UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCES

Centre of Biological Sciences

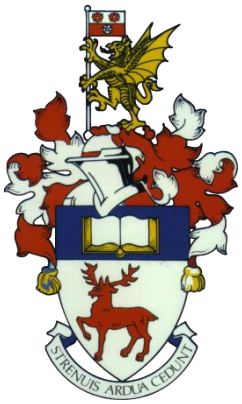
# Evidence for Plant Adaptation to a Future High $CO_2$ World

by

Alex Watson-Lazowski

Supervised by

Professor Gail Taylor, Dr Mark Chapman and Dr Richard Edwards




Thesis for the degree of Doctor of Philosophy

May 2015

# Abstract

FACULTY OF NATURAL AND ENVIRONEMNTAL SCIENCES

Doctor of Philosophy

**EVIDENCE FOR PLANT ADAPTATION TO A FUTURE HIGH $CO_2$ WORLD**

By Alex Watson-Lazowski

Plant morphology and function are sensitive to rising atmospheric carbon dioxide ($CO_2$) concentrations, but evidence that $CO_2$ concentration can act as a selective pressure driving evolution is sparse. Plants originating from naturally high $CO_2$ springs are subjected to elevated $CO_2$ concentration over multiple generations, providing an opportunity to predict how adaptation to future atmospheres may occur, with important implications for future plant conservation and crop breeding strategies. Using *Plantago lanceolata* L. from such a site (the 'spring' site) and from an adjacent ambient $CO_2$ site ('control' site), and growing the populations in ambient and elevated $CO_2$ at 700 μmol $mol^{-1}$, I have characterised, for the first time, the functional and population genomics, alongside morphology and physiology, of plant adaptation to elevated $CO_2$ concentrations. Growing plants in elevated $CO_2$ caused relatively modest changes in gene expression, with fewer changes evident in the spring than control plants (33 vs 131 genes differentially expressed [DE], in spring and control plants respectively). In contrast, when comparisons were made between control and spring plants grown in either ambient or elevated $CO_2$, there were a much larger number of loci showing DE (689 in the ambient and 853 in the elevated $CO_2$ environment). Population genomic analysis revealed that genetic differentiation between the spring and control plants was close to zero with no fixed differences, suggesting that plants are adapted to their native $CO_2$ environment at the level of gene expression. Growth at elevated $CO_2$ led to an unusual phenotype, with an increase in stomatal density and index in the spring, but not in control plants. Focussing on previously characterised stomatal patterning genes revealed significant DE (FDR < 0.05) between spring and control plants for three loci (*YODA*, *CDKB1;1*, and *SCRM2*) and between ambient and elevated $CO_2$ for four (*ER*, *YODA*, *MYB88*, and *BCA1*). We propose that the up-regulation in spring plants of two positive regulators of stomatal numbers (*SCRM2* and *CDKB1;1*) act here as key controllers of stomatal adaptation to elevated $CO_2$ on an evolutionary timescale. Significant transcriptome reprogramming of the photosynthetic pathway was identified, with an overall decrease in expression across the pathway in control plants, and an increase in spring plants, in response to elevated $CO_2$. This was followed up by physiological measurements, where a significant increase ($P < 0.05$) in photosynthetic capacity and regeneration rate was exhibited in spring plants, compared to control plants, at both elevated and ambient $CO_2$ concentrations. Through this comprehensive analysis, we have identified the basis of plant adaptation to elevated $CO_2$ likely to occur in the future.

# Table of Contents

# List of Figures

**Chapter 1**



**Chapter 2**



**Chapter 3**

**Chapter 4**



**Chapter 5**



**Chapter 6**

# List of Tables

**Chapter 1**



**Chapter 2**



**Chapter 3**



**Chapter 4**



**Chapter 5**

**Appendix**

# List of Formulas

**Chapter 1**



**Chapter 2**



**Chapter 3**



**Chapter 4**



**Chapter 5**

# List of Abbreviations

| | |
|---|---|
| **3-PGA** | 3-phosphoglyceric acid |
| **A** | Assimilation rate of $CO_2$ |
| **ABA** | Abscisic acid |
| **ABCB14** | ABC Transporter B Family Member 14 |
| **AD** | Anno Domini |
| **ANOVA** | Analysis of variance |
| **Asat** | Light saturated rate of $CO_2$ uptake |
| **ATP** | Adenosine triphosphate |
| **BASL** | Breaking of asymmetry in the stomatal lineage |
| **BC** | Before Christ |
| **βCA1** | Beta carbonic anhydrase 1 |
| **βCA4** | Beta carbonic anhydrase 4 |
| **BLAST** | Basic Local Alignment Search Tool |
| **bp** | Base pairs |
| **$Ca^{2+}$** | Calcium |
| **CAM** | Crassulacean-acid metabolism |
| **cDNA** | Complementary DNA |
| **CHISM** | Chloroform/Isoamyl alcohol |
| **$C_i$** | Intracellular $CO_2$ concentration |
| **CNBH** | Carbon–nutrient balance hypothesis |
| **$CO_2$** | Carbon dioxide |
| **$CO_2$-eq** | Carbon dioxide equivalents |
| **C(t)** | Cycle threshold |
| **ddNTP** | Chain-terminating di-deoxynucleotides |
| **DNA** | Deoxyribonucleic acid |
| **DNMT** | DNA methyltransferase |
| **E** | Efficiency |
| **EPF1** | Epidermal patterning factor 1 |
| **EPF2** | Epidermal patterning factor 2 |
| **EST** | Expressed Sequence Tag |

| | |
|---|---|
| **ER** | ERECTA |
| **ERL1** | ERECTA like 1 |
| **ERL2** | ERECTA like 2 |
| **FACE** | Free air $CO_2$ enrichment |
| **FDR** | False discovery rate |
| **FLP** | Four lips |
| **FPKM** | Fragments per kilobase of exon per million mapped |
| **FST** | Fixation index |
| **G3P** | Glyceraldehyde 3-phosphate |
| **GA** | Gibberellic acid |
| **GC** | Guard cell |
| **GCA2** | Growth controlled by abscisic acid 2 |
| **GHGs** | Greenhouse gases |
| **GLM** | Generalized linear model |
| **GMC** | Guard mother cell |
| **GO** | Gene ontology |
| **$g_s$** | Stomatal conductance |
| **$H^+$** | Hydrogen |
| **HT1** | High leaf temperature 1 |
| **$H_2S$** | Hydrogen sulphide |
| **HSD** | Honestly significant difference |
| **IAA** | Indole-3-acetic acid |
| **IGA** | Istituto Di Genominca Applicata |
| **IRGA** | Infra-red gas analysis |
| **iWUE** | Water use efficiency |
| **JA** | Jasmonic acid |
| **Jmax** | PAR-saturated electron transport rate |
| **$K^+$** | Potassium |
| **Kbp** | Kilo-base pairs |
| **LTA** | Long term acclimation |
| **M** | Mole |

| | |
|---|---|
| **MET** | Metalliferous |
| **miRNA** | microRNA |
| **MMC** | Meristamoid mother cell |
| **N** | Nitrogen |
| **NADPH** | Nicotinamide adenine dinucleotide phosphate |
| **NGS** | Next generation sequencing |
| **NMET** | Non-metalliferous |
| **ns** | Not significant |
| **$O_2$** | Oxygen |
| **$O_3$** | Ozone |
| **PAR** | Photosynthetically active radiation |
| **Pb** | Lead |
| **PC** | Protodermal cell |
| **PCA** | Principle component analysis |
| **PCR** | Polymerase chain reaction |
| **PEPC** | Phosphoenolpyruvate carboxylase |
| **PHD** | Plant homeodomain |
| **PIN** | Pin-formed |
| **PPFD** | Photon flux density |
| **PPM** | Parts per million |
| **PSI** | Photosystem I |
| **PSII** | Photosystem II |
| **QTL** | Quantitative trait loci |
| **R** | Reported flourescence |
| **rbcL** | RuBisCO large subunits |
| **rbcS** | RuBisCO small subunits |
| **RNA** | Ribonucleic acid |
| **RNAi** | RNA interfering |
| **RNA-Seq** | RNA sequencing |
| **RPKM** | Reads per kilobase of exon per million mapped |
| **RT-PCR** | Quantitative real-time PCR |

| | |
|---|---|
| **R-Type** | Rapid transient |
| **RuBisCO** | Ribulose-1,5-*bis*phosphate carboxylase |
| **RubP** | Ribulose-1,5-bisphosphate |
| **S** | Spring |
| **SA** | Spring ambient |
| **SBL** | Sequence by ligation |
| **SBS** | Sequence by synthesis |
| **SCA** | Salicylic acid |
| **SD** | Stomatal density |
| **SDD1** | Stomatal density and distribution 1 |
| **SE** | Spring elevated |
| **SI** | Stomatal index |
| **siRNA** | Small interfering RNA |
| **SE** | Standard error |
| **SLAC1** | Slow Anion Channel-Associated 1 |
| **SLGC** | Stomatal lineage guard cell |
| **SNP** | Single nucleotide polymorphism |
| **$SO_2$** | Sulphur dioxide |
| **SPCH** | Speechless |
| **STA** | Short term acclimation |
| **S-Type** | Slow-activating sustained |
| **TPM** | Transcripts per million |
| **TPU** | Triose phosphate utilization |
| **TTM** | Too many mouths |
| **Vcmax** | Maximal carboxylation rate |
| **Zn** | Zinc |
| **Δ** | Carbon discrimination |
| **π** | Pi |
| **θ** | Theta |

# Author's Declaration

I Alex Watson-Lazowski, declare that this thesis entitled 'Evidence for Plant Adaptation to a Future High $CO_2$ World' and the work presented are both my own, and have been generated by me as the result of my own original research. I confirm that;

- This work was done wholly or mainly while in candidature for a doctoral degree at the University of Southampton.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institute, this has been clearly stated.
- Where I have consulted the published work of others, this has clearly been attributed.
- Where I have quoted from the work of others, the source has always been given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where this thesis is based on work jointly carried out with others, I have made clear what was done by others and what I have contributed myself.
- None of this work has been published before submission.

Signed:

Date:

# Acknowledgements

I would like to thank Professor Gail Taylor for securing the funding for this project, and for her fantastic supervision throughout. I would also like to thank NERC and EXPEER for funding the project. I have had the privilege of having several supervisors, and would like to thank Dr Mark Chapman and Dr Richard Edwards for their continued and enlightening supervision.

Access to the spring site in Bossoleto could not have been possible without the help of Professor Franco Miglietta of CNR, along with assistance collecting the seed itself from the site.

Numerous colleagues have assisted me on this project, including Dr Yunan Lin, of who I took over the project from, and trained and advised me at the start of my project. I would also like to thank Dr Maud Viger, Dr Adrienne Payne, Dr Hazel Smith and Dr Suzie Miler, who helped teach me an array of techniques which I have used throughout my project. I would also like to thank Dr Caitriona Murray and Dr Slivia Baronti who also assisted with seed collection from the spring site in Bossoleto. Thank you to the lab as a whole for help with the running and analysis of several experiments. Their enthusiastic assistance helped solidify both my experiments and analysis.

Assistance was provided by Dr Tracy Lawson and Jack Matthews of the University of Essex on the physiological analysis, of which I am very grateful, I could not have completed the chapter without their help.

Thank you to Peter Willson and Stephanie Ivie, both 3rd year project students who carried out their projects with work related to my thesis, and helped with the maintenance of my experiments and analysis of the data.

Thank you to my family for their continued support throughout my PhD, I could have not done it without them. I would also like to thank all of my friends, who have made the experience all the better.

*AW TL JWL JW DW SL ZL DM SJ NL MM JP MB DM DP HW AB DF TL MD RE JE JC ST JF SB ZH AD HG NW*

# General Introduction

## 1.1 Rising Atmospheric Carbon Dioxide

Since industrialisation, atmospheric carbon dioxide ($CO_2$) has risen and, over the past hundred years or so, there has been a consistent increase in the concentration of carbon dioxide in the atmosphere. $CO_2$ is the highest contributing greenhouse gas (Figure 1.1a), and so it is essential to understand this rise in terms of the atmosphere, but also how it will affect plants due to the role of $CO_2$ in plant photosynthesis. The concentration of atmospheric $CO_2$ currently stands at 398.03 µmol $mol^{-1}$ (also known as parts per million (ppm)) (Tans and Keeling 2014), and has the potential to rise to up to 1000 µmol $mol^{-1}$ by the year 2100 due to increased use of fossil fuels and decreased value of the rainforests as carbon sinks (Field, Barros *et al.* 2014). Most atmospheric $CO_2$ is contributed by the decay of organic material and forest fires, but is counter acted by biological processes such as photosynthesis removing $CO_2$ from the atmosphere.

