0.58	0.37	
$[CO_2]$ treatment $_{AMB- ELEV}$ x Population $_{CONTROL - HYBRID}$	-1.15	-1.97	-0.23	0.01	**
Number of seeds per spike					
Intercept	3.10	-0.24	6.55	0.06	†
$[CO_2]$ treatment $_{AMB - ELEV}$	1.97	-1.13	5.23	0.22	
Population $_{CONTROL - SPRING}$	2.49	-1.86	6.72	0.27	
Population $_{CONTROL - HYBRID}$	2.94	-1.21	7.19	0.17	
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL - SPRING}$	-4.73	-10.40	0.83	0.09	†
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL - HYBRID}$	-5.91	-10.29	-1.24	0.01	**
Seed weight					
Intercept	3.94	0.02	7.99	0.04	*
$[CO_2]$ treatment $_{AMB - ELEV}$	6.78	2.63	11.29	0.00	**
Population $_{CONTROL - SPRING}$	0.14	-4.90	5.23	0.95	
Population $_{CONTROL - HYBRID}$	-4.42	-9.24	0.20	0.06	†
Log single leaf area					
Intercept	2.68	2.55	2.80	<3e-04	***

$[CO_2]$ treatment $_{AMB - ELEV}$	0.15	0.06	0.24	0.00	**
Population $_{CONTROL - SPRING}$	0.09	-0.07	0.27	0.30	
Population $_{CONTROL - HYBRID}$	0.21	0.02	0.40	0.02	*
Log leaf dry weight					
Intercept	4.74	4.58	4.87	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$	0.12	0.04	0.20	0.00	**
Population $_{CONTROL - SPRING}$	-0.07	-0.27	0.14	0.47	
Population $_{CONTROL - HYBRID}$	-0.12	-0.31	0.08	0.22	
Log abaxial stomatal conductance					
Intercept	5.46	5.30	5.67	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$	-0.59	-0.74	-0.46	<3e-04	***
Population $_{CONTROL - SPRING}$	0.07	-0.16	0.29	0.55	
Population $_{CONTROL - HYBRID}$	0.21	-0.04	0.45	0.10	†
Log abaxial stomatal density					
Intercept	5.23	5.12	5.36	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$	-0.04	-0.10	0.03	0.24	
Population $_{CONTROL - SPRING}$	-0.09	-0.24	0.08	0.28	
Population $_{CONTROL - HYBRID}$	-0.45	-0.70	-0.21	<3e-04	***
Log abaxial stomatal index					
Intercept	3.38	3.35	3.41	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$	0.02	0.00	0.04	0.05	*
Population $_{CONTROL - SPRING}$	0.01	-0.03	0.05	0.67	
Population $_{CONTROL - HYBRID}$	0.04	-0.01	0.09	0.08	†
Log abaxial epidermal cell area					
Intercept	7.67	7.46	7.87	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$	0.12	-0.01	0.24	0.06	†
Population $_{CONTROL - SPRING}$	0.11	-0.15	0.33	0.39	
Population $_{CONTROL - HYBRID}$	0.44	0.12	0.76	0.00	**
Log abaxial guard cell length					
Intercept	3.27	3.22	3.32	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$	0.04	0.01	0.07	0.01	**
Population $_{CONTROL - SPRING}$	0.02	-0.04	0.09	0.58	
Population $_{CONTROL - HYBRID}$	0.09	0.02	0.16	0.01	*

Appendix Table B. 3: Posterior mode of variance estimates for each fixed effect and interaction effect models of reproductive and fitness traits where interactions were not significant.

Estimates are given with 95 % Credible Interval lower (95 % CI Lb) and upper bounds (95 % CI Ub) and p-value (pMCMC) significance indicated as *** p<0.001, ** p<0.01, *p<0.01, † p<0.1. For the contrast estimates for the two categorical moderator variables [CO_2] treatment and population, contrasts are given in subscript.

	Posterior mode	95% Lb CI	95% Lb CI	pMCMC	
Log days to germination					
Intercept	4.16	3.50	4.83	<3e-04	***
[CO_2] treatment $_{AMB-ELEV}$	-0.11	-0.34	0.08	0.30	
Population $_{CONTROL-SPRING}$	0.32	-0.16	0.85	0.20	
Population $_{CONTROL-HYBRID}$	0.47	0.01	0.95	0.05	†
Generation	-0.58	-0.80	-0.36	<3e-04	***
[CO_2] treatment $_{AMB-ELEV}$ x Population $_{CONTROL-SPRING}$	0.09	-0.19	0.39	0.51	
[CO_2] treatment $_{AMB-ELEV}$ x Population $_{CONTROL-HYBRID}$	-0.16	-0.52	0.19	0.40	
Seed weight					
Intercept	11.80	5.93	17.73	0.00	***
[CO_2] treatment $_{AMB-ELEV}$	1.98	-1.45	4.94	0.21	
Population $_{CONTROL-SPRING}$	2.12	-2.59	6.70	0.36	
Population $_{CONTROL-HYBRID}$	3.82	-0.21	8.31	0.07	†
Generation	0.12	-0.73	0.89	0.77	
[CO_2] treatment $_{AMB-ELEV}$ x Population $_{CONTROL-SPRING}$	-2.33	-8.15	2.85	0.42	
[CO_2] treatment $_{AMB-ELEV}$ x Population $_{CONTROL-HYBRID}$	-0.15	-4.58	4.24	0.94	
Log single leaf area					
Intercept	2.46	2.27	2.66	<3e-04	***
[CO_2] treatment $_{AMB-ELEV}$	0.05	-0.10	0.20	0.47	
Population $_{CONTROL-SPRING}$	0.06	-0.10	0.24	0.48	
Population $_{CONTROL-HYBRID}$	0.18	-0.02	0.39	0.10	†
Generation	0.11	0.03	0.18	0.01	**
[CO_2] treatment $_{AMB-ELEV}$ x Population $_{CONTROL-SPRING}$	0.14	-0.06	0.34	0.18	
[CO_2] treatment $_{AMB-ELEV}$ x Population $_{CONTROL-HYBRID}$	-0.04	-0.29	0.21	0.78	
Log leaf dry weight					
Intercept	4.82	4.61	5.03	<3e-04	***
[CO_2] treatment $_{AMB-ELEV}$	0.13	-0.01	0.26	0.07	†
Population $_{CONTROL-SPRING}$	-0.08	-0.27	0.14	0.45	
Population $_{CONTROL-HYBRID}$	-0.10	-0.31	0.13	0.38	
Generation	-0.04	-0.11	0.03	0.27	

$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL – SPRING}$	0.02	-0.17	0.21	0.88
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL – HYBRID}$	0.02	-0.22	0.24	0.82

Appendix Table B. 4: Posterior mode of variance estimates for each fixed effect and interaction effect for models of abaxial leaf traits where interaction effects were not significant.

Estimates are given with 95 % Credible Interval lower (95 % CI Lb) and upper bounds (95 % CI Ub) and p-value (pMCMC) significance indicated as *** p<0.001, ** p<0.01, *p<0.01, † p<0.1. Models included random effects of the individual, mothers, fathers, and additionally maternal grandmother and paternal grandfather for the two traits: 'number of seeds per spike' and 'seed weight'. For the contrast estimates for the two categorical moderator variables [CO_2] treatment and population, contrasts given in subscript.

	Posterior mode	95% Lb CI	95% Ub CI	pMCMC	
Log abaxial stomatal conductance					
Intercept	6.05	5.72	6.40	<3e-04	***
[CO_2] treatment $_{AMB - ELEV}$	-0.40	-0.61	-0.16	<3e-04	***
Population $_{CONTROL - SPRING}$	0.12	-0.09	0.35	0.28	
Population $_{CONTROL - HYBRID}$	0.30	0.02	0.57	0.03	*
Generation	-0.26	-0.39	-0.14	<3e-04	***
[CO_2] treatment $_{AMB - ELEV}$ x Population $_{CONTROL - SPRING}$	-0.20	-0.51	0.09	0.20	
[CO_2] treatment $_{AMB - ELEV}$ x Population $_{CONTROL - HYBRID}$	0.01	-0.34	0.38	0.96	
Log abaxial stomatal density					
Intercept	5.72	5.59	5.85	<3e-04	***
[CO_2] treatment $_{AMB - ELEV}$	0.00	-0.10	0.10	0.99	
Population $_{CONTROL - SPRING}$	-0.06	-0.16	0.02	0.16	
Population $_{CONTROL - HYBRID}$	-0.04	-0.16	0.08	0.54	
Generation	-0.33	-0.38	-0.29	<3e-04	***
[CO_2] treatment $_{AMB - ELEV}$ x Population $_{CONTROL - SPRING}$	-0.03	-0.16	0.11	0.66	
[CO_2] treatment $_{AMB - ELEV}$ x Population $_{CONTROL - HYBRID}$	-0.04	-0.20	0.12	0.63	
Log abaxial stomatal index					
Intercept	3.32	3.28	3.37	<3e-04	***
[CO_2] treatment $_{AMB - ELEV}$	0.02	-0.01	0.05	0.24	
Population $_{CONTROL - SPRING}$	0.01	-0.03	0.05	0.63	
Population $_{CONTROL - HYBRID}$	0.04	-0.01	0.09	0.16	
Generation	0.03	0.01	0.05	0.00	***
[CO_2] treatment $_{AMB - ELEV}$ x Population $_{CONTROL - SPRING}$	-0.01	-0.05	0.03	0.72	

$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL – HYBRID}$	-0.03	-0.08	0.02	0.21	
Log abaxial epidermal cell area					
Intercept	6.92	6.73	7.15	<3e-04	***
$[CO_2]$ treatment $_{AMB – ELEV}$	0.03	-0.14	0.20	0.77	
Population $_{CONTROL – SPRING}$	0.09	-0.07	0.25	0.25	
Population $_{CONTROL – HYBRID}$	0.02	-0.20	0.23	0.84	
Generation	0.42	0.34	0.50	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL – SPRING}$	-0.08	-0.30	0.16	0.51	
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL – HYBRID}$	0.08	-0.20	0.34	0.56	
Log abaxial guard cell length					
Intercept	3.13	3.06	3.19	<3e-04	***
$[CO_2]$ treatment $_{AMB – ELEV}$	0.01	-0.04	0.06	0.78	
Population $_{CONTROL – SPRING}$	0.02	-0.03	0.08	0.43	
Population $_{CONTROL – HYBRID}$	0.02	-0.06	0.08	0.67	
Generation	0.08	0.05	0.10	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL – SPRING}$	0.01	-0.05	0.08	0.67	
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL – HYBRID}$	0.03	-0.04	0.11	0.36	

Appendix Table B. 5: The effect of the specification of different priors for random effects in the G structure for reproductive and fitness traits modelled using a Gaussian distribution.

Priors tested were parameter expanded priors (presented model, V=1, ν=1, α.μ=0, α.V=1000), parameter expanded half Cauchy priors (V=1, ν=1, α.μ=0, α.V=25^2) and inverse gamma priors (V=1, ν=0.002). Inverse gamma priors were specified for fixed effects in all models. Estimates are given as posterior mode with 95 % Credible Interval lower (95 % CI Lb) and upper bounds (95 % CI Ub).