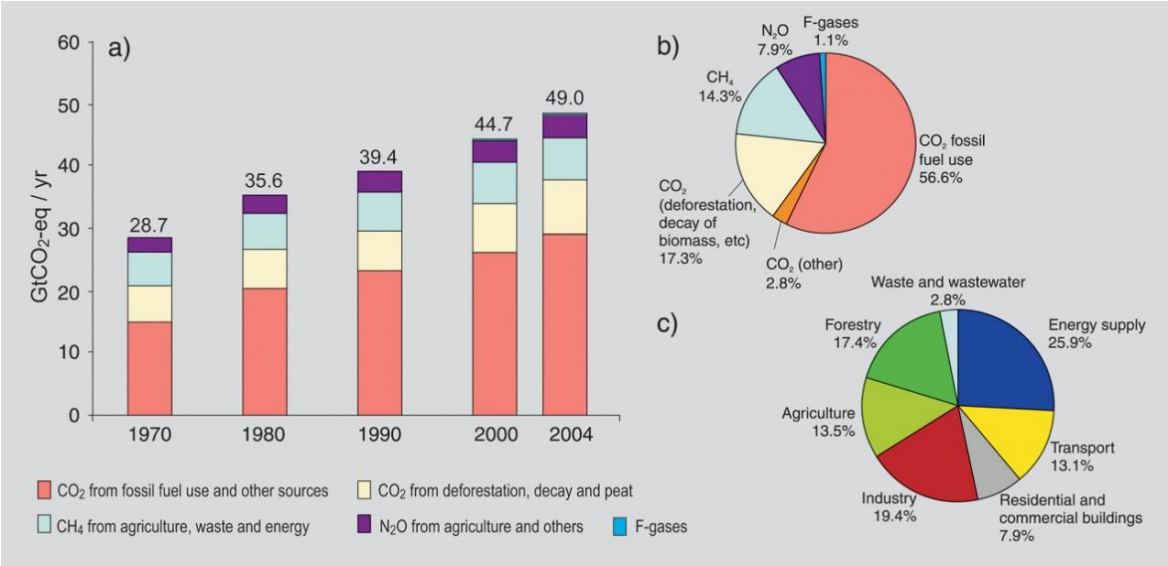



Figure 1.1 – (**a**) Global annual emissions of anthropogenic greenhouse gases (GHGs) from 1970 to 2004 (**b**) Share of different anthropogenic GHGs in total emissions in 2004 in terms of carbon dioxide equivalents ($CO_2$-eq). (c) Share of different sectors in total anthropogenic GHG emissions in 2004 in terms of $CO_2$-eq. (Forestry includes deforestation) (Bernstein, Bosch *et al.* 2007).

The balance of biological processes and organic decay can be disrupted by deforestation (Figure 1.1b). This, alongside an uncertain future of $CO_2$ emissions, has led to a range of predicted concentrations in the future (Griggs and Noguer 2002). Depending on how issues of carbon emissions and deforestation are addressed, the $CO_2$ concentration in 2100 could lie anywhere between 500 and 1000 µmol $mol^{-1}$ (Griggs and Noguer 2002), with the most realistic predictions lying around 700 µmol $mol^{-1}$. Numerous studies have been conducted to assess how elevated $CO_2$ concentrations will affect the utilisation of $CO_2$ by plants, and several general conclusions with respect to photosynthesis, respiration, plant growth, stomatal responses and utilisation of other

resources, have been suggested. These have been effectively summarised in a number of meta-analyses (Ainsworth, Davey *et al.* 2002, Leakey, Ainsworth *et al.* 2009), but due to the uncertainty of the mechanism of plant respiratory responses to elevated $CO_2$ consistent conclusions have not always been possible (Leakey, Xu *et al.* 2009). For example, both increases (Davey, Hunt *et al.* 2004) and decreases (Drake, Azcon-Bieto *et al.* 1999) in night time respiration levels of plants exposed to long term elevated $CO_2$ have been reported leading to the inability to predict whether plants will act as carbon sinks or sources in the future. This highlights the need to focus on understanding how the biosphere operates under elevated (relative to pre-industrial) atmospheric $CO_2$.

The ratio of gases in the atmosphere is not a constant, and the concentration of atmospheric $CO_2$ was previously low compared to today. At these previous lows concentrations dropped to less than half of today's concentration, for example around 10,000 years ago the concentration was around 180 μmol $mol^{-1}$ (Gerhart and Ward 2010). Intriguingly, over the 400 million years of the Phanerozoic era, periods of low atmospheric $CO_2$ concentrations are associated not only with adaptations such as high stomatal density (Beerling and Woodward 1997) but also the emergence of entirely new plant groups such as the ferns, pteridosperms and angiosperms (Woodward 1998). Atmospheric $CO_2$ has fluctuated gradually through geological time (Dippery, Tissue *et al.* 1995), and it is only recently that a rapid increase has occurred, since the beginning of industrialization from around 280 μmol $mol^{-1}$ to its present value of around 398 μmol $mol^{-1}$ (Tans and Keeling 2014). Evidence suggests that in the recent geological past, over 500,000 years, atmospheric $CO_2$ has fluctuated between 180 and 300 μmol $mol^{-1}$ (Figure 1.2), with the implications that adaptation may have occurred in response to changing conditions. This has led to the study of fossil plants used as a proxy to investigate the likely impacts of varying $CO_2$ concentration on plant traits. However, current conditions are significant in that they represent the first time that plants have been exposed to atmospheric $CO_2$ above 300 μmol $mol^{-1}$, for at least 500,000 years and probably considerably longer (Tans and Keeling 2014).

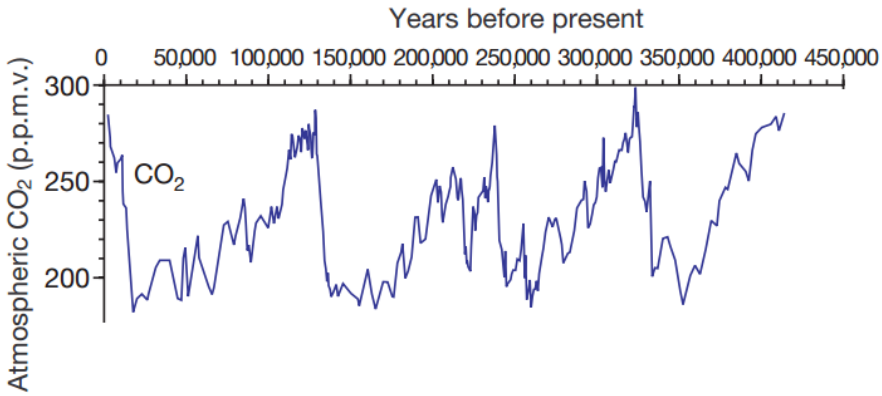



Figure 1.2 - Shows the concentration of $CO_2$ present in the atmosphere over the past 450,000 years. Modified from Sigman and Boyle (2000).

## 1.2 Adaptation and Acclimation

When a plant is placed under a selection pressure there are two possible modes of change enabling survival: adaptation or acclimation. An acclimation is a functional change which occurs rapidly and can occur within seconds or minutes. There are two types of acclimation, short term acclimation (STA) and long term acclimation (LTA) (Dietzel, Bräutigam *et al.* 2008). STA is the quicker response occurring within minutes of an environmental change, but it can be easily reversed. Due to the speed of the acclimation this typically involves pre-existing components within a biochemical pathway. In terms of elevated $CO_2$, an example of this could be the plant reducing the aperture of their stomata (Morison 1998). Another example of an STA is *Spinacia oleracea* (spinach) to high temperatures (Weis 1984). When incubated at 35˚C, spinach quickly exhibited a considerable decrease in the variable part of room temperature fluorescence. This gives rise to the suggestion that this transition lowered the reduction level of plastoquinone, i.e. increased electron flow through photosystem I, relative to photosystem II. LTA also begins within minutes but a phenotypic change does not become apparent until days, or weeks after the environmental change. These responses involve altered patterns of gene expression, reallocation of resources between the component process of photosynthesis and morphological change. An LTA will lead to the development of a different phenotype, and is a manifestation of phenotypic plasticity. It is LTA which represents an acclimation towards an advantage in an altered environment as these changes may not be immediately reversible. The most documented LTA to elevated $CO_2$ is a decrease in the content of ribulose-1,5-*bis*phosphate carboxylase (RuBisCO), the plant enzyme which fixes $CO_2$ (Sage 1994, Crafts-Brandner and Salvucci 2000), with the plant no longer requiring the same amount of the enzyme under elevated $CO_2$. Source–sink interactions, particularly with regard to carbon and nitrogen, are key determinants of photosynthetic acclimation. The biological control of these mechanisms, in response to elevated $CO_2$, is another example of a LTA acclimating to shifts in the source–sink balance (Wolfe, Gifford *et al.* 1998).

An adaptation is a response that occurs over multiple generations, leading to changes in the genetic structure of the population, due to natural selection towards a desired trait in response to a modified environment. Adaptive responses can involve phenology, growth and development, morphology and biochemistry. These responses occur to aid the survival of a plant and improve fitness by helping utilise nutrients, or protect itself from harsh conditions within its environment. An example of an adaptation is the rise of CAM (Crassulacean-Acid metabolism) plants, which have the ability to store $CO_2$ in vacuoles overnight, allowing stomata to remain closed during the day and increase water use efficiency (iWUE) (Han and Felker 1997). They have adapted to use the enzyme Phosphoenolpyruvate carboxylase to catalyse the fixation of $CO_2$ during the night.

This results in the formation of malic acid, which is stored in the vacuole. Malate is then released from the vacuole during the day and decarboxylated, with the resulting $CO_2$ being fixed via the Calvin cycle (Nimmo 2000). Plants found in arid environments, such as cacti, exhibit this adaptation. Another adaptation found is to high toxic metal concentrations in the soil, resulting in the adaptation of metallophyte species, such as *Thlaspi caerulescens* and *Arabidopsis halleri*, which are Cadmium (Cd) hyper-accumulators. The adaptation allows them to survive in soils with high levels of Cd; soils in which other species would not survive (Clemens 2006). Samples of the species *Anthyllis vulneraria* were taken from two metalliferous (MET) Zinc (Zn)- Lead (Pb) mine sites, and two non-metalliferous (NMET) sites and exhibited the same type of adaptation. Plants from each population were reciprocally grown in soil from the provenance of each population. The two *A. vulneraria* NMET populations exhibited high mortality and low growth rates in soil from the metalliferous sites and the two *A. vulneraria* MET populations exhibited a high growth rate in MET soils, but showed high mortality in NMET soils (Mahieu, Soussou *et al.* 2012). This highlights how the same species can adapt to survive in two different environments.

## 1.3 Plant Response to Elevated $CO_2$

### 1.3.1 Historic Studies

Present day seed plants exposed to historic conditions, demonstrate a clear response to increasing atmospheric $CO_2$, with a notable increase in overall size and plant mass across geological time. This was visualised by Dippery *et. al.* (1995) (Figure 1.3), where samples of *Abutilon theophrasti* were grown at glacial (150 µmol $mol^{-1}$) through to future (700 µmol $mol^{-1}$) $CO_2$ concentrations.

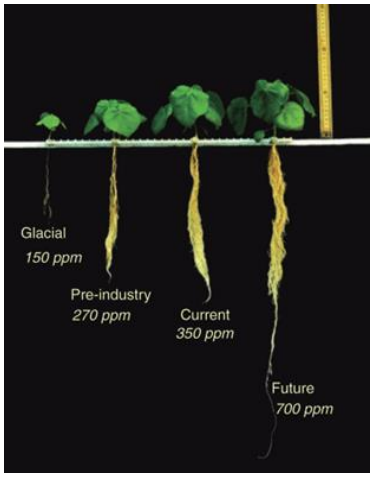



Figure 1.3 - Images of *Abutilon theophrasti* grown at four different concentrations of atmospheric $CO_2$ based on evidence of concentrations from previous eras (all plants were 14 days of age and were grown under similar water, light, and nutrient conditions). Modified from Dippery *et al.* (1995).

The change in plant size from historic to present atmospheric $CO_2$, is greater than size change from present to future atmospheric $CO_2$, despite the latter being a bigger absolute increase in atmospheric $CO_2$; and is consistent with most studies involving plant responses to past through to future $CO_2$ concentrations (Sage and Reid 1992, Polley, Johnson *et al.* 1993, Ward and Strain 1997). It has been suggested that plants may have already exhausted much of their potential to respond to rising $CO_2$, which raises concern, as it appears only a major evolutionary change in the future could provide a solution to the ever increasing atmospheric $CO_2$ (Gerhart and Ward 2010). However, these studies are limited by the fact that 'present day' seeds were exposed to the altered conditions, which provides only a limited snapshot of adaptive potential and how this may evolve in multi-generational exposure to future $CO_2$. However, it remains valuable to consider these response functions as part of the establishment of 'reaction norms' and a quantification of plastic responses to $CO_2$ that may ultimately provide insight into adaptive mechanisms.

Fossilised samples have provided material for analysis of plants from past millennia, providing physical evidence of trait changes induced by the altering atmospheric climate (Crane, Herendeen *et al.* 2004). Control of the stomatal pore complex is a vital physiological processes in plants, and is well-documented in fossilised plants because of the ease with which stomatal density can be assessed in fossil samples (McElwain and Chaloner 1995). Stomata are pores on the leaf surface which act as the plants airways, regulating the flow of gases between the plant and the environment. Stomata can adapt and acclimate to both local and global environmental changes on all timescales, from minutes to millennia (Hetherington and Woodward 2003). Monitoring the regulation of stomata in response to differential μmol $mol^{-1}$ of $CO_2$ is vital, because stomata are one of the limiting factors in the utilisation of $CO_2$. This has led to in-depth analysis of the stomatal density of fossilised plants, in a range of atmospheric $CO_2$ concentrations.