	Posterior mode	95% Lb CI	95% Ub CI
Log days to germination			
Parameter expanded priors			
Heritability	0.31	0.00	0.58
Maternal effects	0.23	0.00	0.45
Paternal effects	0.06	0.00	0.23
Parameter expanded half Cauchy priors			
Heritability	0.31	0.00	0.58
Maternal effects	0.23	0.00	0.44
Paternal effects	0.06	0.00	0.23
Inverse gamma priors			
Heritability	0.26	0.00	0.56
Maternal effects	0.25	0.00	0.47
Paternal effects	0.05	0.00	0.19
Log days to first spike			
Parameter expanded priors			
Heritability	0.01	0.00	0.06
Maternal effects	0.02	0.00	0.09
Paternal effects	0.69	0.48	0.85
Parameter expanded half Cauchy priors			
Heritability	0.01	0.00	0.05
Maternal effects	0.02	0.00	0.09
Paternal effects	0.69	0.49	0.85
Inverse gamma priors			
Heritability	0.02	0.00	0.06
Maternal effects	0.03	0.00	0.08
Paternal effects	0.67	0.46	0.84
Log single leaf area			
Parameter expanded priors			
Heritability	0.07	0.00	0.21
Maternal effects	0.06	0.00	0.18
Paternal effects	0.06	0.00	0.22
Parameter expanded half Cauchy priors			
Heritability	0.07	0.00	0.20
Maternal effects	0.06	0.00	0.18
Paternal effects	0.06	0.00	0.22
Inverse gamma priors			
Heritability	0.06	0.00	0.18

Maternal effects	0.05	0.00	0.15
Paternal effects	0.05	0.00	0.18
Log leaf dry weight			
Parameter expanded priors			
Heritability	0.26	0.00	0.47
Maternal effects	0.06	0.00	0.16
Paternal effects	0.04	0.00	0.14
Parameter expanded half Cauchy priors			
Heritability	0.26	0.00	0.48
Maternal effects	0.06	0.00	0.16
Paternal effects	0.04	0.00	0.14
Inverse gamma priors			
Heritability	0.23	0.01	0.44
Maternal effects	0.06	0.00	0.15
Paternal effects	0.04	0.00	0.13

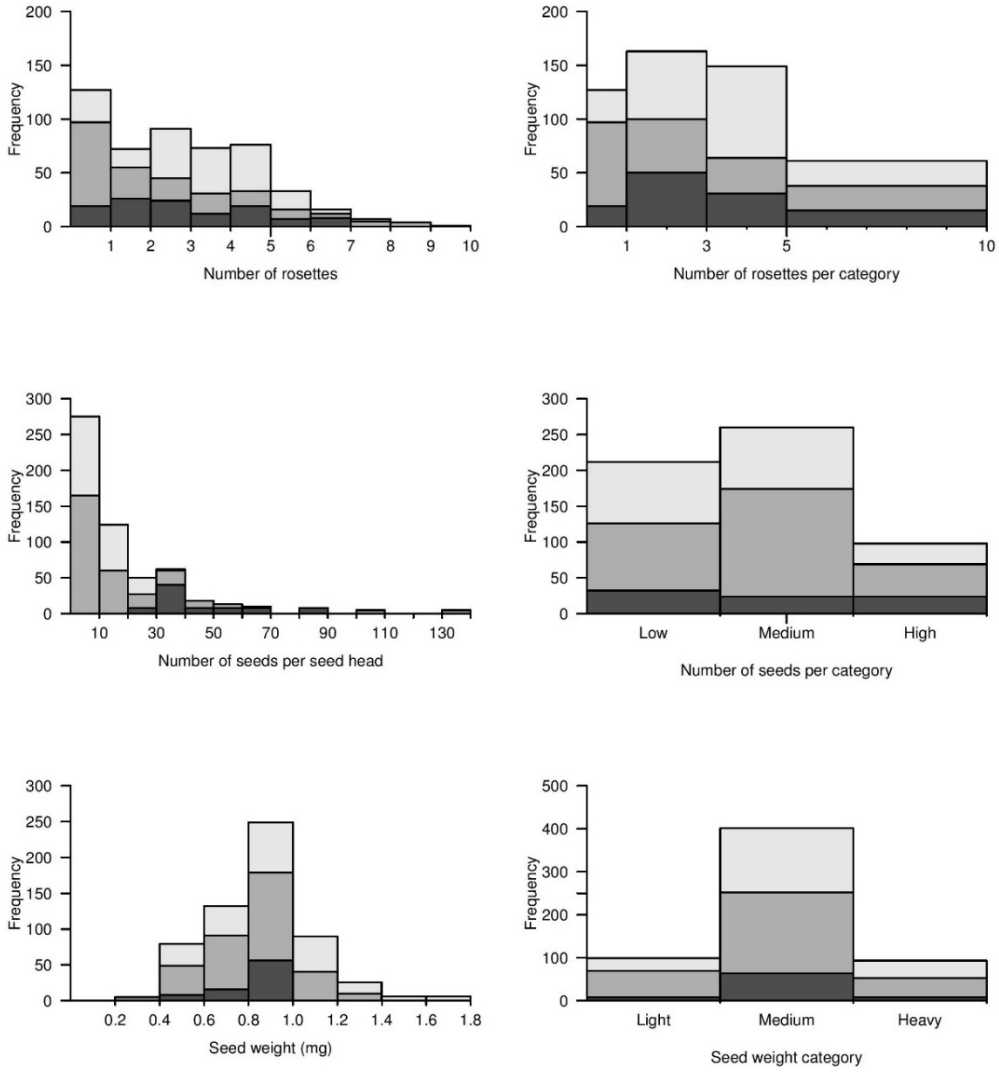
Appendix Table B. 6: The effect of the specification of different priors for random effects in the G structure for abaxial leaf traits modelled using a Gaussian distribution.

Priors tested were parameter expanded priors (presented model, V=1, ν=1, α.μ=0, α.V=1000), parameter expanded half Cauchy priors (V=1, ν=1, α.μ=0, α.V=25^2) and inverse gamma priors (V=1, ν=0.002). Inverse gamma priors were specified for fixed effects in all models. Estimates are given as posterior mode with 95 % Credible Interval lower (95 % CI Lb) and upper bounds (95 % CI Ub).

	Posterior mode	95% Lb CI	95% Ub CI
Log abaxial stomatal conductance			
Parameter expanded priors			
Heritability	0.03	0.00	0.09
Maternal effects	0.02	0.00	0.09
Paternal effects	0.02	0.00	0.09
Parameter expanded half Cauchy priors			
Heritability	0.03	0.00	0.09
Maternal effects	0.02	0.00	0.09
Paternal effects	0.03	0.00	0.10
Inverse gamma priors			
Heritability	0.02	0.00	0.08
Maternal effects	0.02	0.00	0.07
Paternal effects	0.02	0.00	0.08
Log abaxial stomatal density			
Parameter expanded priors			
Heritability	0.02	0.00	0.08
Maternal effects	0.02	0.00	0.07
Paternal effects	0.03	0.00	0.11
Parameter expanded half Cauchy priors			
Heritability	0.02	0.00	0.08
Maternal effects	0.02	0.00	0.07
Paternal effects	0.03	0.00	0.11
Inverse gamma priors			
Heritability	0.02	0.00	0.08
Maternal effects	0.02	0.00	0.07
Paternal effects	0.03	0.00	0.11
Log abaxial stomatal index			
Parameter expanded priors			
Heritability	0.11	0.00	0.29
Maternal effects	0.06	0.00	0.19
Paternal effects	0.07	0.00	0.23
Parameter expanded half Cauchy priors			
Heritability	0.11	0.00	0.29
Maternal effects	0.06	0.00	0.19
Paternal effects	0.07	0.00	0.26
Inverse gamma priors			
Heritability	0.14	0.02	0.30
Maternal effects	0.11	0.02	0.21

Paternal effects	0.12	0.02	0.26
Log abaxial epidermal cell area			
Parameter expanded priors			
Heritability	0.03	0.00	0.10
Maternal effects	0.02	0.00	0.07
Paternal effects	0.04	0.00	0.14
Parameter expanded half Cauchy priors			
Heritability	0.03	0.00	0.10
Maternal effects	0.02	0.00	0.07
Paternal effects	0.04	0.00	0.14
Inverse gamma priors			
Heritability	0.03	0.00	0.09
Maternal effects	0.02	0.00	0.06
Paternal effects	0.04	0.00	0.12
Log abaxial guard cell length			
Parameter expanded priors			
Heritability	0.04	0.00	0.14
Maternal effects	0.05	0.00	0.17
Paternal effects	0.19	0.00	0.39
Parameter expanded half Cauchy priors			
Heritability	0.04	0.00	0.13
Maternal effects	0.05	0.00	0.18
Paternal effects	0.19	0.00	0.41
Inverse gamma priors			
Heritability	0.04	0.00	0.13
Maternal effects	0.05	0.00	0.17
Paternal effects	0.20	0.00	0.41

Appendix Figure B. 1: The impact of discretization on the distribution of trait values for traits categorised and modelled using an ordinal distribution.



Appendix Table B. 7: Posterior mode estimates of heritability and transgenerational indirect genetic effects for number of seeds per spike and seed weight when alternative categorisation is used.

These traits are categorised into three categories according to whether they were less than 1 standard deviation from the generation mean, within a standard deviation or more than a standard deviation. 95 % Credible Interval lower (95% Lb CI) and upper bounds (95 % Ub CI) are also given.

	Posterior mode	95% Lb CI	95% Ub CI
Number of seeds per spike			
Heritability	0.00	0.00	0.02
Maternal effects	0.66	0.47	0.86
Paternal effects	0.05	0.00	0.17
Maternal grandmaternal effects	0.22	0.02	0.41
Paternal grandpaternal effects	0.04	0.00	0.13
Seed weight			
Heritability	0.00	0.00	0.02
Maternal effects	0.65	0.33	0.94
Paternal effects	0.17	0.00	0.50
Maternal grandmaternal effects	0.07	0.00	0.22
Paternal grandpaternal effects	0.08	0.00	0.23

Appendix Table B. 8: Posterior mode estimates of fixed effects for number of seeds per spike and seed weight when alternative categorisation is used.

These traits are categorised into three categories according to whether they were less than 1 standard deviation from the generation mean, within a standard deviation or more than a standard deviation. 95 % Credible Interval lower (95% Lb CI) and upper bounds (95 % Ub CI) are also given with p-value (pMCMC) significance indicated as *** p<0.001, ** p<0.01, *p<0.01, † p<0.1.

	Posterior mode	95% Lb CI	95% Ub CI	pMCMC	
Number of seeds per spike					
Intercept	5.63	1.56	9.65	0.08	†
$[CO_2]$ treatment $_{AMB - ELEV}$	3.18	-0.22	6.56	0.08	†
Population $_{CONTROL - SPRING}$	2.39	-2.26	6.74	0.15	
Population $_{CONTROL - HYBRID}$	3.62	-0.93	7.73	0.25	
Generation	-0.99	-1.89	-0.15	0.07	†
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL - SPRING}$	-4.42	-10.20	1.48	0.13	
$[CO_2]$ treatment $_{AMB - ELEV}$ x Population $_{CONTROL - HYBRID}$	-7.41	-12.10	-2.50	0.00	
Seed weight					
Intercept	9.31	2.83	15.70	<3e-04	***
$[CO_2]$ treatment $_{AMB - ELEV}$	5.15	-0.23	10.07	0.14	
Population $_{CONTROL - SPRING}$	1.64	-2.80	6.89	0.08	
Population $_{CONTROL - HYBRID}$	3.67	-0.89	8.38	0.35	†
Generation	1.18	-0.77	3.10	0.80	

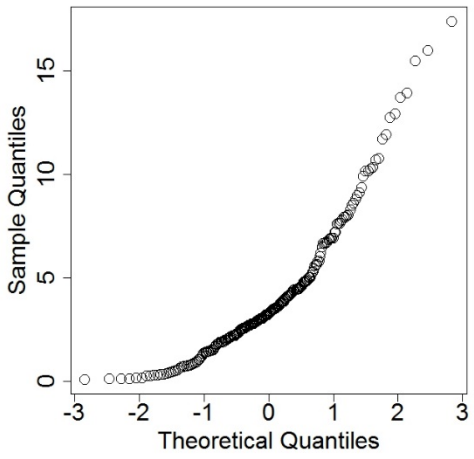
Appendix Table B. 9: The effect of the specification of different priors for random effects for biomass traits modelled using an ordinal distribution.

Priors tested were parameter expanded and inverse gamma priors, with V=1, R=fix retained for fixed effects.

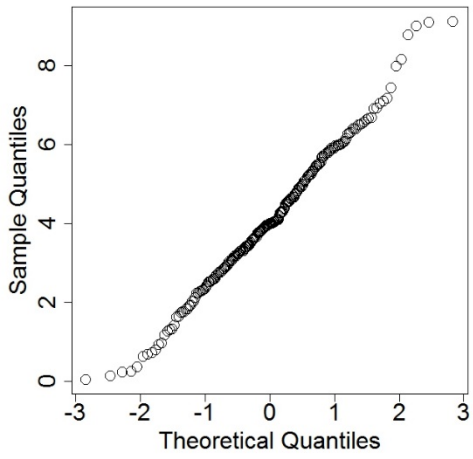
	Posterior mode	95% Lb CI	95% Ub CI
Number of rosettes			
Parameter expanded priors			
Heritability	0.09	0.00	0.31
Maternal effects	0.06	0.00	0.22
Paternal effects	0.13	0.00	0.32
Inverse gamma priors			
Heritability	0.05	0.00	0.22
Maternal effects	0.04	0.00	0.18
Paternal effects	0.10	0.00	0.30
Seeds per spike			
Parameter expanded priors			
Heritability	0.00	0.00	0.02
Maternal effects	0.67	0.47	0.88
Paternal effects	0.04	0.00	0.16
Maternal grandmaternal effects	0.22	0.00	0.40
Paternal grandpaternal effects	0.04	0.00	0.13
Inverse gamma priors			
Heritability	0.04	0.02	0.06
Maternal effects	0.73	0.48	0.98
Paternal effects	0.02	0.00	0.10
Maternal grandmaternal effects	0.20	0.00	0.47
Paternal grandpaternal effects	0.02	0.00	0.11
Seed weight			
Parameter expanded priors			
Heritability	0.00	0.00	0.02
Maternal effects	0.75	0.53	0.95
Paternal effects	0.12	0.00	0.31
Maternal grandmaternal effects	0.03	0.00	0.13
Paternal grandpaternal effects	0.07	0.00	0.22
Inverse gamma priors			
Heritability	0.00	0.00	0.00
Maternal effects	0.91	0.65	1.00
Paternal effects	0.05	0.00	0.27
Maternal grandmaternal effects	0.01	0.00	0.04
Paternal grandpaternal effects	0.04	0.00	0.24

Appendix Figure B. 2: Q-Q plots used to evaluate the assumption of normality for biomass traits.

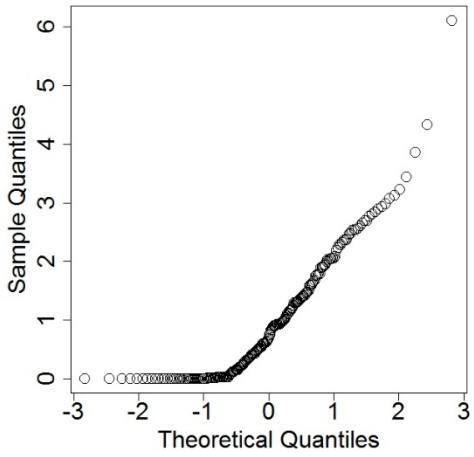
a) Root biomass Normal Q-Q plot



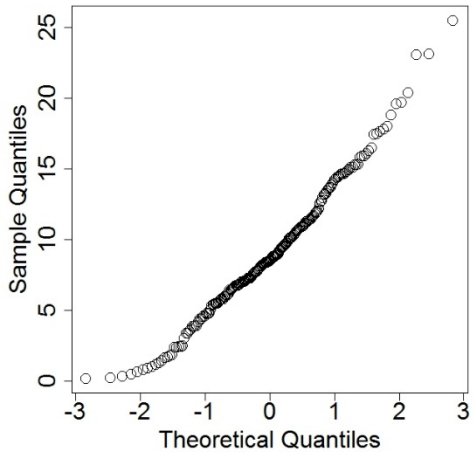
b) Leaf biomass Normal Q-Q plot



c) Reproductive biomass Normal Q-Q plot



d) Total biomass Normal Q-Q plot



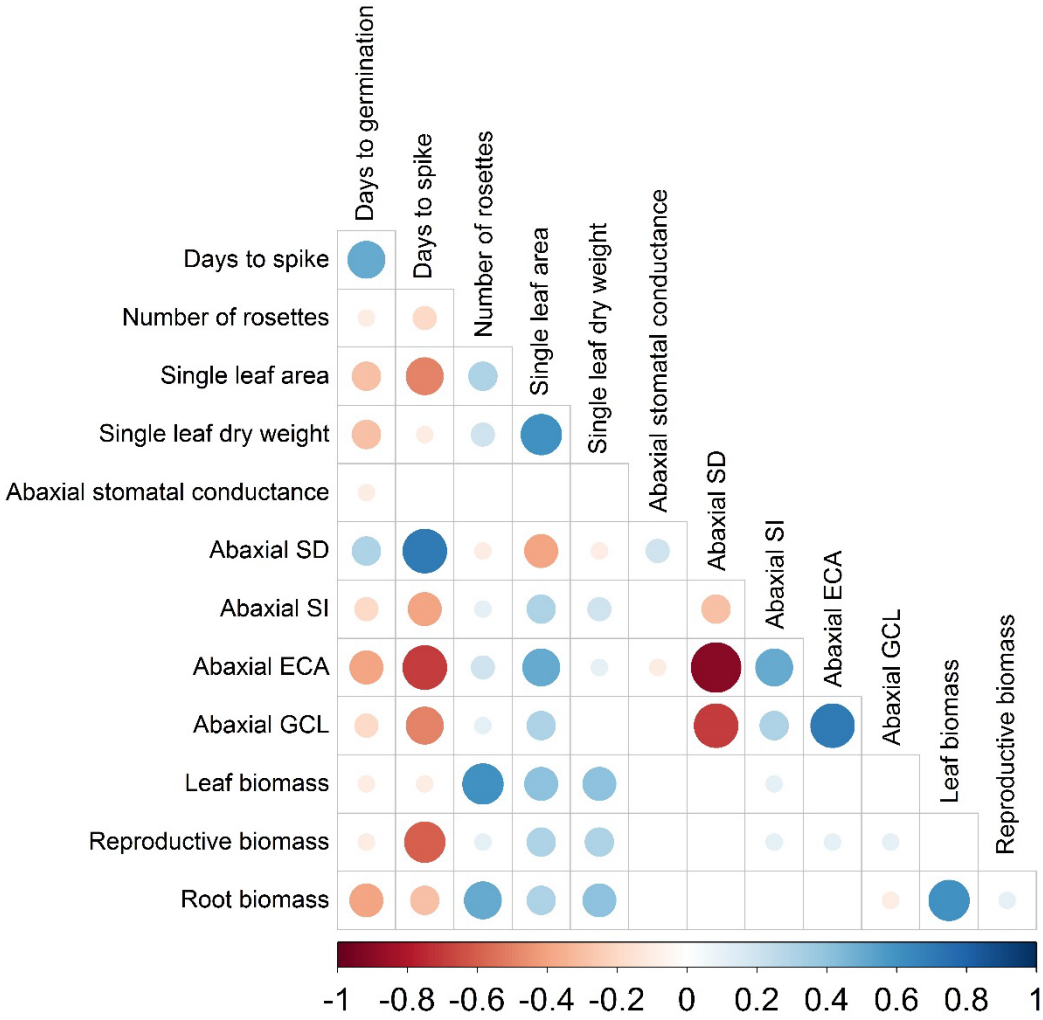
Appendix Table B. 10: Effect of [CO_2] treatment and ancestry of plants in generation 3 on whether or not plants reached the floral to vegetative transition before harvest.

Modelled using a glm with family=binomial so summary table gives values on the logit scale. P-value significance indicated as *** p<0.001, ** p<0.01, *p<0.01, † p<0.1

	Estimate	Standard Error	Z value	
Floral to vegetative transition				
Intercept	2.12	0.41	5.11	***
[CO_2] treatment $_{AMB - ELEV}$	0.12	0.38	0.32	
Population $_{CONTROL - SPRING}$	-0.83	0.47	-1.76	†
Population $_{CONTROL - HYBRID}$	-0.80	0.50	-1.58	