Beerling & Chaloner (1992) examined fossil leaves of *Salix herbacea* and found evidence of increased stomatal density and stomatal index during the low $CO_2$ of the Pleistocene, suggesting an adaptation for increased carbon uptake (Beerling and Chaloner 1992). Low $CO_2$ concentrations near 200 μmol $mol^{-1}$ (Sage 1995) during the Pleistocene era lasted tens of thousands of years, providing sufficient time for adaptation to the environment. A more comprehensive study showed how stomatal density has decreased over the last 3000 years (Beerling and Chaloner 1993) using samples of *Olea europaea* (olive), from material found in King Tutankhamun's tomb, which dates back to 1327 B.C. (Hepper 1990), and their own plants grown in 1991. Their results showed a decrease in stomatal density over the past 3000 years, including a more pronounced decrease over the past 200 years (Figure 1.4), which ties in with the start of the industrial era and the rise in $CO_2$ concentration.

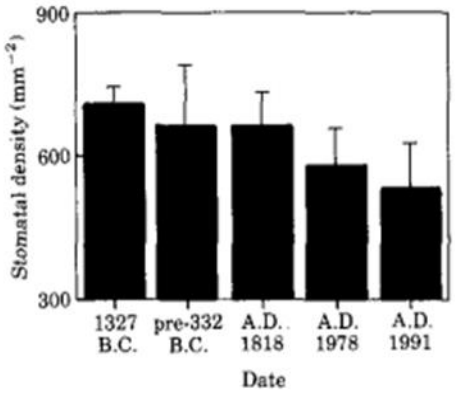



Figure 1.4 - Shows the stomatal density of *Olea europaea* samples from a range of dates over the past 3000 years (Beerling and Chaloner 1993).

Beerling & Chaloner along with other colleagues have since carried out numerous studies on the stomata of fossil plants (Beerling and Chaloner 1994, McElwain and Chaloner 1995, Chen, Li *et al.* 2001, Kouwenberg, McElwain *et al.* 2003). All of these studies, over a range of both species and time-scales, found that stomatal density has decreased over time (Table 1.1).

| ***Species*** | ***$CO_2$ Range*** | ***Stomatal Density*** | ***Stomatal Index*** | ***Reference*** |
|---|---|---|---|---|
| *Fagus sylvatica* | ↑(21%) | ↓(43%) | ^^ | (Paoletti, A. Raschi *et al.* 1997) |
| *Uqercus ilex* | ↑(21%) | ↓(28%) | | (Van de Water, Leavitt *et al.* 1994) |
| *Pinus flexilis* | ↑(30%) | ↓(17%) | ^^ | |
| Decidious woodland species | ↑(21%) | ↓(29%) | ^^ | (Woodward 1987) |
| *Salix cinerea* | ↑(27%) | ↓(36%) | ↓(24%) | (McElwain, Mitchell *et al.* 1995) |
| | ↑(19%) | ↓(22%) | ↓(16%) | |
| *Salix herbacea* | ↑^ | ↓^ | ^^ | (Beerling, Chaloner *et al.* 1993) |
| *Quercus petraea* | ↑^ | ↓^ | ↓^ | (Van Der Burgh, Visscher *et al.* 1993) |
| 14 species of trees, shrubs and herbs | ↑(20%) | ↓(21%) | n.s. | (Penuelas and Matamala 1990) |
| *Betula nana* | ↑^ | ↓^ | ^^ | (Beerling 1993) |
| *Olea europaea* | ↑(23%) | ↓(50%) | ^^ | (Beerling and Chaloner 1993) |
| | ↑(10%) | ↓(25%) | | |

Table 1.1 - Table of studies and results of stomatal density and index related to fossil samples. Modified from (McElwain and Chaloner 1995).^ Percentage change not provided, ^^ not recorded. All changes significant to P=≤0.05 unless n.s.(not significant) is stated.

The literature for increased atmospheric $CO_2$ beyond the ambient concentrations now found, is less consistent, with decreased stomatal numbers found in approximately 75% of the species studied (Woodward and Kelly 1995). Although this adaptation may not be detrimental to plants themselves, secondary effects also occur, such as decreased transpiration, which has been shown to alter the fresh water cycle and climate; and will continue to do so in the future (de Boer, Lammertsma *et al.* 2011). Additionally, reduced nutrient quality for insect and animal grazers due to lower nitrogen content has also been reported (Thompson and Drake 1994, Chamaillé-Jammes and Bond 2010).

### 1.3.2 Current Studies

A number of meta-analyses have been conducted on how plants adapt and acclimate to differential atmospheric $CO_2$; using artificial environments in which atmospheric $CO_2$ concentrations can be adjusted. Experiments scale from small controlled environments, to full field trials using FACE (Free Air $CO_2$ Enrichment) facilities (Li, Bohnert *et al.* 2007). FACE experiments have been used for decades to expose plants to expected future atmospheric $CO_2$ concentrations, and have provided us with a true insight into what we should expect from plants in the future. However, many of these studies focus on different aspects of phenotypic plasticity and acclimation, rather than on adaptation to elevated $CO_2$ over multiple generations. Many aspects can be monitored when looking for responses to $CO_2$, including radiation capture, canopy water use efficiency, harvest index (crop species), leaf nitrogen content and daily carbon gain (LeCain and Morgan 1998, Monje and Bugbee 1998). Leakey, Ainsworth *et al.* (2009) provided an overview of the 15 major elevated $CO_2$ FACE experiments (concentrations between 475-600 µmol $mol^{-1}$) which used fully replicated ($n \geq 3$) large plots (>100$m^2$), on different ecosystems in different parts of the world. Analysis of the data provided five important findings:

1. Elevated $CO_2$ stimulates photosynthetic carbon gain and net primary production over the long term, despite down-regulation of RuBisCO activity. Down-regulation of RuBisCO has been documented at the level of gene expression.
2. Elevated $CO_2$ improves nitrogen use efficiency.
3. Elevated $CO_2$ decreases water use in the leaf, and can lead to improved instantaneous leaf water use efficiency (iWUE), but canopy-scale results vary.
4. Elevated $CO_2$ does not directly stimulate C4 photosynthesis, but can indirectly stimulate carbon gain in times and places of drought.
5. The stimulation of yield by elevated $CO_2$ in crop species is much smaller than expected from studies conducted in highly controlled environments, but nevertheless there is still a positive stimulation of yield in future $CO_2$ .

The first four points indicate plants have a higher efficiency under elevated $CO_2$, which is expected as elevated $CO_2$ is increasing one of their main source of carbon for metabolism (Makino and Mae 1999). Elevated $CO_2$ concentrations increase net photosynthesis by increasing the carboxylation rate of RuBisCO, and competitively inhibiting the oxygenation of Ribulose-1,5-bisphosphate (RuBP) (Drake, Gonzalez-Meler *et al.* 1997). The down regulation of RuBisCO itself, is due to the increase in efficiency elevated $CO_2$ provides, meaning less enzyme is required (Sage 1994). A decrease in RuBisCO activity, stomatal conductance and even assimilation rate, have been shown to not lead to a decrease in net photosynthesis (Harley, Thomas *et al.* 1992). A 10% decrease in carbon assimilation, and a 32% decrease in stomatal conductance was identified in cotton grown in elevated $CO_2$ compared to ambient $CO_2$; but a 25% increase in net photosynthesis was measured (Harley, Thomas *et al.* 1992).

The increases in yield, correlates with findings of increased biomass under elevated $CO_2$, where at twice ambient $CO_2$; biomass was always significantly increased in woody species, regardless of other conditions (Curtis and Wang 1998). Point 5 also ties in the findings of Dippery *et al.* (1995), where the overall mass of samples under future atmospheric $CO_2$ does not increase as much as expected, when the amount of extra assimilates available to the plant is taken into consideration. These findings again suggest that plants may be beginning to exhaust their plasticity to acclimate to increased atmospheric $CO_2$, and are therefore, beginning to lose their ability to utilise the excess $CO_2$. This may also be due to other factors becoming limited, such as nitrogen availability (Reich and Hobbie 2013), which may be halting the potential extra growth.

Point 4 highlights the differences this effect will have on C3 plants compared to C4 plants. C3 and C4 plants differ in their photosynthetic pathway; C3 plants incorporate $CO_2$ into a 3-carbon molecule, whereas C4 incorporate it into 4-carbon molecule. C4 plants use the enzyme PEP Carboxylase (Phosphoenolpyruvate carboxylase) to increase the uptake of $CO_2$ and transport to RuBisCO, but C3 plants use RuBisCO itself to uptake $CO_2$, which is a less efficient process (Gowik and Westhoff 2011). Consequently C4 plants are a lot more efficient, which is why C4 species can survive in high temperature and drought environments, due to their stomata only opening for short durations. $CO_2$ levels therefore, affect C4 plants less than C3 plants, because their photosynthetic pathway is already saturated, and the excess $CO_2$ does not increase efficiency.

The stomatal mechanisms plants employ to utilise $CO_2$ leads to secondary advantages, which are detailed in points 3 and 4. The water use efficiency of plants under elevated $CO_2$ is increased, which in turn leads to an increase in drought resistance. All of these results suggest the plant is

moving towards more efficient mechanisms, both in the photosynthetic pathway, and $CO_2$ uptake through the stomata.

An empirical test of carbon acquisition under elevated $CO_2$ comes from soybean (*Glycine max*), one of the most studied species for response to elevated $CO_2$, and has been subject to extensive meta-analysis under the effects of elevated $CO_2$. A total of 111 studies have been brought together and have been tested for 25 different variables describing its physiology, growth and yield (Ainsworth, Davey *et al.* 2002). Table 1.2 shows some of the results from Ainsworth *et al.* (2002).

| ***Activity in Soybean*** | ***Increase/Decrease under elevated $CO_2$*** |
|---|---|
| Stimulation of soybean leaf $CO_2$ assimilation rate | 39% Increase |
| Stomatal conductance | 40% Decrease |
| RuBisCO activity | 11% Decrease |
| Canopy photosynthetic rate | 59% Increase |
| Total dry weight | 37% Increase |
| Seed yield | 24% Increase |
| Shoot : Root ratio | Remained the same |

Table 1.2 – The average differences in activity for several processes in soybean under elevated $CO_2$ (Ainsworth, Davey *et al.* 2002).

The results correspond with those found in the FACE studies, and again imply that the primary plant response to rising atmospheric $CO_2$ is to increase resource use efficiency (Drake, Gonzalez-Meler *et al.* 1997). The reduced activity of RuBisCO is again responsible for this increase in efficiency, whilst maintaining higher photosynthetic rates. A reduction of stomatal conductance and transpiration was noted, which in turn improves water use efficiency, also found in FACE studies. This may have important consequences for the ability of plants to avoid water stress, as the restriction of water loss is likely to benefit turgor maintenance and limit dehydration injury. This response appears consistent and beneficial to the plant. It would therefore follow that a plant would adapt/acclimate towards a decrease in the number of stomata to further increase resource use efficiency, as at high $CO_2$ concentrations a plant can decrease the development of stomata, while still maintaining the required $CO_2$ influx. In most cases this was true, and explained

by a decrease in stomatal density in response to elevated $CO_2$ leading to a decrease in stomatal conductance (Eamus 1991). Again, this ties in with the Beerling and Chaloners (1993) study of stomatal density of fossil samples. This is advantageous not only due to stomatal conductance, but also means construction costs of stomata can be saved. It is therefore, expected that adaptations of stomatal density to increased $CO_2$ concentrations are of selective value (Konrad, Roth-Nebelsick et al. 2008). This being said, studies examining the effect of a future increase in atmospheric $CO_2$ over 350 µmol $mol^{-1}$ on stomatal morphology have not been consistent between species, leaf age, or environments (Woodward and Bazzaz 1988, Ferris and Taylor 1994, Knapp, Cocke *et al.* 1994). Furthermore, Ainsworth and Rogers (2007) FACE experiment showed reduced stomatal conductance of 22%, yet the reduction was not associated with any change in stomatal density. Their results showed that stomatal conductance is consistently decreased in both C3 and C4 species, yet stomatal density does not change significantly and stomatal conductance does not acclimate to elevated $CO_2$ independently of photosynthesis (Ainsworth and Rogers 2007). It has therefore, been suggested that the short-term change in stomatal aperture is likely to determine most of the long-term response of stomatal conductance to elevated $CO_2$. This illustrates that future stomatal adaptations and acclimations remain to be completely elucidated.