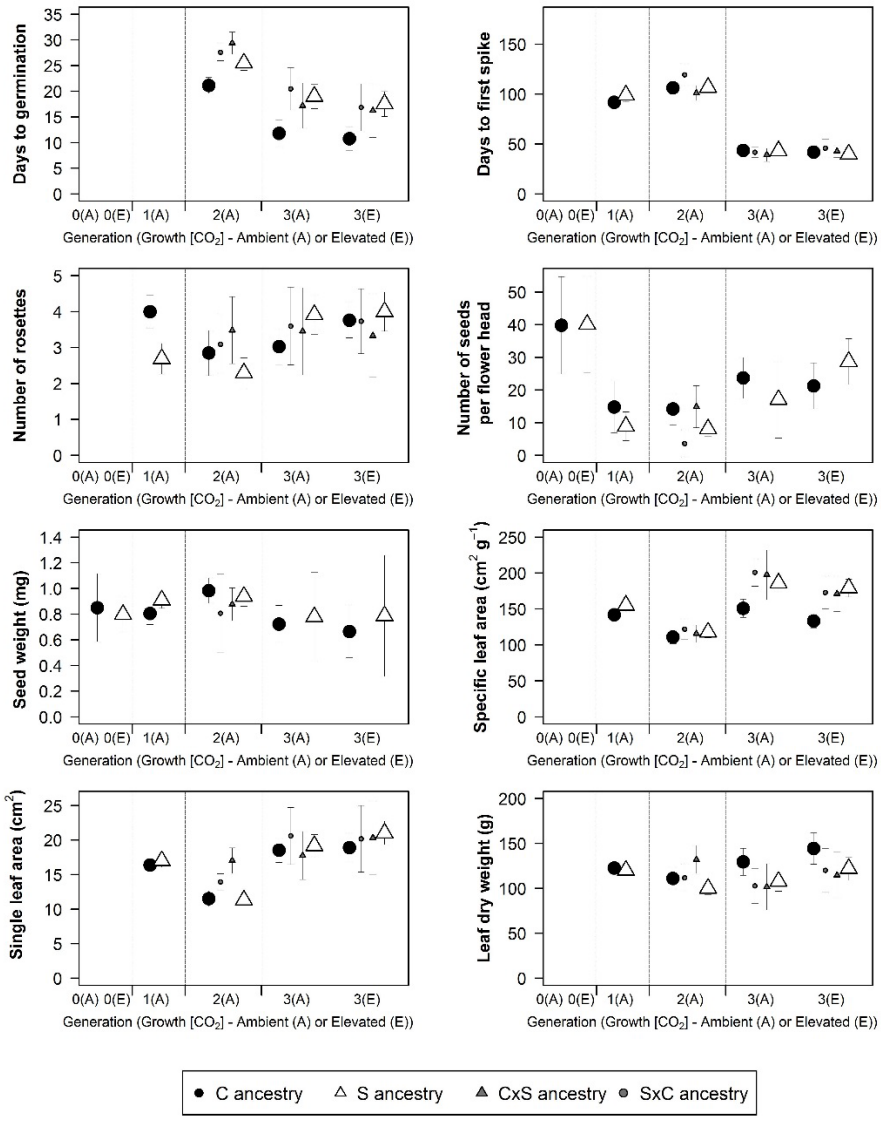
Appendix Figure B. 3: Pairwise correlation between traits in a three generation experiment with *P. lanceolata*.

The area of the circle and colour intensity indicates the Spearman's rank correlation coefficient between the two traits, with red signifying a positive correlation and blue signifying a negative correlation. Circles are present where the p value of the correlation was significant with p<0.01.



Appendix Figure B.4: Raw multigenerational reproductive and fitness traits of *Plantago lanceolata* in response to ambient and elevated growth [CO_2].

Response of *P. lanceolata* originating from *in situ* spring (G0, E) and control (G0, A) populations to elevated and ambient [CO_2]. Labels on the x-axis give generation number (growth [CO_2], A - ambient, E - elevated). Progeny of plants originating from the spring (S) or control (C) site population are identified by a white triangle or black circle respectively, with progeny resulting from spring-control (SxC or CxS) plant crosses identified by smaller grey points and with shape according to maternal parent. Trait responses are given as mean with 95 % confidence intervals.



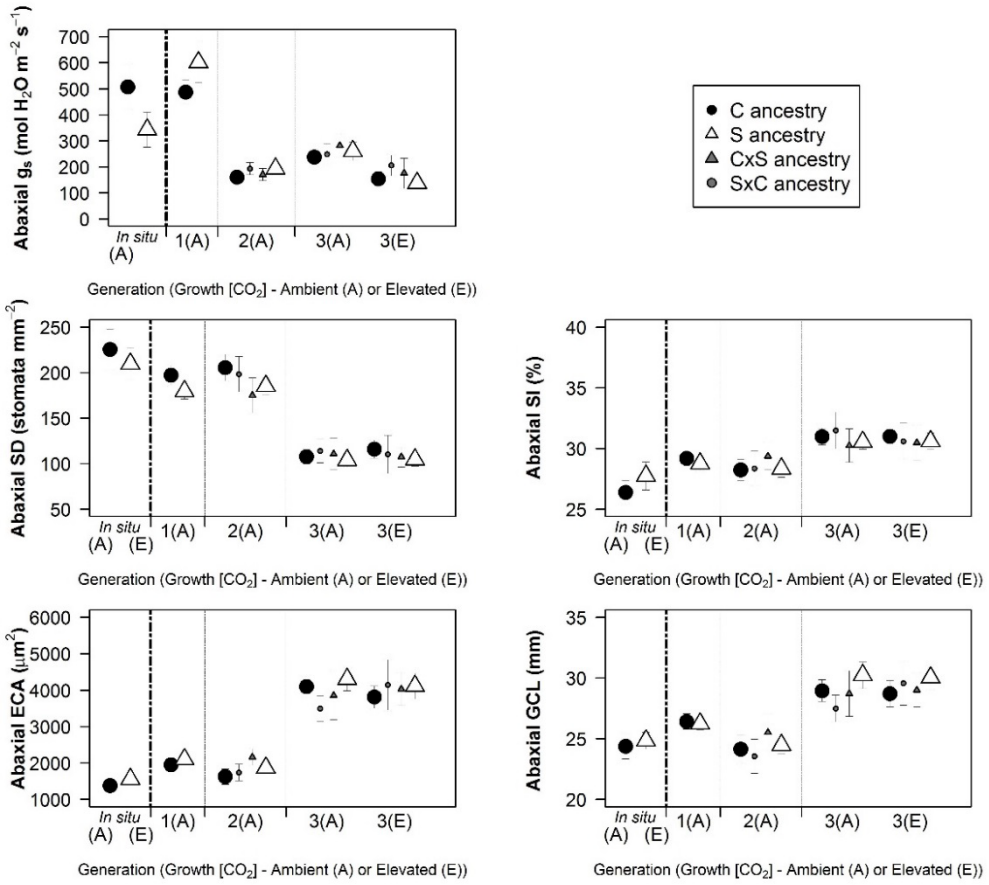
Appendix Table B.11: Effect estimates of [CO_2] treatment, population site of origin (Population), generational environment (Generation) and interaction effects on *Plantago lanceolata* traits.

Posterior mode of fixed effect variance estimates are given with 95 % Credible Interval lower and upper bounds in square brackets and P-value significance indicated as: *, P < 0.05; **, P < 0.01; ***, P < 0.001, †, P < 0.1. Interaction effects of [CO_2] treatment x ancestral location are included for models where they had a significant impact on variation. Models included random effects of the individual, mothers, fathers, (and maternal grandmother and paternal grandfather for two traits). For the two categorical moderator variables [CO_2] treatment and Population, contrasts are given in subscript. Abbreviations; A – ambient, E – elevated, C – control, S – spring, H - hybrid.

	Posterior mode [95 % CI]											
	Log days to germination	**Log days to first spike**	**Rosette number**	**Seeds per spike**	**Seed weight**	**Log single leaf area**	**Log single leaf dry weight**	**Log abaxial g_s**	**Log abaxial SD**	**Log abaxial SI**	**Log abaxial ECA**	**Log abaxial GCL**
Intercept	4.10 [3.5, 4.8] ***	4.94 [4.8, 5.1] ***	0.42 [-0.3, 1.1]	5.58 [1.7, 9.8] **	9.93 [3.7, 17.0] ***	2.42 [2.2, 2.6] ***	4.80 [4.6, 5.0] ***	6.14 [5.8, 6.5] ***	5.73 [5.6, 6.5] ***	3.32 [3.3, 3.4] ***	6.95 [6.7, 7.2] ***	3.13 [3.1, 3.2] ***
[CO_2] treatment $_{A-E}$	-0.11 [-0.3, 0.2] †	-0.08 [-0.2, 0.0] †	0.32 [-0.3, 0.9]	3.19 [0.0, 6.7] †	3.08 [-2.1, 8.3]	0.09 [0.0, 0.2] †	0.13 [0.1, 0.2] **	-0.49 [-0.6, -0.3] ***	-0.03 [0.0, 0.5]	0.01 [0.0, 0.0]	0.00 [-0.1, 0.1]	0.02 [0.0, 0.1]
Population $_{C-S}$	0.39 [-0.1, 0.8]	0.05 [-0.1, 0.2]	0.07 [-0.6, 0.8]	2.45 [-1.7, 7.1]	0.70 [-4.7, 5.6]	0.11 [-0.1, 0.3]	-0.13 [-0.3, 0.1]	0.05 [-0.2, 0.3]	-0.08 [-0.2, 0.0] †	0.02 [0.0, 0.1]	0.07 [-0.1, 0.2]	0.02 [0.0, 0.1]
Population $_{C–H}$	0.36 [-0.1, 0.8] †	-0.25 [-0.5, 0.0] *	-0.69 [-1.6, 0.2]	3.66 [-0.6, 7.8] †	-2.52 [-7.4, 2.5]	0.13 [0.0, 0.3]	-0.07 [-0.2, 0.1]	0.31 [0.1, 0.5] **	-0.05 [-0.2, 0.0]	0.01 [0.0, 0.1]	0.05 [-0.1, 0.2]	0.03 [0.0, 0.1]
Generation	-0.54 [-0.8, -0.3] ***	-0.44 [-0.5, -0.4] ***	0.60 [0.3, 0.9] ***	-0.99 [-1.9, -0.2] *	-2.60 [-4.7, -0.3] *	0.12 [0.0, 0.2] **	-0.04 [-0.1, 0.0]	-0.29 [-0.4, -0.2] ***	-0.34 [-0.4, -0.3] ***	0.03 [0.0, 0.1] **	0.42 [0.3, 0.5] ***	0.07 [0.0, 0.1] ***
[CO_2] treatment $_{A-E}$ x Population $_{C-S}$		-0.4 [-0.2, 0.1]	-0.88 [-1.7, 0.0] *	-4.50 [-10.0, 1.4]								
[CO_2] treatment $_{A-E}$ x Population $_{C-H}$		0.16 [0.0, 0.3] *	-0.6 [-1.0, 1.0]	-7.39 [-12.3, -2.6] ***								

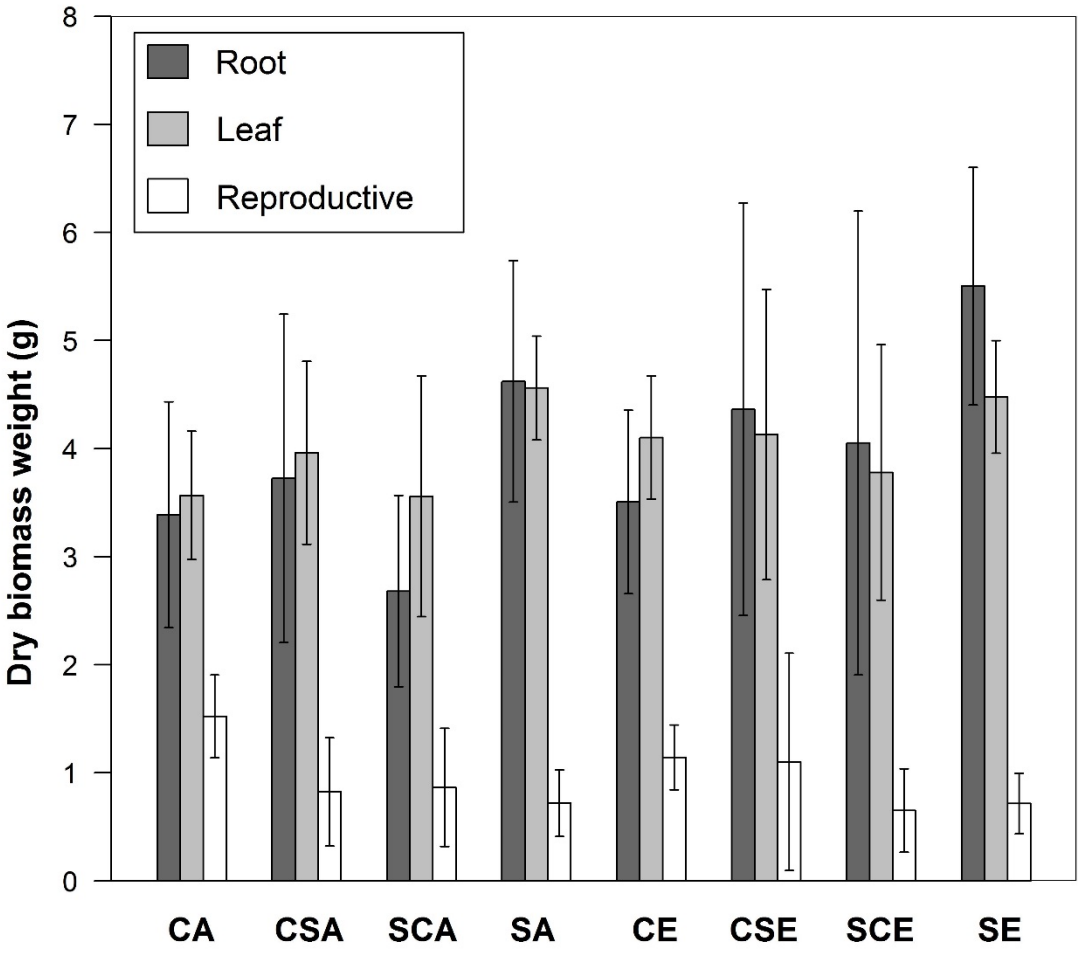
Appendix Figure B.5: Raw multigenerational abaxial stomatal patterning traits of *Plantago lanceolata* in response to ambient and elevated growth [CO_2].

Response of *P. lanceolata* originating from *in situ* spring (*in situ*, E) and control (*in situ*, A) populations to elevated and ambient [CO_2]. Labels on the x-axis give generation number (growth [CO_2], A - ambient, E - elevated). Progeny of plants originating from the spring (S) or control (C) site population are identified by a white triangle or black circle respectively, with progeny resulting from spring-control (SxC or CxS) plant crosses identified by smaller grey points and with shape according to maternal parent. Trait responses are given as mean with 95 % confidence intervals. Note that the first generation of responses (labelled as *in situ* rather than generation zero) represent the *in situ* stomatal patterning response but were recorded two years after the seed for this experiment were collected, and have not contributed to the multigenerational analysis of these traits.



Appendix Figure B. 6: Mean biomass components of *P. lanceolata* originating from spring and control site population (or hybrids of these ancestries) when grown in ambient or elevated [CO_2] following two generations of growth at ambient [CO_2].

The eight treatment groups are CA - control ambient, CSA – control-spring hybrid ambient, SCA - spring-control hybrid ambient, SA – spring ambient, CE – control elevated, CSE – control-spring elevated, SCE – spring-control elevated, SE – spring elevated. For hybrids the first letter gives the maternal ancestry with second letter giving paternal ancestry. Bars indicate mean biomass weight and error bars are 95% confidence intervals.



Appendix C **Chapter 4 Supplementary Information**

Appendix C.1 | Supplementary Tables and Figures

Appendix Figure C.1 | Frequency distribution of relative sequence divergence (F_{ST} between spring and control populations) for all SNPs identified prior to filtering out singleton SNPs and filtering by eight individuals per population.

Appendix Figure C.2 | Density distribution of annotation values informing SNP discovery.

Appendix Table C.1 | Assembly statistics of short read assembled genomes using kmer size of 105 and 110.

Appendix Table C.2 | Impact of filtering genome to > 2 Kb contigs.

Appendix Figure C.3 | The size distribution of assembled genome contigs.

Appendix Figure C.4 | Genome completeness following size filtration.

Appendix Figure C.5 | Chloroplast genome assembly.

Appendix Table C.3 | Conversion efficiency rate for each sample as determined by the number of cytosines retained in sequences mapping to the chloroplast following bisulfite conversion.

Appendix Figure C.6 | Transcriptome completeness.

Appendix Table C.4 | The percentage of DMS in each methylation sequence context that were also the location of a C-T SNP.

Appendix Figure C.7 | Difference in expression and percentage methylation across predicted gene features between plants originating from the spring vs control population, as well as the methylation of regions of the genome to which expressed transcripts map and expression of those transcripts.

Appendix Table C.5 | Linear models of difference in expression and methylation across predicted gene features.

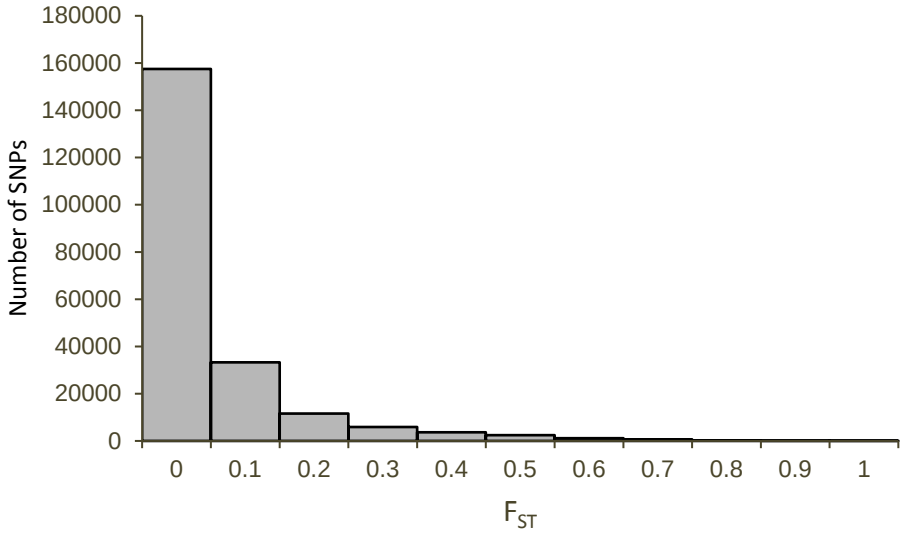
Appendix Table C.6 | Description of differentially expressed (DE) transcripts that mapped to within 1 kb of a differentially methylated site (DMS) or a single nucleotide polymorphism (SNP).

Appendix Figure C.8 | Features of DMSs with DE transcripts mapping to within 500 bases (with one exception, which had multiple transcripts mapping to the same site, see Appendix Figure C. 9).

Appendix Figure C.9 | Transcripts encoding LRR and NB-ARC domains-containing disease resistance protein mapped to two locations in the genome (start position on contig given), were differentially expressed and located within 500 bp of an outlier FST SNP.

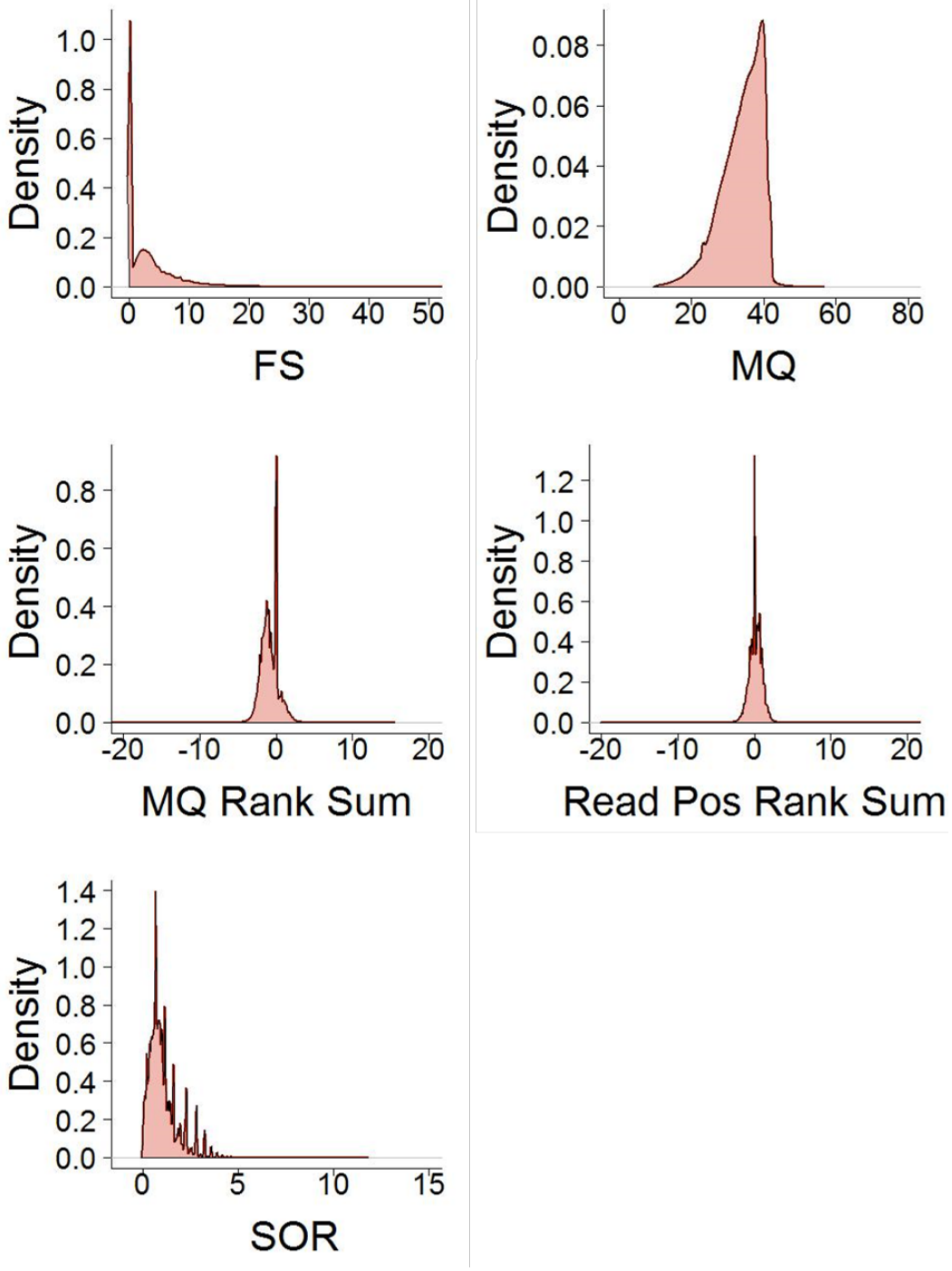
C.1 Supplementary Tables and Figures

Appendix Figure C. 1: Frequency distribution of relative sequence divergence (F_{ST} between spring and control populations) for all SNPs identified prior to filtering out singleton SNPs and filtering by eight individuals per population.



Appendix Figure C. 2: Density distribution of annotation values informing SNP discovery.