Respiration is another aspect that has been extensively examined with regards to elevated $CO_2$. An increase in photosynthesis would suggest an increase in carbohydrate substrate for respiration, which has been confirmed (Cheng, Moore *et al.* 1998), but it has been suggested that dark respiration is inhibited by elevated $CO_2$ (Gonzalez-Meler, Ribas-Carbo *et al.* 1996). A wide range of results have been returned over the years in regards to respiration, from a 45% reduction (Gifford, Lambers *et al.* 1985), to an 11% increase (Davey, Hunt *et al.* 2004), or no change at all (Davey, Hunt *et al.* 2004). Long *et al.* (2004) showed that, re-evaluation of the methods used to measure dark respiration under elevated $CO_2$, showed that down regulation of respiration appeared to be an artefact of earlier measurement systems. The artefact arises due to adsorption, absorption and leakage of $CO_2$, both via chamber seals and via the intercellular air spaces of leaves (Long, Ainsworth *et al.* 2004). Alternatively, respiration can be measured using $O_2$ uptake, which avoids these limitations. When this method was used and the artefacts eliminated, down regulation of respiration was absent, explaining why there may be such a wide array of results in this area. This provides one explanation, but it is still important for further research to understand the reasons for this variation. A significant change in either direction could represent an increase or decrease in carbon released to the atmosphere, similar in size to current anthropogenic carbon emissions (Leakey, Xu *et al.* 2009).

The biosynthesis of secondary metabolites is an important component of plant defence against both biotic and abiotic stress. Elevated $CO_2$ has been shown to elicit a stress responses in plants, such as higher leaf temperature (Darbah, Sharkey *et al.* 2010), and so a secondary metabolite increase would be expected. The Carbon–Nutrient Balance Hypothesis (CNBH), proposed by Bryant *et al.* (1983), also indicated that this would be the case. The CNBH predicts that carbon-based defence compounds such as phenolics and terpenoids will increase as a result of the 'excess' carbon under elevated $CO_2$. Conversely, nitrogen-based defence compounds such as alkaloids, cyanogenic glycosides and glucosinolates, will decrease as a result of the scarce N (Bryant, Chapin III *et al.* 1983). Support for this hypothesis has been put forward (Matros, Amme *et al.* 2006), but a meta-analysis showed that the pattern is not consistent (Robinson, Ryan *et al.* 2012). The hypothesis appears to hold up in terms of N based secondary metabolites, but the individual groups within carbon based secondary metabolites seem to vary. Robinson *et al.* (2012) showed phenolics expression was increased, whilst terpeniods expression was decreased under elevated $CO_2$. This could be due to the specific nature of secondary metabolites to all environmental conditions, and so it is important to not generalise secondary metabolites.

Elevated $CO_2$ has been shown to alleviate the negative responses of other stresses such as heat, and drought, through the increase of antioxidants (Naudts, Van den Berge *et al.* 2014, Zinta, AbdElgawad *et al.* 2014). Biomass reduction, photosynthesis inhibition, chlorophyll fluorescence decline, $H_2O_2$, production and protein oxidation were all significantly mitigated when grown under drought and heat stress conditions, along with elevated $CO_2$ (compared to ambient $CO_2$). Analysis of enzymatic and molecular antioxidants revealed that the stress-mitigating $CO_2$ effect operates through up-regulation of antioxidant defence metabolism, resulting in lowered oxidative pressure (Zinta, AbdElgawad *et al.* 2014). The up-regulation is probably due to the already elevated secondary metabolites as a result of elevated $CO_2$ concentrations, and not a direct response to decrease oxidative stress of drought and heat, but serves to mitigate the oxidative stress (Pérez-López, Robredo *et al.* 2009). In Future climates therefore, where drought and elevated temperatures are expected, the negative impact may not be as great as predicted, as increasing $CO_2$ levels may alleviate the negative effects.

Research has been carried out into how this increase in resources in elevated $CO_2$ will affect the nutrient content of crops. This is important as currently a significant percentage of the human population is deficient in one or more essential dietary nutrients (Bloem and de Pee 2013), therefore any decrease in nutrition could exacerbate this already high threat to humanity. Research into grain crops found decreased concentrations of zinc, iron, and proteins in elevated

$CO_2$ in wheat (Manderscheid, Bender *et al.* 1995, Fangmeier, Grüters *et al.* 1997, Pleijel, Gelang *et al.* 2000), barley (Manderscheid, Bender *et al.* 1995), and rice (Seneweera and Conroy 1997). All of these investigations were chamber experiments, and therefore the changes maybe a consequence of what is known as, potting effect (Lieffering, Kim *et al.* 2004), because a small root volume in the chambers leads to nutrient dilution in the root–soil interface. Elevated $CO_2$ FACE experiments were carried out to test this hypothesis, with all the studies showing either no change, or a non-significant decrease (Prior, Runion *et al.* 2008, Högy, Wieser *et al.* 2009). Results that showed a non-significant decrease in nutrition were probably due to low number of replications. Recent research has looked into how elevated $CO_2$ will affect the nutrient content of crops on a meta scale, to see if the changes were significant on a larger scale (McGrath and Lobell 2013, Myers, Zanobetti *et al.* 2014). Myers *et al.* (2014) looked at both C3 and C4 crops to see how elevated $CO_2$ would affect their nutrient content. In C3 crops there was a significant decrease in Zinc, Iron and protein in wheat, rice and field peas. C4 grasses only showed limited significant differences, probably due to their photosynthetic system already being saturated under ambient $CO_2$. The decrease in nutrients in staple crops is a concern where nutrient rich crops will be required to feed the future population.

There are a limited amount of studies which have utilised controlled environments to investigate the multi-generational effect of $CO_2$ (Frenck, Van der Linden *et al*. 2013, Teng, Jin *et al.* 2009, Wieneke, Prati *et al.* 2004, Rey and Jarvis 1998). The limited number of studies looking at plants exposed to elevated $CO_2$ over several growing seasons have found a range of results, ranging from the $CO_2$ response being reduced when compared to the short-term responses to elevated $CO_2$ (Wieneke, Prati *et al.* 2004, Rey and Jarvis 1998), a heightened response to elevated $CO_2$ (Frenck, Van der Linden *et al*. 2013), and no change across generations (Teng, Jin *et al.* 2009). Wieneke *et al.* (2004) showed reductions in biomass after six growing seasons exposed to elevated $CO_2$ in *Sanguisorba minor.* On the contrary, Frenck *et al*. (2013) used *Brassica napus*, and showed over four growing seasons of exposure to elevated $CO_2$, aboveground biomass increased. Teng *et al.* (2009) found that after 15 growing seasons exposed to elevated $CO_2$ using *A. thaliana*, there were no changes in the response to elevated $CO_2$ across the growing seasons. These results do not give a clear indication of what plant responses to multi-generational exposure to elevated $CO_2$ will cause, but as the experiments were carried out on different species and over different numbers of generations there is potential for variation. These experiments were also carried out on a very limited number of generations in comparison to what plants will be exposed to naturally over time, and may not show the true multi-generational adaptations that could over hundreds of generations.

The predictions for future adaptations and acclimations can still not be fully clarified, but the outcomes of both adaptation and acclimation of plants seem similar, and give a good overview of what we should expect to see. Overall, plant response appears positive, resulting in a more efficient plant with the ability to accumulate more biomass. The negative effects, such as reduced nutrient content, effect human consumption more than the plants themselves. As $CO_2$ continues to rise, genotypes that are able to utilise the greater carbon supply may have a selective advantage, since additional carbon could be allocated to the acquisition of other limiting resources (Geber and Dawson 1993). Elevated $CO_2$ may also act as a strong selective agent through indirect effects on plants, such as increasing competition, in which genotypes with greater ability to utilise the excess carbon, would be favoured (Bazzaz, Jasieński *et al.* 1995).

## 1.4 Gene Expression Analysis

Recent advances in technology have allowed a comprehensive approach to be taken in identifying genetic explanations for phenotypic changes. Quantitative trait loci (QTL) investigations have been employed to identify differences in traits between plants exposed to elevated and ambient $CO_2$ concentrations (Fan, Cai *et al.* 2005, Rae, Tricker *et al.* 2007). Their results quantified and identified genetic variation in traits in response to elevated $CO_2$, providing an insight into the genomic response to the changing environment. This provides a good basis to investigate the expression of specific genes, particularly which genes are being switched on or off, or being up or down regulated in response to elevated $CO_2$; but this is the limit of QTL analysis. Dubbed "a revolutionary tool for transcriptomics," (Wang, Gerstein *et al.* 2009) RNA-Seq, a Next Generation Sequencing (NGS) method, refers to the use of high-throughput, deep-sequencing technologies to sequence cDNA in order to get information about the transcriptome of a given biological sample. The transcriptome is the complete set of transcripts (RNA molecules) in a cell, and their quantity, for a specific developmental stage or physiological condition (Wang, Gerstein *et al.* 2009). Prior to RNA-Seq, transcriptomics was originally measured using Northern blots (Van Oosten and Besford 1996), advancing onto more sensitive techniques such as microarrays (De Souza, Gaspar *et al.* 2008) and quantitative real-time PCR (RT-PCR) (Bandres, Cubedo *et al.* 2006). The limitation of these techniques was the number of genes that could be analysed, as only a selection of genes could be chosen for analysis. This meant potentially important gene changes could be missed. Furthermore, for microarrays and RT-PCR, only species with extensive genomic resources could be assayed. RNA-Seq provides a far more complete and precise measurement of levels of transcripts, and their isoforms than any other method, allowing for whole pathway analysis across the genome. RNA-Seq not only provides expression levels of genes without specific prior knowledge of the genome, but also a comparison of expression levels between genes can be made. RNA-Seq has opened up exciting new opportunities to elucidate adaptive and acclamatory responses, due to prior genomic knowledge not being required and providing comparisons on a much larger scale by allowing whole genome analysis across a high number of samples. Not only does RNA-Seq provide a comprehensive expression profile, but also produces the sequence data for each gene analysed.

## 1.5 Gene Expression Changes in Response to Elevated $CO_2$

Phenotypic and physiological changes have the potential to be explained by underlying gene expression changes, whether these are small increases or decreases in expression, or genes no longer being transcribed. Identifying changes in gene expression is important in elucidating the pathways which underpin variation in phenotypic traits, including plants exposed to elevated $CO_2$. Gene expression changes in response to elevated $CO_2$ were initially investigated using Northern blots (Van Oosten and Besford 1996). They showed that after exposure to elevated $CO_2$, RuBisCO small subunits (rbcS) and RuBisCO large subunits (rbcL) were both down regulated in expression (Figure 1.5). Work on RuBisCO has almost exclusively continued to show the down regulation of the enzyme under elevated $CO_2$ when no additional factors are involved (Gesch, Kang *et al.* 2003, Gupta, Duplessis *et al.* 2005, Taylor, Street *et al.* 2005, Rae, Tricker *et al.* 2007, Fukayama, Sugino *et al.* 2011). It has also been shown that when leaves are supplied with increased levels of glucose and sucrose (which is a downstream effect of elevated $CO_2$), the response is mimicked, suggesting it is caused by the accumulation of carbohydrates (Van Oosten, Wilkins *et al.* 1994).

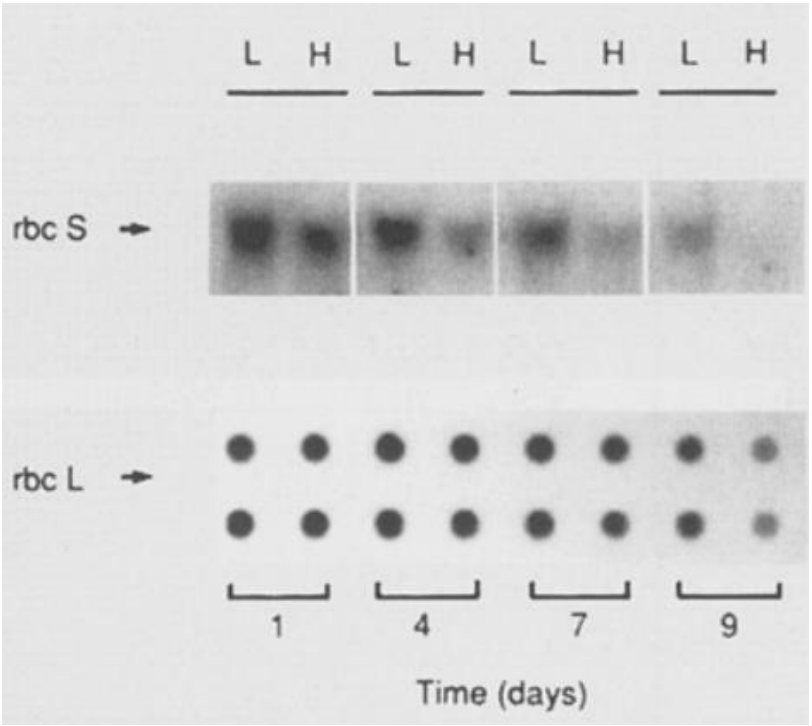



Figure 1.5 – Northern blot analysis of abundance of rbcS or rbcL mRNA in an attached leaf of tomato plants growing under ambient (L) or elevated (H) $CO_2$ for up to 9 days (Van Oosten and Besford 1996).

Since these early studies, experiments have moved towards global gene expression analysis using microarrays and NGS, with the first papers published from controlled environments and FACE experiments (Taylor, Street *et al.* 2005, Ainsworth, Rogers *et al.* 2006, Li, Sioson *et al.* 2006). Using micro-arrays, a screening of transcripts involved in leaf development in soybean (using SoyFACE), under elevated $CO_2$ was looked into by Ainsworth *et al.* (2006). They found 327 (out of 5,314) transcripts showed significant differential expression in developing and mature leaves in response to elevated $CO_2$, and categorised these transcripts to see where change was occurring (Figure 1.6).

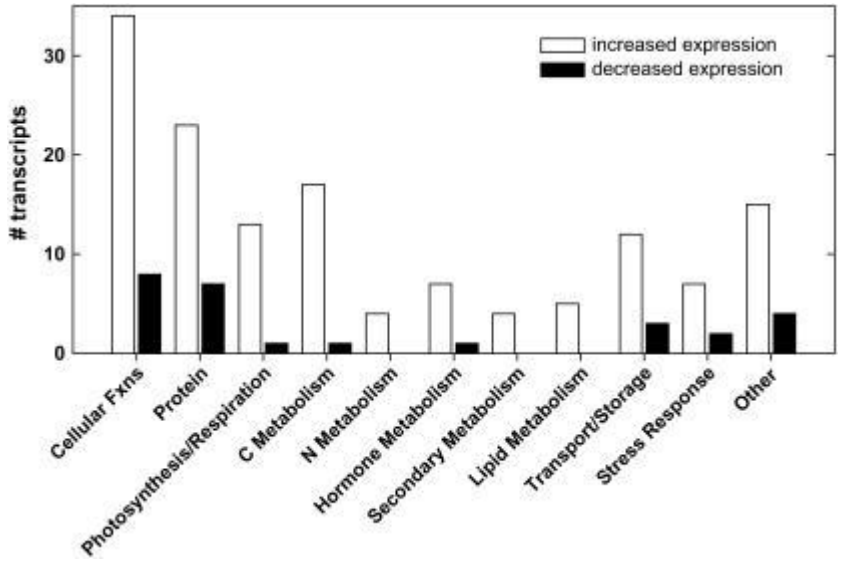



Figure 1.6 - Categorical distribution of genes showing differential expression under elevated $CO_2$ (Ainsworth, Rogers *et al.* 2006).

The profile suggests the stimulation of the respiratory breakdown of carbohydrates, which would provide increased energy and biochemical precursors for leaf expansion and growth at elevated $CO_2$. Increased photosynthesis and carbohydrate production at elevated $CO_2$ has been extensively noted in the literature, but this was the first evidence that at the transcript level, respiratory breakdown of starch is also increased with elevated $CO_2$. Taylor *et al.* (2005) used POPFace (*Poplar* grown using FACE) and microarrays to study the effects in *Poplar*. They found fewer genes to be significantly differentially expressed; only 28 out of 35,000 ESTs (Expressed Sequence Tag) were differentially expressed in semi-mature leaves. One of the genes identified, pyruvate kinase, is involved in the step between carbohydrate metabolism and protein synthesis, but interestingly this gene was down regulated. These variations in the respiration pathway have been noted in other cases (Gonzalez-Meler, Taneva *et al.* 2004), and has led to extensive evaluation in terms of gene expression under elevated $CO_2$ in attempt to clarify the response of the pathway. Leakey *et*

*al.* (2009) looked at gene expression of the whole of the respiration pathway in soybean, and found that dark respiration appeared to be mainly up regulated in elevated $CO_2$ (Figure 1.7).

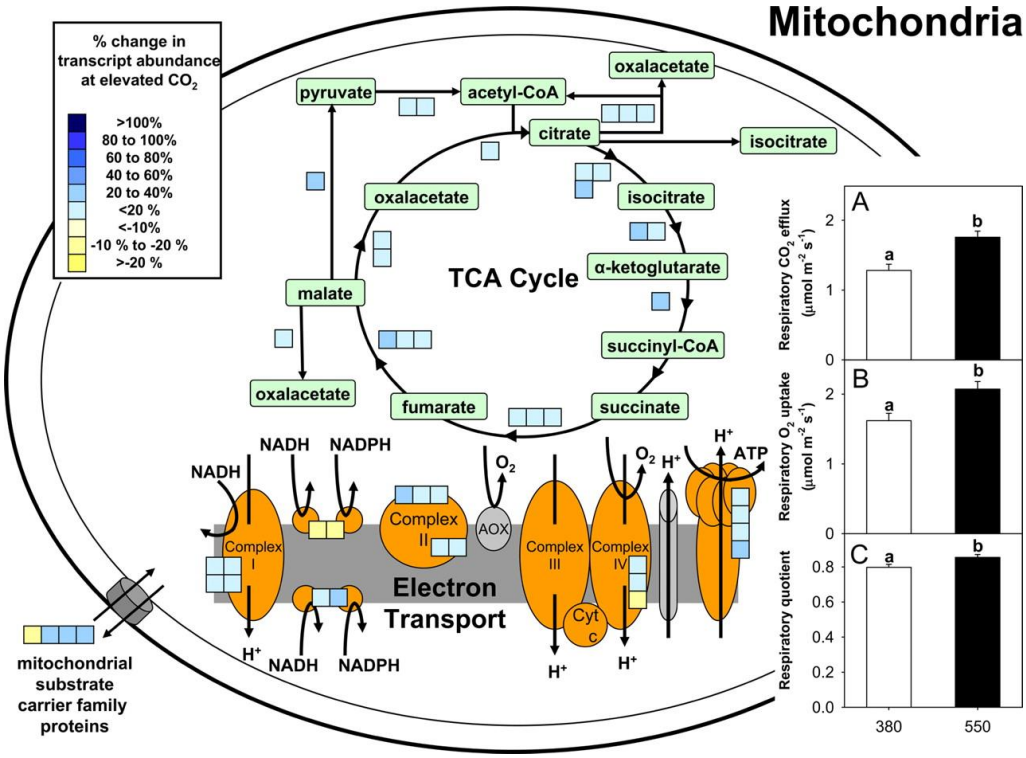



Figure 1.7 - Graphical representation of carbon metabolism in mitochondria of mature, sun leaves of soybean grown at SoyFACE. (Inset) Mean treatment values (± SE) of the night time rate of respiratory $CO_2$ efflux ($\mu mol \cdot m^{-2} \cdot s^{-1}$) (**A**), rate of respiratory $O_2$ uptake ($\mu mol \cdot m^{-2} \cdot s^{-1}$) (**B**), and respiratory quotient (**C**). Each blue or yellow box represents the statistically significant treatment response ($P < 0.05$) of a unique transcript encoding an enzyme or protein structure (Leakey, Xu *et al.* 2009).

This response has been also noted in *Arabidopsis* (Fukayama, Sugino *et al.* 2011), where most genes exhibited an increase in expression across the pathway. Markelz *et al.* (2014) looked into how the different developmental stages affect respiration transcript abundance in *Arabidopsis* under elevated $CO_2$; and showed mature leaves exhibited more respiratory transcripts than young leaves. This conclusion is not unique, Carrie, Murcha *et al.* (2013) showed that mitochondrial transcript abundance for components of the respiratory pathway increase per mesophyll cell, as cells expand and move through the developmental stages. The expression of the respiratory system was investigated across several stages of development (Figure 1.8), primordia (16 days after germination), expanding (23 days after germination) and mature (30 days after germination).

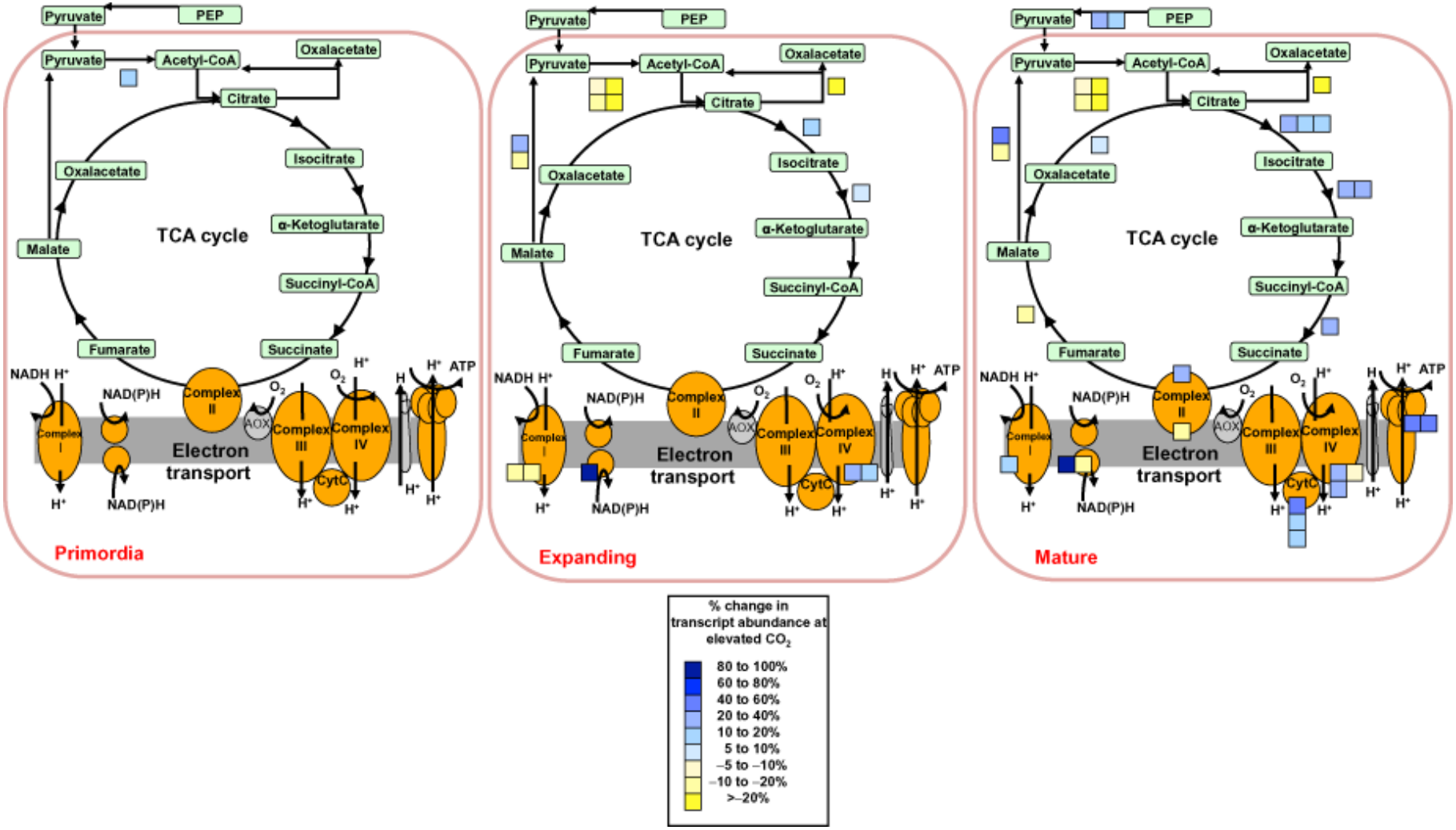



Figure 1.9 - A graphical representation of the relative abundance at midnight of transcripts encoding components of carbon metabolism pathways in mitochondria in leaves of *Arabidopsis thaliana* grown at ambient $CO_2$ (370 µmol $mol^{-1}$) or elevated $CO_2$ (750 µmol $mol^{-1}$), at 16 (Primordia), 23 (Expanding) and 30 (Mature) days after germination. Green boxes represent metabolites. Arrows represent reactions. Each box represents the average treatment response of a unique transcript that responded significantly ($P < 0.05$) to elevated $CO_2$ in a given leaf developmental stage (Markelz, Vosseller *et al.* 2014).

Markelz *et al.* (2014) showed that in the primordia stage, elevated $CO_2$ had a minimal effect on transcript abundance of the respiratory pathway. In expanding leaves, elevated $CO_2$ induced greater glucose content and transcript abundance for some respiratory genes, but did not alter respiratory $CO_2$ efflux. Finally in mature leaves, elevated $CO_2$ led to greater glucose, sucrose and starch content, plus greater transcript abundance for many components of the respiratory pathway, and greater respiratory $CO_2$ efflux; the effect previously noted by Leakey *et al.* (2009). These results suggest that growth at elevated $CO_2$ stimulates dark respiration only after leaves transition from carbon sinks into carbon sources. This could explain the variation in respiration measurements from previous studies, as no consistent time point is used throughout studies on respiration. Taylor *et al.* (2005) looked into differences between young and semi-mature leaves in terms of carbon-compounds and carbohydrate metabolism. They showed 15 genes were up regulated, and 2 down regulated in semi-mature leaves compared to young leaves, and this was largely not affected by elevated by $CO_2$ (one less gene was up regulated and two more were down regulated). This agrees with more carbohydrate metabolism occurring in mature leaves, increasing respiration, but suggests it occurs irrelevant of $CO_2$ concentration, although it has been previously shown to be enhanced by elevated $CO_2$ (Ainsworth, Rogers *et al.* 2006).