Annotations are; the Phred scaled probablility of strand bias at the site (FisherStrand FS), the root mean square mapping quality over all the reads at the site (RMSMapping Quality, MQ), the u-based z-approximation from the rank sum test for mapping qualities (MappingQualityRankSumTest, MQRankSum), the u-based z-approximation from the rank sum test for site position within reads (ReadPosRankSum), and the strand offs ratio to estimate strand bias (StrandOddsRatio, SOR)



Appendix Table C. 1: Assembly statistics of short read assembled genomes using kmer size of 105 and 110.

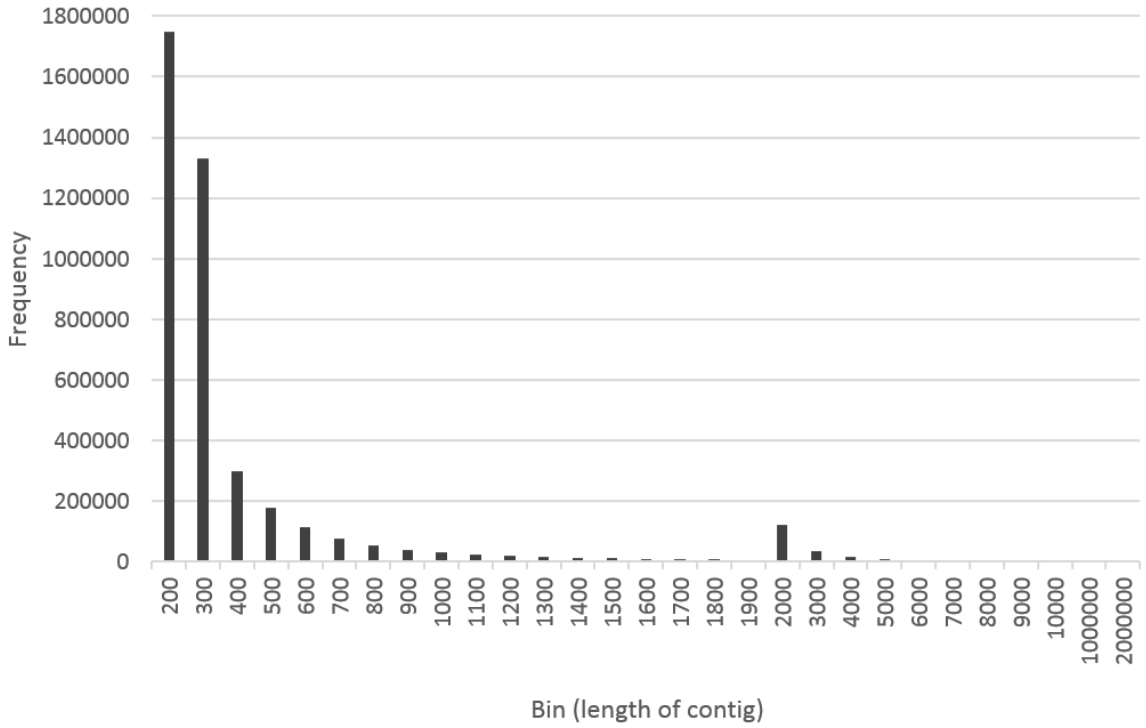
Genomes were assembled using ABySS (Simpson *et al*., 2009) and SOAPdenovo2 (Luo *et al*., 2012).

Assembler	kmer	# contigs	Largest contig (Kb)	Main genome contig sequence total (Kb)	GC (%)	N50 (b)
Abyss	105	4,075,744	1321	1,425,357	39.88	1855
	110	3,847,997	1321	1,433,684	39.88	1811
SOAPdenovo2	105	12,996,685	777	2,389,772	40.26	501
	110	12,058,795	928	2,357,958	40.29	477

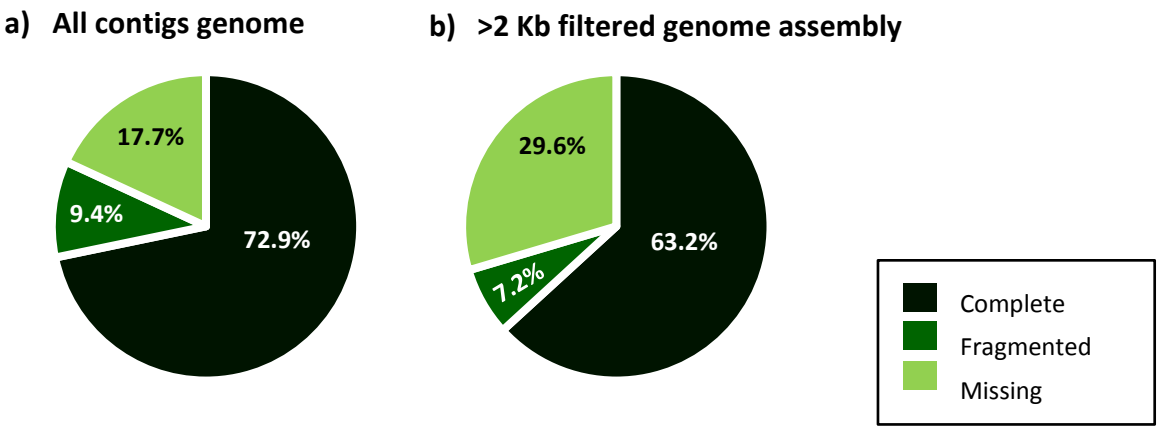
Appendix Table C. 2: Impact of filtering genome to > 2 Kb contigs.

Genome	# contigs	Main genome contig sequence total (kb)	Percentage # contigs of full genome	Percentage of full genome contig sequence total
***P. lanceolata* full genome**	4,075,744	1,425,357	100%	100%
***P. lanceolata* >2 Kb filtered genome**	29,103	195,204	0.7%	13.7%

Appendix Figure C. 3: The size distribution of assembled genome contigs.

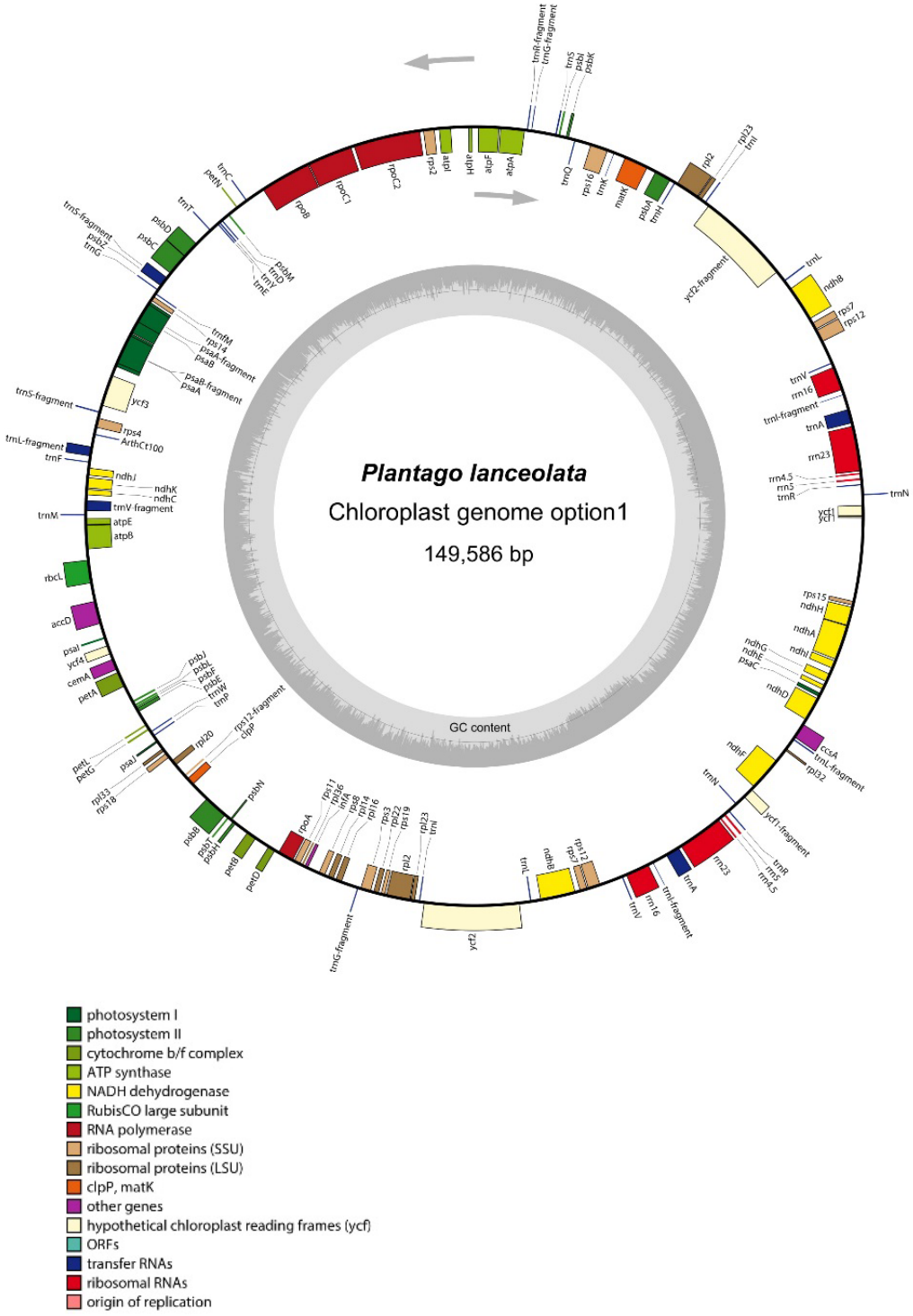


Appendix Figure C. 4: Genome completeness following size filtration.
Percentage of 1440 single copy orthologues from Arabidopsis thaliana that are complete (dark green), fragmented (mid green) or missing (light green) as quantified by BUSCO analysis in a) the all contigs genome and b) the >2 Kb contig filtered genome.



Appendix Figure C. 5: Chloroplast genome assembly.

Chloroplast genome assembled using Novoplasty and visualised with OGDRAW. Two options were resolved based on the orientation of the inverted repeat with this option most likely based on phylogeny.