Gene expression of the photosynthetic pathway has been expanded from genes involved in the Calvin cycle, to include gene expression of both photosystem I (PSI) and photosystem II (PSII). Although Ainsworth *et al.* (2006) showed photosynthesis transcripts were increased in expression, recent results showed an overall decrease in expression in genes across the pathway (Li, Sioson *et al.* 2006, Li, Ainsworth *et al.* 2008, Leakey, Xu *et al.* 2009, Kaplan, Zhao *et al.* 2012). When C3 leaves are exposed to elevated $CO_2$, there is an immediate increase in net photosynthesis (Kramer 1981, Cure and Acock 1986) because of decreased photorespiration, which is negatively affected by an increase in atmospheric $CO_2$:$O_2$ ratios (Bowes 1991). Stimulation of net photosynthesis has also been observed even when plants are grown for long periods at high $CO_2$ (Davey, Parsons *et al.* 1999, Ainsworth, Rogers *et al.* 2003, Ainsworth and Long 2005). However, photosynthetic gene expression is often decreased in plants grown for long periods at elevated $CO_2$, due the plants becoming acclimated, and not requiring as much investment in the pathway to increase/maintain photosynthesis (Paul and Foyer 2001). When C3 plants are grown for long periods under high $CO_2$, they tend to become carbohydrate rich, and as with RuBisCO, alterations of gene expression patterns in response to elevated $CO_2$ are in part, thought to be related to altered sugar signalling, as the expression of photosynthetic genes is rapidly inhibited when plants are supplied with exogenous sugars (Foyer, Neukermans *et al.* 2012). The acclimation of photosynthesis observed in C3 plants grown for long periods at elevated $CO_2$, is related to the

substantial re-programming of gene expression in the photosynthesis pathway (Leakey, Xu *et al.* 2009, Kaplan, Zhao *et al.* 2012).

Other aspects of plant development have also been investigated, including genes involved in cell wall synthesis and strengthening (Taylor, Street *et al.* 2005, Ainsworth, Rogers *et al.* 2006). Cell wall synthesis and cell wall loosening genes have been shown to be up regulated in elevated $CO_2$ (Taylor, Street *et al.* 2005, Ainsworth, Rogers *et al.* 2006), which would indicate a stimulation or prolonged expansion in leaves This was confirmed in the literature, where cell wall extensibility has been measured in elevated $CO_2$ which was found to be stimulated (Taylor, Tricker *et al.* 2003).

## 1.6 Stomata

The phenotypic changes identified in response to elevated $CO_2$ are diverse, but one of the main proxies which is used to measure response to $CO_2$ is stomatal numbers. Stomata are pores formed by a pair of specialised cells called the guard cells (Willmer and Fricker 1996). They are found on the surface of the aerial part of the majority of higher plants, and open and close to control gas exchange between a plant and its environment. They are portals for entry of $CO_2$ into the leaf for photosynthesis, and also an exit for oxygen and water vapour through transpiration. The stomata's major function is to allow sufficient $CO_2$ to enter the leaf to optimise photosynthesis under the prevailing conditions, whilst conserving as much water as possible. Stomata mainly form part of the lower epidermis, which sit below the mesophyll cells, in which photosynthesis occurs (Figure 1.9).

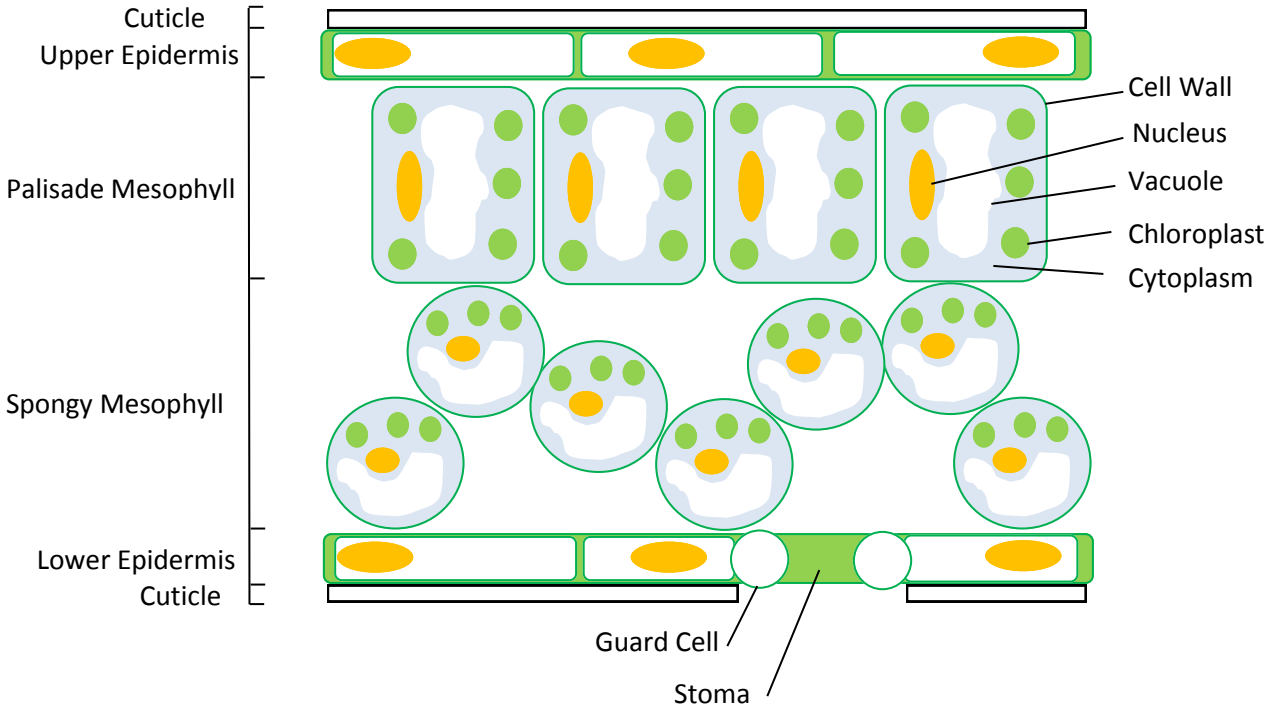



Figure 1.9 – Diagram of the internal structure of a leaf.

There are two types of mesophyll cells, spongy and palisade (Takebe, Otsuki *et al.* 1968). Both contain chloroplasts, but they are most concentrated in the palisade cells, hence the majority of photosynthesis occurs within these cells. The spongy mesophyll cells are closer to the stomata and their main role is to allow for the interchange of gases (mainly $CO_2$). The palisade mesophyll cells are closer to the upper epidermis and so are more exposed to light, which is why the chloroplasts are more concentrated in those cells.

Stomata can greatly differ in numbers and patterns between species, and these are often dependent on environment factors of the species habitat, such as $CO_2$ concentration and drought (Croxdale 2000). Monitoring these factors is essential in determining whether atmospheric $CO_2$ is the factor affecting the plant on a physiological level. There are two main measures to determine leaf stomatal numbers, stomatal density and stomatal index. These calculations have been widely used for decades on both current and fossil samples to determine response retrospectively to atmospheric $CO_2$.

### 1.6.1 Stomatal Density

Stomatal density can be calculated as follows:

$$\text{Stomatal Density} = \frac{\text{No. of Stomata}}{\text{Field of View } (mm^2)}$$

Formula 1.1 – The formula used to calculate stomatal density.

One of the most documented responses to elevated concentrations of $CO_2$ in plants is a change in stomatal density (Woodward and Kelly 1995). Numerous experiments have measured stomatal density, on a variety of species under controlled $CO_2$ environments. Stomatal density responses to elevated $CO_2$ are not always consistent, with studies showing a decrease in stomatal density (Beerling and Chaloner 1993, McElwain, Mitchell *et al.* 1995), studies showing that elevated $CO_2$ has no consistent effect (Ryle and Stanley 1992, Gay and Hauck 1994), plus fewer studies noting an increase in stomatal density (Malone, Mayeux *et al.* 1993). The most documented response to an increase in atmospheric $CO_2$ is a decrease in stomatal density (Penuelas and Matamala 1990). This was confirmed by Woodward and Kelly (1995) who reported that from 122 observations of 100 species, 74% exhibited a decrease in stomatal density in elevated $CO_2$.

One of the species exhibiting the less common, 'increased numbers' phenotype was *Ranunculus glacialis*, an arctic-alpine plant which is found in southern Europe in the high mountains, which Woodward and Kelly (1995) found in elevated $CO_2$ conditions increased its stomatal density. Malone *et al.* (1993) carried out a study on the same species and showed no consistent effect on stomatal density. They concluded this response may be due to a lack of plasticity to significantly alter stomatal density after only a single generation or growing season, which is something we may have to account for when designing any experimental plans (Malone, Mayeux *et al.* 1993). This conclusion has to be taken into account when concluding that a species does not show a consistent effect. The most likely explanation however for varying responses of stomatal density

to elevated $CO_2$ is the fact that this parameter is an indirect measure only of stomatal numbers. Stomata are embedded in epidermal cells and it is well-documented that plant exposure to elevated $CO_2$ can lead to both increased and decreased epidermal cell expansion (Ferris, Sabatti *et al.* 2001, Taylor, Tricker *et al.* 2003). This being the case, stomatal density is a reflection not only of stomatal initiation but also how much expansion has occurred in leaf epidermal cells. All data from experiments where stomatal density alone are presented and no stomatal index (see below) or epidermal cell size, should be treated with extreme caution. In such circumstances no 'inferences' may be made about how $CO_2$ affects stomatal development and yet a glance at Table 1.1 reveals that this is the most-often used metric in studies on ancient specimens.

### 1.6.2 Stomatal Index

Although stomatal density is regularly used as an estimate of stomatal numbers, stomatal index is an assessment of stomatal development, as it also takes into account the relative number of epidermal cells. Stomatal index is calculated as follows (Salisbury 1928);

$$\text{Stomatal Index (\%)} = \frac{\text{Stomatal Density}}{(\text{Stomatal Density} + \text{Epidermal Cell Density})} * 100$$

Formula 1.2 – The formula used to calculate stomatal index.

Stomatal index should be used alongside stomatal density, as together they give a more complete understanding of the stomatal patterning that is occurring, but as stated above, for many studies this measure is absent. This is usually because the measure cannot always be carried out, for example when ancient samples of leaf fragments are studied, stomatal index calculations are impossible. Stomatal index largely follows the same pattern as stomatal density and is decreased under elevated $CO_2$. McElwain and Chaloner (1995) looked into prehistoric stomatal density studies, and how many of these studies also used stomatal index (Table 1.1). Out of nine studies three also measured stomatal index, with two showing a significant decrease in stomatal index as $CO_2$ increased, and one showing a no significant difference. Royer (2001) went on to later produce more of a meta-study on stomatal index comparing both prehistoric (sub-ambient $CO_2$) to ambient $CO_2$, and the latter in relation to elevated $CO_2$.Royer (2001) found 89% of species showed a decrease in stomatal index when sub-ambient $CO_2$ was compared to ambient $CO_2$, and 29% showing a decrease when comparing ambient to elevated $CO_2$.Only 4% in either showed the opposite reaction (an increase in stomatal index in response to elevated $CO_2$), and the rest

showed no significant difference. Interestingly the number of species that decreased stomatal numbers dramatically decreased when comparing sub-ambient to ambient, and ambient to elevated $CO_2$, again suggesting plants are already exhausting their ability to respond to elevated $CO_2$, or are now adapting/acclimating in novel ways.

### 1.6.3 Stomatal Patterning

Stomata can from on the adaxial (upper) or abaxial (lower) part of the leaf surface, and there are species which have stomata only on the abaxial surface (hypostomatous), or on both (amphistomomatus). Whether this pattering effects how the plant responds to elevated $CO_2$ has been investigated (Woodward and Kelly 1995, Royer 2001). Woodward and Kelly (1995) showed there was no strong relationship between either hypostomatous and amphistomomatus species, or the abaxial and adaxial surface of amphistomomatus species. Royer (2001) found hypostomatous species more often inversely responded to $CO_2$for both stomatal density ($P$<0.03) and stomatal index ($P$<0.001), and within amphistomatous species, neither the abaxial nor adaxial surface yield statistically different responses. No significant difference between abaxial and adaxial surfaces appears consistent, though there are still cases where these do differ (Marchi, Tognetti *et al.* 2004). The significant difference between hypostomatous and amphistomomatus species is interesting, as it could relate to the $CO_2$ sensing mechanisms contained within the stomata acting differently whether they are on the abaxial or adaxial surface of the leaf. The number of stomata has also been shown to be negatively correlated with the size of the stomata (Doheny-Adams, Hunt *et al.* 2012), and as numbers decrease, the size of the stomata increases. This decrease in numbers and increase in size has proven to be more water use efficient for that plant, indicating that fewer larger stomata control stomatal conductance more efficiently, over more, but smaller stomata.