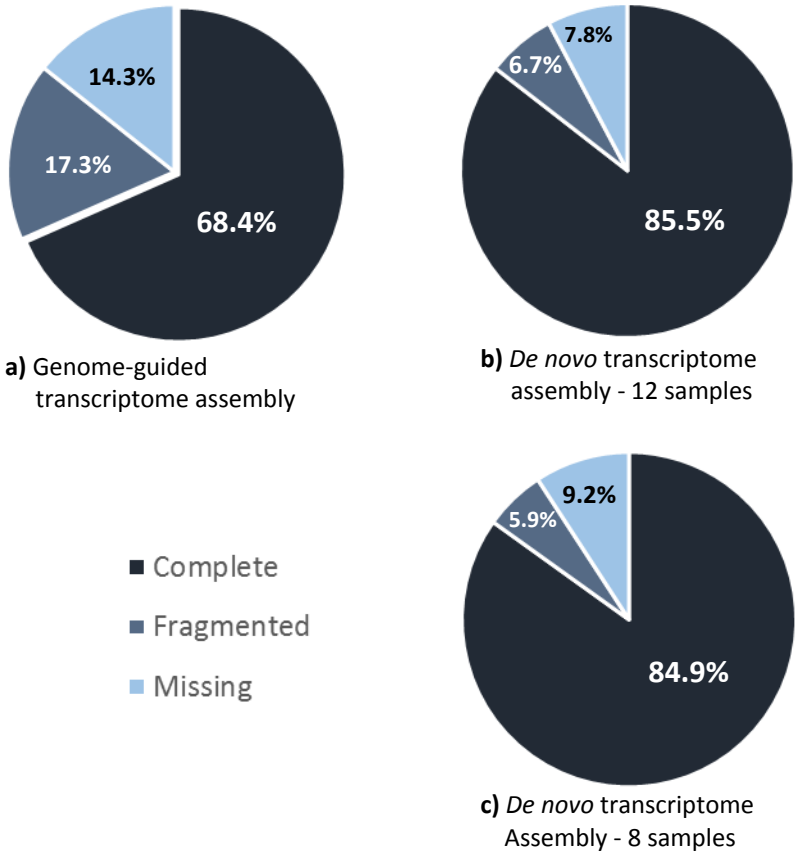


Appendix Table C. 3: Conversion efficiency rate for each sample as determined by the number of cytosines retained in sequences mapping to the chloroplast following bisulfite conversion.

Percentage of reads mapping to the chloroplast is given as mapping efficiency.

Sample	Percentage mapping efficiency	Percentage conversion efficiency error rate	Percentage of C's correctly converted
CA3-3	1.30	1.32	98.68
CA3-8	2.00	1.31	98.69
CA6-7	1.90	1.26	98.74
CA6-14	1.60	1.46	98.54
CA9-3	1.70	1.28	98.72
CA9-6	1.80	1.28	98.72
CE3-1	1.80	1.35	98.65
CE3-4	1.50	1.48	98.52
CE6-5	2.10	1.13	98.87
CE6-12	1.50	1.29	98.71
CE9-2	1.60	1.39	98.61
CE9-8	1.90	1.34	98.66
SA13-1	2.50	1.29	98.71
SA13-3	2.10	1.45	98.55
SA15-3	2.20	1.35	98.65
SA15-7	1.20	1.44	98.56
SA19-1	2.50	1.32	98.68
SA19-6	2.10	1.28	98.72
SE13-4	1.70	1.38	98.62
SE13-5	1.90	1.19	98.81
SE15-11	2.20	1.18	98.82
SE15-12	1.80	1.27	98.73
SE19-2	1.60	1.35	98.65
SE19-5	2.10	1.20	98.80

Appendix Figure C. 6: Transcriptome completeness.
Percentage of 1440 single copy orthologues from *Arabidopsis thaliana* that are complete (dark blue), fragmented (mid blue) or missing (light blue) in three assembled transcriptomes **a)** the genome guided trinity assembled transcriptome **b)** the *de novo* trinity assembled transcriptome using 12 samples, **c)** the *de novo* trinity assembled transcriptome using 8 samples (published in Watson-Lazowski *et al*., 2016)



a) Genome-guided transcriptome assembly

b) *De novo* transcriptome assembly - 12 samples

c) *De novo* transcriptome Assembly - 8 samples

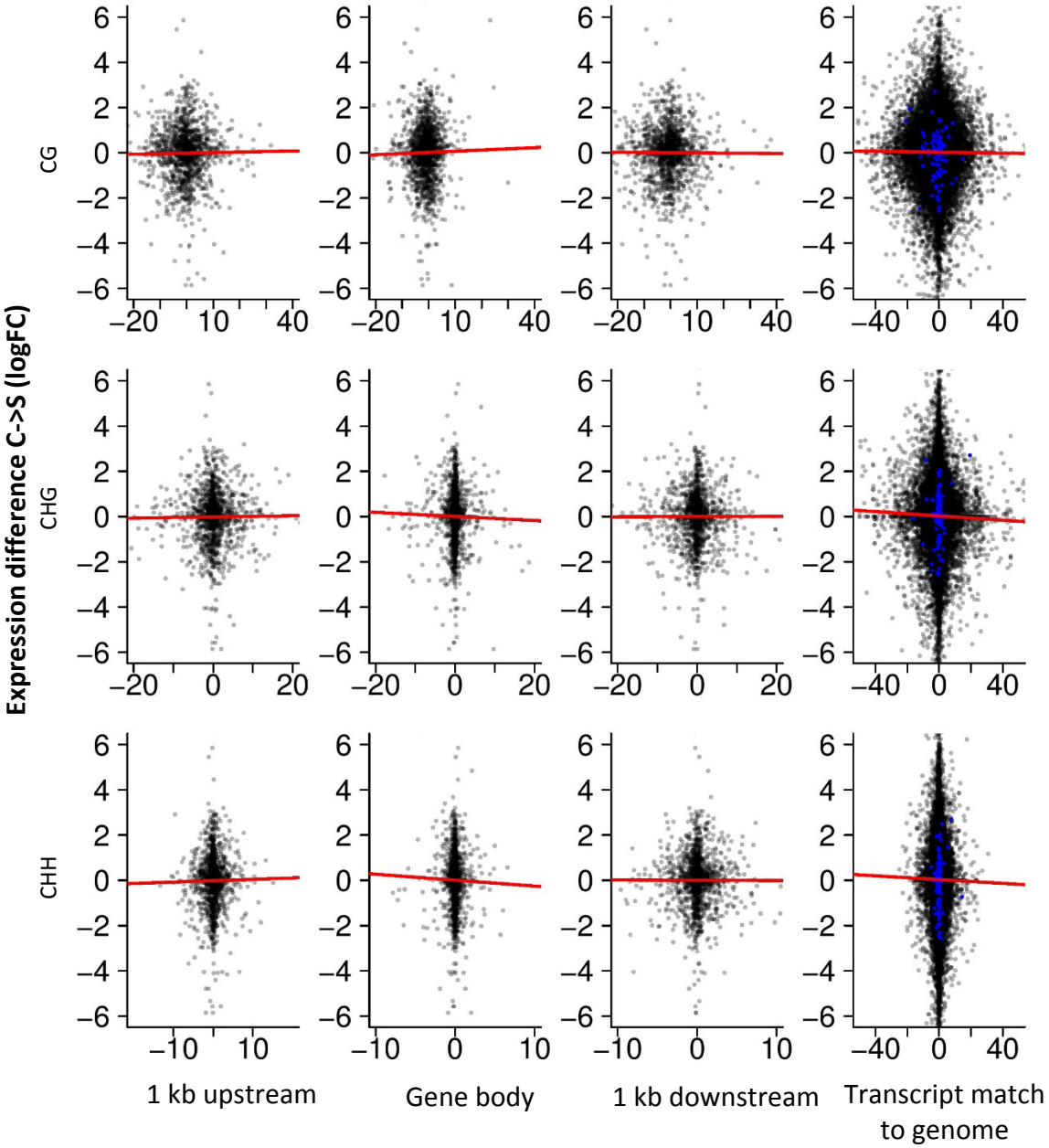
Appendix Table C. 4: The percentage of DMS in each methylation sequence context that were also the location of a C-T SNP.

These could result in the incorrect calling of a DMS if not aligned to individual reference genomes, or removed from the analysis

	Site of origin	Growth [CO_2]	Site of origin x Growth [CO_2]
CG	15.7%	14.9%	13.6%
CHG	11.1%	13.0%	13.4%
CHH	3.3%	2.1%	3.1%

Appendix Figure C. 7: Difference in expression and percentage methylation across predicted gene features between plants originating from the spring vs control population, as well as the methylation of regions of the genome to which expressed transcripts map and expression of those transcripts.

Linear models are given as red lines and differentially expressed transcripts are plotted with blue dots in the fourth column



Appendix Table C. 5: Linear models of difference in expression and methylation across predicted gene features.

F-statistics, R^2 values and p-values for simple linear regressions are given for Difference in expression and percentage methylation across predicted gene features between plants originating from the spring vs control population, as well as the methylation of regions of the genome to which expressed transcripts map and expression of those transcripts

Context	Position	Coefficient estimate	F-statistic (df)	R^2	p-val
CG	1 kb upstream	0.0023	0.23 (1, 1431)	0.00162	(NS)
	Gene body	0.0052	0.50 (1, 1732)	0.00029	(NS)
	1 kb downstream	-0.0009	0.04 (1, 1491)	<0.00001	(NS)
	Across transcript	-0.0011	1.31 (1, 33736)	<0.00001	(NS)
CHG	1 kb upstream	0.0026	0.16 (1, 1455)	0.00011	(NS)
	Gene body	-0.0091	1.00 (1, 1739)	0.00058	(NS)
	1 kb downstream	0.0006	0.01 (1, 1513)	<0.00001	(NS)
	Across transcript	-0.0047	14.34 (1, 39553)	0.00036	(***)
CHH	1 kb upstream	0.0061	0.20 (1, 1496)	0.00014	(NS)
	Gene body	-0.0263	0.74 (1, 1745)	0.00042	(NS)
	1 kb downstream	-0.0019	0.01 (1, 1554)	<0.00001	(NS)
	Across transcript	-0.0042	1.96 (1, 44397)	<0.00001	(NS)

Appendix Table C. 6: Description of differentially expressed (DE) transcripts that mapped to within 1 kb of a differentially methylated site (DMS) or a single nucleotide polymorphism (SNP).

Where DE effect is Site (site of origin) log fold change (logFC) associated with the transcript is given for spring relative to control, where DE effect is Growth [CO_2], logFC is given as elevated relative to ambient.