It is important to be aware of any previously documented genes that affect stomatal patterning, as these will be the genes investigated when attempting to explain any stomatal phenotypes. An extensive set of genes have been documented which affect stomatal patterning, with confirmed effects using mutant phenotypes (in *Arabidopsis thaliana*), Table 1.3. These have provided an insight into the development of the stomata.

| *Gene name* | *Function* | *Null Mutant* |
|---|---|---|
| Too many mouths (TMM) (Yang and Sack 1995) | Encodes putative cell-surface receptor, repress divisions. | Fail to orient the asymmetry of spacing divisions, fail to inhibit asymmetric divisions in cells adjacent to two or more stomata or their precursor cells and have a reduced number of amplifying divisions which leads to the premature conversion of meristemoids into guard mother cells. |
| ERECTA (ER) (Masle, Gilmore *et al.* 2005) | Involved in specification of organs originating from the shoot apical meristem. Contains a cytoplasmic protein kinase catalytic domain, a transmembrane region, and an extracellular leucine-rich repeat. ER has been identified as a quantitative trait locus for transpiration efficiency by influencing epidermal and mesophyll development, stomatal density and porosity of leaves. Together with ERL1 and ERL2, ER governs the initial decision of protodermal cells to either divide proliferatively to produce pavement cells or divide asymmetrically to generate stomatal complexes. | Loss of function in all Erecta family gene (*er,erl1* and *erl2*) shows striking stomatal over-proliferation and spacing defects. |
| Erecta Like 1 (ERL1) (Shpak, McAbee *et al.* 2005) | Encodes a receptor-like kinase that, together with ER and ERL2 governs the initial decision of protodermal cells to either divide proliferatively to produce pavement cells or divide asymmetrically to generate stomatal complexes. Along with erl2 functionally compensates for loss of erecta during integument development. | Loss of function in all Erecta family gene (*er,erl1* and *erl2*) shows striking stomatal over-proliferation and spacing defects. |
| Erecta Like 2 (ERL2) (Shpak, McAbee *et al.* 2005) | Encodes a receptor-like kinase that, together with ER and ERL1 governs the initial decision of protodermal cells to either divide proliferatively to produce pavement cells or divide asymmetrically to generate stomatal complexes. It is also important for maintaining stomatal stem cell activity and preventing terminal differentiation of the meristemoid into the guard mother cell. | Loss of function in all Erecta family gene (*er,erl1* and *erl2*) shows striking stomatal over-proliferation and spacing defects. |
| YODA (Gray and Hetherington 2004) | Alter stomatal density and spacing; possess a long N-terminal extension with negative regulatory activity. | Mutant exhibits severe reduction in overall plant height and internode length as well as excess production of guard cells. |
| Stomatal Density and Distribution 1(SDD1) (Schlüter, Muschak *et al.* 2003) | Expressed in meristemoids and GMCs, represses stomatal divisions and also causes arrest of meristemoids and GMCs. | Excessive entry divisions and fewer amplification divisions and can fail to orient spacing divisions; increase the stomatal index and disorient spacing divisions. |
| Four Lips (FLP) (Lai, Nadeau *et al*. 2005) | The *FLP* gene is involved in generating normal stomatal patterning. | Recessive mutations in the *flp* gene abnormally induce at least four guard cells in contact with one another. |
| MYB88 (Lai, Nadeau *et al.* 2005) | Encodes a putative transcription factor involved in stomata development. | Double loss of *myb88* and *flp* activity results in a failure of guard mother cells to adopt the guard cell fate, thus they continue to divide resulting in abnormal stomata consisting of clusters of numerous guard cell-like cells. This phenotype is enhanced in double mutants over the single mutant *flp* phenotype. |

| | | |
|---|---|---|
| FAMA (Ohashi-Ito and Bergmann 2006) | Encodes a basic helix-loop-helix transcription factor whose activity is required to promote differentiation of stomatal guard cells and to halt proliferative divisions in their immediate precursors. Both transcript and protein are expressed in and are required for halting divisions at the end of the stomatal lineage. It also has a role in the promotion of guard cell fate and in controlling the transition from guard mother cell to guard cell. | *fama* null mutants make meristemoids and GMCs, but the GMCs cannot progress to become guard cells and instead continue rounds of symmetric division while maintaining GMC marker expression, eventually forming small epidermal "tumors" |
| MUTE (Pillitteri, Sloan *et al.* 2007) | Encodes a basic helix-loop-helix (bHLH) protein that controls meristemoid differentiation during stomatal development. Epidermal cells lose their competence to respond to MUTE overexpression during cotyledon development. | In the absence of MUTE, meristemoids abort after excessive asymmetric divisions and fail to differentiate stomata. |
| Breaking of Asymmetry in the Stomatal Lineage (BASL) (Dong, MacAlister *et al*. 2009) | A regulator of asymmetric divisions. In asymmetrically dividing stomatal-lineage cells, BASL accumulates in a polarized crescent at the cell periphery before division, and then localizes differentially to the nucleus and a peripheral crescent in self-renewing cells and their sisters after division. | Mutants produce excessive numbers of small epidermal cells, and their stomata are found in a miss-patterned (clustered) distribution, the defects in *basl* centre on a loss of asymmetry intrinsic to the divisions. |
| CDKB1;1 (Boudolf, Barrôco *et al.* 2004) | Arabidopsis homolog of yeast cdc2, a protein kinase (cyclin-dependent kinase) that plays a central role in control of the mitotic cell cycle. | Plants that overexpress a dominant negative allele of *cdkb1;1* have abnormal stomata and a decreased number of stomatal complexes. |
| SCRM2 (Hofmann 2008) | Encodes ICE2 (Inducer of CBF Expression 2), a transcription factor of the bHLH family that participates in the response to deep freezing through the cold acclimation-dependent pathway. Overexpression of ICE2 results in increased tolerance to deep freezing stress after cold acclimation. | Loss of *scrm* and *scrm2* recapitulated the phenotypes of *fama*, *mute*, and *spch.* Gain-of-function mutant has an epidermis made entirely of stomata. |
| Epidermal Patterning Factor 1(EPF1) (Hara, Kajita *et al.* 2007) | Encodes a secretory peptide EPF1 involved in stomatal development. EPF1 is related to EPF2 which controls asymmetric cell divisions during stomatal devlopment. | Induction of e*pf1* or *epf2* in *Arabidopsis* inhibits stomatal differentiation, resulting in seedling lethality. Over expression leads to decreased stomatal density. |
| Epidermal Patterning Factor 2 (EPF2) (Hara, Yokoo *et al.* 2009) | Encodes a secretory peptide EPF2 expressed in proliferating cells of the stomatal lineage, known as meristemoids, and in guard mother cells, the progenitors of stomata. Controls asymmetric cell divisions during stomatal development. | Induction of *epf1* or *epf2* in *Arabidopsis* inhibits stomatal differentiation, resulting in seedling lethality. Over expression leads to decreased stomatal density. |
| Beta Carbonic Anhydrase 1 (βCA1) (Hu, Boisson-Dernier *et al.* 2010) | Encodes a putative beta-carbonic anhydrase BCA1. Together with BCA4 regulates $CO_2$-controlled stomatal movements in guard cells. | Double-mutant plants in the β-carbonic anhydrases *βca1* and *βca4* show impaired $CO_2$-regulation of stomatal movements and increased stomatal density, but retain functional abscisic-acid and blue-light responses. |

| | | |
|---|---|---|
| Beta Carbonic Anhydrase 4 (βCA4) (Hu, Boisson-Dernier *et al.* 2010) | Encodes a putative beta-carbonic anhydrase BCA4. Together with BCA1 regulates $CO_2$-controlled stomatal movements in guard cells. | Double-mutant plants in the β-carbonic anhydrases *βca1* and *βca4* show impaired $CO_2$-regulation of stomatal movements and increased stomatal density, but retain functional abscisic-acid and blue-light responses. |
| STOMAGEN/EPFL9 (Sugano, Shimada *et al*. 2010) | Encodes a cysteine-rich peptide, a secretory factor that is produced in the mesophyll cells and acts on the epidermis to increase stomatal formation. Its mature form is a 45-aa peptide with three intramolecular disulfide bonds. It is proposed that STOMAGEN increases stomatal number by competing with two negative regulators of stomatal density, EPF1 and EPF2, possibly through direct interaction with the receptor-like protein TMM. | Reduced stomatal density. |
| AGO1(Yang, Jiang *et al.* 2014) | Encodes an RNA Slicer that selectively recruits microRNAs and siRNAs. Mutants are defective in post-transcriptional gene silencing and have pleiotropic developmental and morphological defects. Through its action on the regulation of ARF17 expression, the protein regulates genes involved at the cross talk between auxin and light signaling during adventitious root development. | Increased stomatal density |
| Speechless (SPCH) (MacAlister, Ohashi-Ito *et al.* 2007) | Encodes a bHLH transcription factor that is necessary and sufficient for the asymmetric divisions that establish the stomatal lineage in *Arabidopsis thaliana*. It has been demonstrated that *SPCH* and two paralogues are successively required for the initiation, proliferation and terminal differentiation of cells in the stomatal lineage. | No stomata |
| HIC (Gray, Holroyd *et al.* 2000) | Encodes a 3-ketoacyl coenzyme A (CoA) synthase which has been shown to be involved in the decrease of stomatal numbers under elevated $CO_2$. | Increase in stomata in response to elevated $CO_2$ |

Table 1.3 – Previously documented stomatal patterning genes, with a functional description, and the effect of the null mutant.

These stomatal patterning genes act at different, and sometimes multiple, points during stomatal development. This can be seen in Figure 1.10.

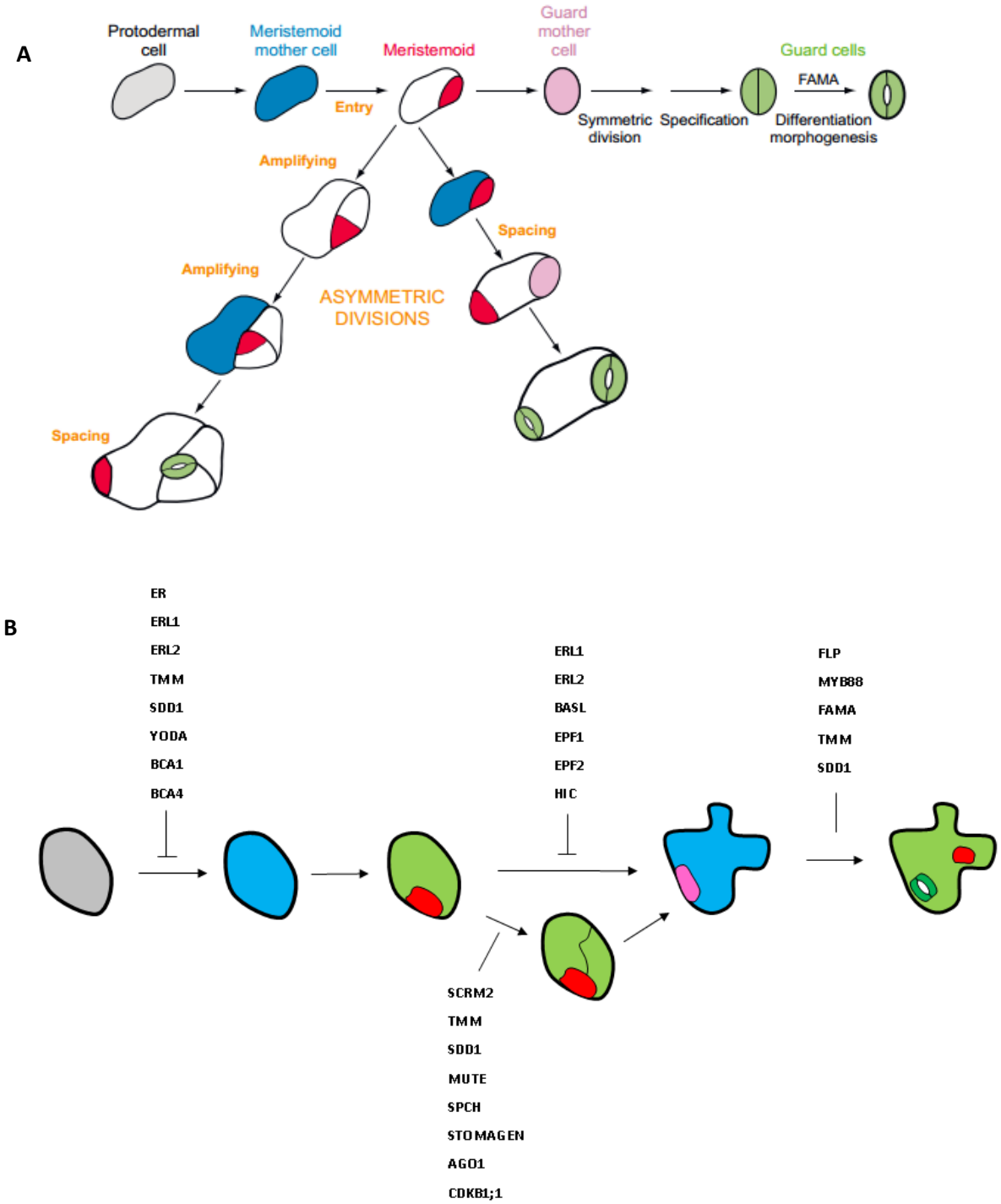



Figure 1.10 – **A** –The pathway from a protodermal cell to stomatal development taken from (Bergmann and Sack 2007) **B** – The position at which stomatal patterning genes act during the stomatal development pathway. Negative regulation is indicated by T shaped lines; positive regulation is indicated by a I shaped line.

Not only do genes act at multiple stages, but some genes have both a negative and positive regulation depending on which stage in development they are acting. Stomatal patterning genes are not independent of each other, and rather interact to produce different phenotypes. Torii (2012) showed how many of the stomatal patterning genes interact with other, through regulations, activations or gene rescue (Figure 1.11).