Transcript name	Description	GO biological process	DMS (context, effect) or SNP	DMS or SNP location relative to transcript	Percentage methylation (Adjusted Wald)				Hyper- or hypo-methylated	DE effect	logFC associated with DE
					CA	CE	SA	SE			
TRINITY_DN32818_c1_g1_i1	Sterile alpha motif domain-containing protein AT3G07760.2	GO:0000004; GO:0007582; physiological process; biological process; single-organism process	DMS (CHG, Site)	933 bp downstream	16.7	33.3	33.3	80.0	Hyper	Site	8.6
TRINITY_DN44997_c0_g8_i1	subtilase family protein (SBTI1.1) AT1G01900.1	GO:0006508; proteolysis	DMS (CG, Site)	175 bp into transcript	94.7	83.3	84.8	97.5	Hypo	Site	4.7
TRINITY_DN42769_c0_g4_i1	ferredoxin-NADP[+]-oxidoreductase 2 (FNR2) AT1G20020.2	GO:0042830; defense response to bacteria GO:0009817; defense response to fungus GO:0055114; oxidation-reduction process GO:0009735; response to cytokinin GO:0015979; photosynthesis	DMS (CG, Site)	220 bp upstream	90.0	90.0	74.0	54.0	Hypo	Site	10.1
TRINITY_DN32740_c0_g1_i1	cleavage and polyadenylation specificity factor	GO:0006378~mRNA polyadenylation GO:0006397~mRNA processing GO:0005847~mRNA cleavage	DMS (CHG, Site)	167 bp downstream	90.0	78.1	75.0	30.8	Hypo	Site	4.4
			DMS (CHH, Site)	240 bp downstream	28.6	43.6	12.6	5.4	Hypo	Site	

	73-I AT1G61010.3	and polyadenylation specificity factor complex	DMS (CHG, Site)	453 bp downstream	87.5	77.8	81.8	33.0	-	Site	
TRINITY_DN37509_c0_g2_i2	LRR and NB-ARC domains-containing disease resistance protein AT1G61190.1	GO:0006952~defense response	SNP	288 bp upstream						Site	5.7
TRINITY_DN37509_c0_g1_i2			SNP	442 bp upstream						Site	6.9
TRINITY_DN37509_c0_g1_i3			SNP	442 bp upstream						Site	-5.9
TRINITY_DN37509_c0_g2_i4			SNP	288 bp upstream						Growth [CO_2]	5.6
TRINITY_DN44489_c3_g1_i3	Sec14p-like phosphatidylinositol transfer family protein AT1G75170.2	GO:0006810~transport	DMS (CG, Site)	985 bp upstream	58.8	16.7	26.3	9.5	Hypo	Growth [CO_2]	5.6
TRINITY_DN44489_c3_g1_i1	Erythronate-4-phosphate dehydrogenase family protein AT1G75180.3	Unknown	DMS (CG, Site)	782 bp upstream	77.3	37.5	65.4	26.3	Hypo	Growth [CO_2]	5.6
			DMS (CG, Site)	649 bp upstream	58.8	16.7	26.3	9.5	Hypo	Growth [CO_2]	5.6
TRINITY_DN43075_c1_g3_i1	Tetratricopeptide repeat (TPR)-like superfamily protein AT2G20710.1	Unknown	DMS (CG, Site)	339 bp downstream	13.0	26.7	71.4	73.8	Hyper	Growth [CO_2]	-5.3
TRINITY_DN22610_c0_g1_i1	hypothetical protein (XBAT31) AT2G28840.2	GO:0016567~protein ubiquitination	SNP	988 bp upstream						Site	4.6

TRINITY_DN34644_c0_g1_i3	Cysteine proteinases superfamily protein AT2G38025.1	GO:0016579~protein deubiquitination GO:0030433~ER-associated ubiquitin-dependent protein catabolic process GO:0030968~endoplasmic reticulum unfolded protein response	SNP	754 bp downstream						Site	-6.7
			SNP	798 bp downstream						Site	
			SNP	805 bp downstream						Site	
			SNP	819 downstream						Site	
TRINITY_DN36297_c1_g2_i1	Ribosomal protein S11 family protein AT3G52580.1	GO:0000028~ribosomal small subunit assembly GO:0000462~maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA) GO:0006412~translation	DMS (CHG, Site)	816 bp downstream	81.1	88.0	41.2	20.8	Hypo	Site	-9.4
TRINITY_DN36297_c1_g3_i1			DMS (CHG, Site)	816 bp downstream	81.1	88.0	41.2	20.8	Hypo	Growth [CO_2]	9.3
TRINITY_DN36297_c1_g7_i1			DMS (CHG, Site)	816 bp downstream	81.1	88.0	41.2	20.8	Hypo	Growth [CO_2]	-7.5
TRINITY_DN37959_c1_g2_i1	Phenazine biosynthesis PhzC/PhzF protein AT4G02860.1	GO:0009058~biosynthetic process	DMS (CG, Site)	910 bp upstream	12.5	11.1	36.4	64.7	Hyper	Site	3.7
TRINITY_DN37959_c1_g2_i2			DMS (CG, Site)	910 bp upstream					Hyper	Site	4.4
TRINITY_DN43852_c1_g2_i1	Cytochrome c oxidase biogenesis protein Cmc1-like protein AT4G21192.2	Unknown	DMS (CHG, Site)	2 bp upstream	93.5	76.0	50.0	27.8	Hypo	Site	-4.7
TRINITY_DN44249_c3_g5_i11	PPPDE thiol peptidase family protein AT4G31980.1	Unknown	DMS (CG, Site)	315 bp upstream	81.8	75.0	21.4	25.0	Hypo	Growth [CO_2]	7.5
			DMS (CG, Site)	302 bp upstream	81.8	66.7	42.9	26.7	Hypo	Growth [CO_2]	
			DMS (CG, Site)	725 bp upstream	13.0	26.9	71.4	81.8	Hyper	Growth [CO_2]	

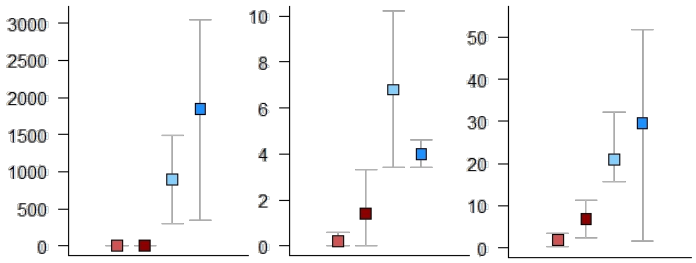
TRINITY_DN16809_c0_g1_i1	NAD(P)-binding Rossmann-fold superfamily protein (FLDH) AT4G33360.2	GO:0006721~terpenoid metabolic process GO:0009788~negative regulation of abscisic acid-activated signaling pathway GO:0016487~farnesol metabolic process	SNP	771 bp upstream						Site	6.4
			SNP	759 bp upstream						Site	
			DMS (CHH, Site)	564 bp upstream	77.3	58.3	52.0	19.2	Hypo	Site	
			DMS (CG, Site)	533 bp upstream	92.7	80.0	48.2	23.5	Hypo	Site	
			SNP	525 bp upstream						Site	
			SNP	170 bp upstream						Site	
			DMS (CHG, Site)	64 bp upstream	81.8	50.0	31.3	23.5	Hypo	Site	
			DMS (CHG, Site)	42 bp upstream	82.8	69.2	25.0	21.1	Hypo	Site	
			DMS (CHG, Site)	53 bp downstream	83.8	85.0	73.3	22.2	Hypo	Site	
			DMS (CHG, Site)	67 bp downstream	84.8	81.8	70.6	24.1	Hypo	Site	
			SNP	238 bp downstream						Site	
			SNP	386 bp downstream						Site	
			SNP	471 bp downstream						Site	
			SNP	510 bp downstream						Site	

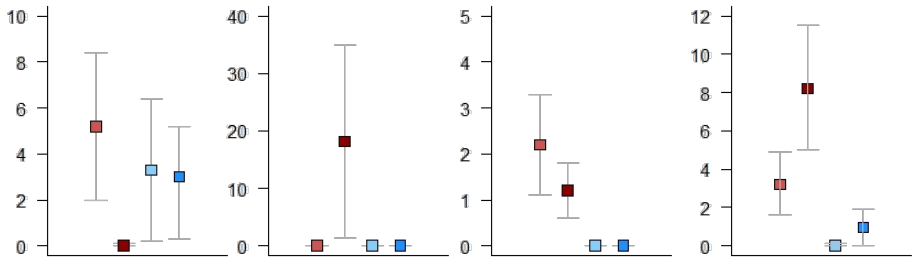
TRINITY_DN33115_c0_g2_i3	alpha-1,3-mannosyl-glycoprotein beta-1,2-N-acetylglucosaminyltransferase (GTN1) AT4G38240.3	GO:0006486~protein glycosylation GO:0006491~N-glycan processing GO:0006972~hyperosmotic response	DMS (CG, Site)	344 bp upstream	34.9	45.2	76.0	89.5	Hyper	Site	-4.1
TRINITY_DN35833_c0_g1_i2	Heavy metal transport/detoxification superfamily protein AT5G14910.1	GO:0009735~response to cytokinin GO:0030001~metal ion transport GO:0009507~chloroplast GO:0009535~chloroplast thylakoid membrane	DMS (CG, Site)	149 bp upstream	22.7	87.0	90.0	86.4	Hyper	Site	3.6
			DMS (CHG, Site)	29 bp downstream	5.6	19.0	26.2	50.0	Hyper		
TRINITY_DN42230_c0_g1_i2	TRYPTOPHAN BIOSYNTHESIS 1 (TRP1) AT5G17990.1	GO:0000162~tryptophan biosynthetic process GO:0008652~cellular amino acid biosynthetic process	DMS (CG, Site)	708 bp downstream	4.2	6.3	45.7	40.0	Hyper	Growth $[CO_2]$	-5.9
TRINITY_DN45278_c0_g1_i2	BURP domain-containing protein (RD22) AT5G25610.1	GO:0009269~response to desiccation GO:0009651~response to salt stress GO:0009737~response to abscisic acid	DMS (CG, Site)	73 bp downstream	12.5	25.0	60.0	61.5	Hyper	Site	5.8
			DMS (CG, Site)	74 bp downstream	13.2	20.5	54.8	55.9	Hyper		
			DMS (CG, Site)	84 bp downstream	12.8	30.2	65.7	66.7	Hyper		
			DMS (CG, Site)	91 bp downstream	7.5	34.1	68.4	48.3	Hyper		
			DMS (CG, Site)	101 bp downstream	5.6	33.3	61.0	53.3	Hyper		
TRINITY_DN43835_c0_g3_i2	Disease resistance protein (CC-NBS-LRR class) family AT5G48620.1	GO:0006952~defense response	DMS (CG, Site)	814 upstream	16.7	41.7	54.5	81.8	Hyper	Site	6.1
			DMS (CG, Site)	240 upstream	8.0	13.3	52.9	70.0	Hyper		

	DMS (CHH, Site)	186 bp upstream	88.9	52.6	18.2	21.4	Hypo

Appendix Figure C. 8: Features of DMSs with DE transcripts mapping to within 500 bases (with one exception, which had multiple transcripts mapping to the same site, see Appendix Figure C. 9).

TMM-normalized expression of the four treatment groups based on site of origin (spring or control) and growth [CO_2] (ambient or elevated) is shown. All DMS in this category were differentially methylated by site of origin (denoted DMS_{SITE}). DMS are described as hyper- or hypo-methylated in plants originating from the spring plant population relative to the control population.

	Subtilase family protein	Cleavage and polyadenylation specificity factor	Heavy metal transport protein
Transcript TMM-normalized expression			
DMS_{SITE} context	CG	2 x CHG 1 x CG	CHG
DMS_{SITE} position	175 bp into the transcript	167, 453 and 240 bp upstream	29 bp downstream
Hyper/hypo-methylated	Hypermethylated	Hypomethylated	Hypermethylated

	TPR-like protein	PPPDE thiol peptidase protein	GNTI protein	CMC1-like protein
Transcript TMM-normalized expression				
DMS_{SITE} context	CG	3 x CG	CG	CHG
DMS_{SITE} position	339 bp downstream	315, 305 and 725 bp downstream	344 bp upstream	2 bp upstream
Hyper/hypo-methylated	Hypermethylated	Hyper-, hyper and hypomethylated respectively	Hypermethylated	Hypomethylated

Appendix Figure C. 9: Transcripts encoding LRR and NB-ARC domains-containing disease resistance protein mapped to two locations in the genome (start position on contig given), were differentially expressed and located within 500 bp of an outlier F_{ST} SNP.