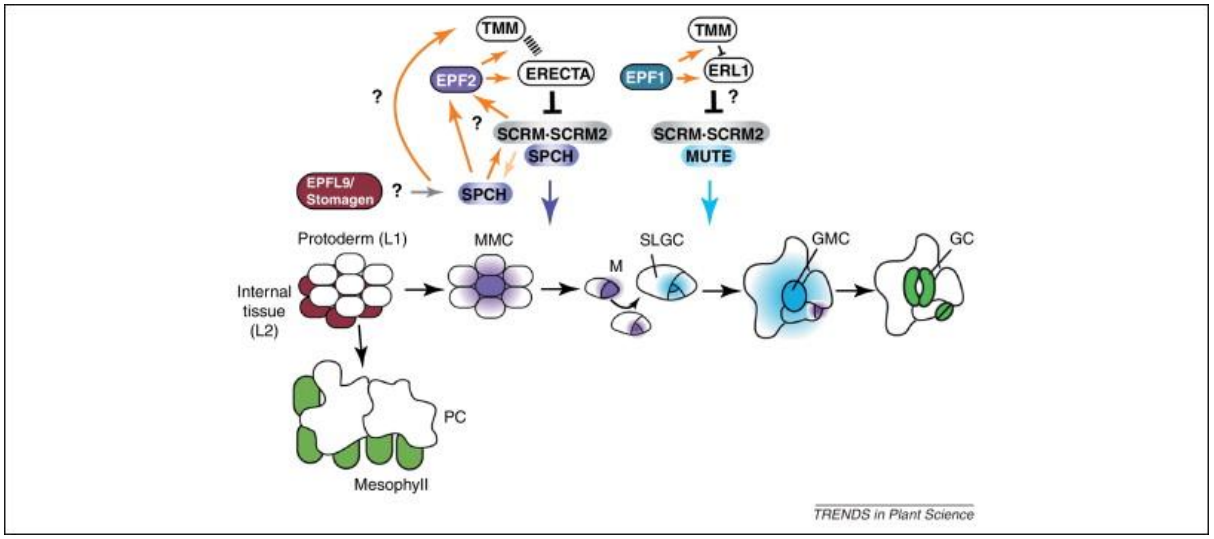



Figure 1.11 – The stomatal patterning pathway with the stomatal patterning genes that interact with each other (Torii 2012). The colours in the pathway relate the colours of the genes. Arrows and T-bars indicate positive and negative genetic interactions, respectively, and do not necessarily imply direct association of gene products. Broken bar indicates a bifunctional relation. For stomatal patterning gene abbreviation see Table 1.3. PC – Protodermal Cell, MMC – Meristemoid Mother Cell, M – Meristemoid, SLGC – Stomatal lineage Guard Cell, GMC – Guard Mother cell, GC – Guard Cell.

Torri (2012) showed in *A. thaliana*, SPCH and SCRM activate EPF2 expression directly or indirectly, and the EPF2 signalling is mediated through ERECTA and is likely to be modulated by TMM, which in turn is likely to inhibit SPCH/SCRMs. TMM is not expressed in *spch*, suggesting SPCH is genetically upstream of TMM. The meristamoid mother cell (MMC) initiates asymmetric entry division and produces a meristemoid, which reiterates asymmetric amplifying divisions. The meristemoids express and secrete EPF1, which orients asymmetric spacing division to enforce the one-cell spacing rule. EPF1 acts upstream of both the ERECTA-family and TMM, suggesting EPF1 signalling is mediated via ERL1 only. They also suggested it is likely that TMM acts in an antagonistic manner to ERL1. SCRM and SCRM2 have also both been shown to specify the sequential actions of SPCH, MUTE, and FAMA, allowing protodermal cells to initiate and execute stomatal cell fate (Kanaoka, Pillitteri *et al.* 2008).This highlights how linked the stomatal

patterning genes are, and to look at expression of a single gene would over look potential interesting interactions, and so the pathway as a whole needs to be analysed.

Other factors have been discovered to effect stomatal patterning, including the hormone auxin (Le, Liu *et al.* 2014). Le *et al.* (2014) suggested auxin depletion is a trigger for a switch from unequal to equal division during stomatal development. Unequal division occurs in the meristemoid cells before they acquire a guard mother cell (GMC), from which equal division occurs. This switch mediates the amount of spacing between stomata. The sister cells of the GMC retain a high auxin level where as the GMC's auxin levels become depleted. It is suggested that this change in auxin concentration was used to determine spacing in stomatal patterning, and so looked at mutants of the pin-formed (PIN) protein mediated transport system. They found no single mutants showed loss of function, but two different quadruple mutants, *pin2, 3, 4, 7* and *pin1, 3, 4, 7* displayed pattern defects in the form of groups of stomata in direct contact (around 20% stomata in clusters), indicating functional redundancy between PINs, which is consistent with the expression of multiple PINs in the epidermis. These mutants show that differences in auxin concentrations play a role in mediating cell spacing in the stomatal patterning pathway.

Spacing between stomata is important to the function of the plant, and stomata are consistently separated by at least one epidermal cell (the 'one-cell-spacing rule'), despite differences in stomatal density between species (Sachs 2005). This spacing allows for proper opening and closure of the stomatal aperture, ensuring that guard cells can efficiently exchange ions and water to regulate pore opening, and reduce unnecessary evaporation (Pillitteri and Dong 2013). This has been proven by Dow *et al.* (2014), who showed genotypes with correct spacing (< 5% of stomata in clusters) achieved diffusive stomatal conductance values comparable to anatomical stomatal conductance across a 10-fold increase in stomatal density, while lines with patterning defects (> 19% clustering) did not (Dow, Berry *et al.* 2014). They also showed that genotypes with clustering had reduced assimilation and impaired stomatal responses.

### 1.6.4 Stomatal Opening and Closing

A plethora of environmental factors are known to influence the opening and closing of stomata, including atmospheric $CO_2$. Although the number and patterning of stomata is important, it is well known that how they function determines a large portion of their efficiency (Mishra, Zhang *et al.* 2006). The aperture of the stomata ultimately determines the stomatal conductance of a plant independently of the number of stomata, and is an effective way to control gas exchange.

The general opening and closing mechanism of the stomata is well understood, and is controlled by the movement of ions to create solute potentials to increase/decrease turgor of the guard cells. To open the stomata $H^+$-ATPases are activated in the plasma membrane, causing membrane hyperpolarization. This activates a $K^+$ ion pump which pumps $K^+$ ions into the vacuole of the guard cells (Assmann, Simoncini *et al.* 1985, Kwak, Murata *et al.* 2001, Lebaudy, Véry *et al.* 2007). This reduces the solute concentration inside the vacuole of the guard cells causing water to enter, increasing the turgor, and opening the stomata. To close stomata $Ca^{2+}$-permeable channels in the plasma membrane of guard cells are activated (Hamilton, Hills et al. 2000, Gobert, Isayenkov et al. 2007). $Ca^{2+}$ in turn activates outward $K^+$ channels in the vacuole (VK channels), which releases the $K^+$ anions from inside the vacuole (Gobert, Isayenkov *et al.* 2007). The $Ca^{2+}$ also activates two different types of anion channels, slow-activating sustained (S-type) and rapid-transient (R-type) anion channels, which cause an anion efflux causing membrane depolarization (Schroeder and Hagiwara 1989, Linder and Raschke 1992). This subsequently drives $K^+$ efflux from guard cells through outward-rectifying $K^+$ channels (Hosy, Vavasseur *et al.* 2003), decreasing solute potential, leading to a decrease in turgor and stomatal closure. A diagram illustrating the mechanism of stomatal opening and closing can be seen in Figure 1.12. The influx and efflux of ions is also associated with the conversion of malate into starch, and vice-versa to open and close stomata (MacRobbie 1998). The production of malate from osmotically inactive starch helps increase turgor and volume in the guard cell, inducing stomatal opening, whereas gluconeogenic conversion of malate back into starch works the opposite way, decreasing turgor and causing stomatal closure.

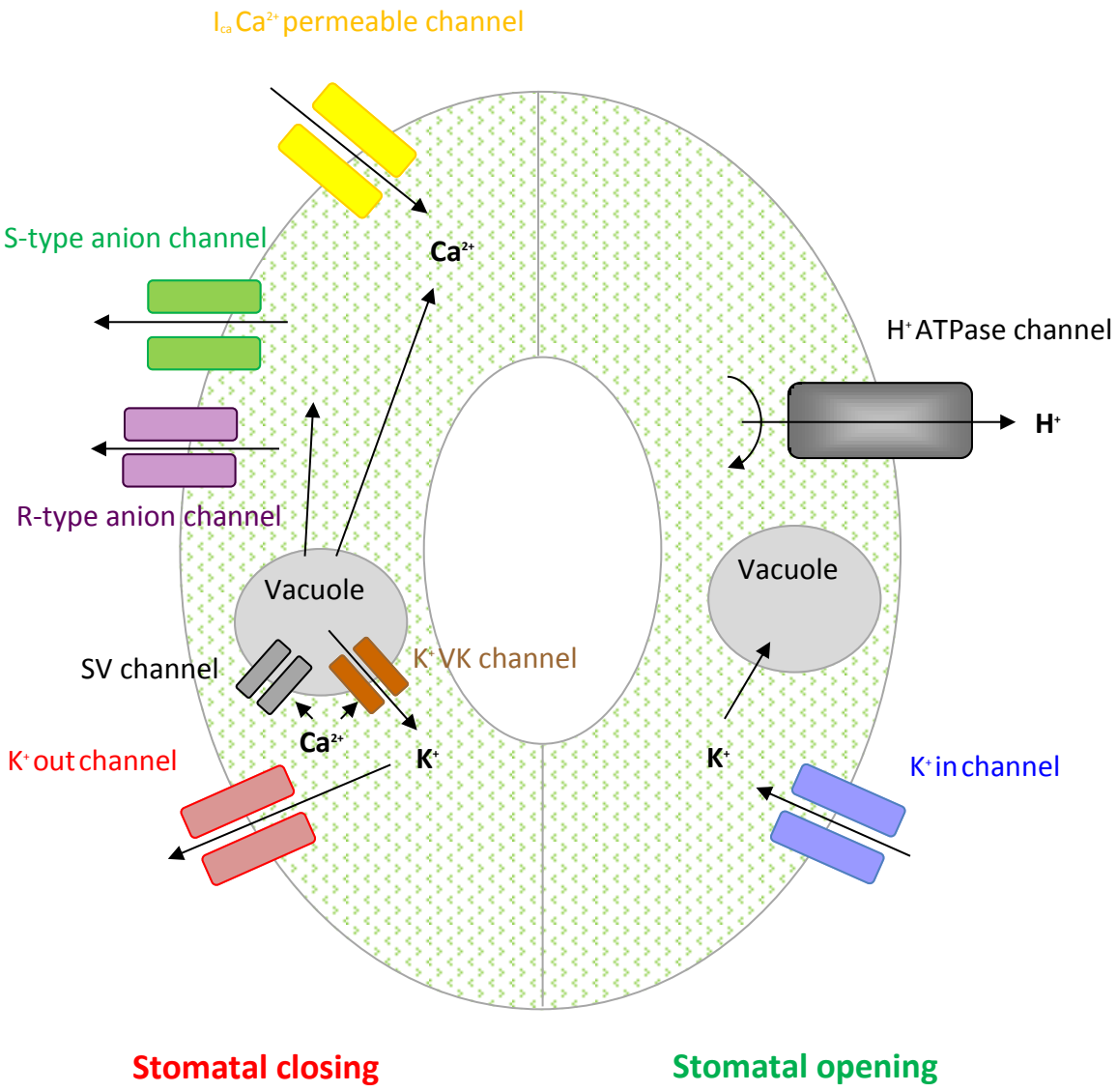



Figure 1.12 - Summary of guard cell signalling and ion channel regulation. Signalling events during stomatal closing are shown in the left guard cell, and stomatal opening mechanisms are shown in the right guard cell.

There are multiple potential messengers in the stomatal $CO_2$ concentration response, including cytosolic free calcium concentration ($[Ca^{2+}]$), apoplastic and cytoplasmic pH gradients, ion channels and membrane potential, chloroplastic zeaxanthin levels, photosynthetically derived ATP and protein phosphorylation/dephosphorylation (Assmann 1999, Hetherington and Woodward 2003, Young, Mehta *et al.* 2006). Many of these sensing systems overlap leading to the hypothesis - multiple sensing is likely in response to $CO_2$. It was thought the sensing mechanism for $CO_2$ was contained within the mesophyll, but there is also evidence that plants respond independently of the mesophyll (Morison 1998). Under elevated $CO_2$ concentrations the intracellular $CO_2$ concentration ($C_i$) has been shown to be unchanged when compared to ambient $CO_2$ (Tuba,

Szente *et al.* 1994, Xu 1994), even though there was a change in stomatal conductance suggesting the differences in $CO_2$ was not sensed in the mesophyll. Analyses of epidermal strips removed from the mesophyll cell environment have also been carried out, and it was shown that $CO_2$ mediated stomatal pore closure even in the epidermal strips (Willmer and Pallas 1974, Lee, Choi *et al.* 2008).

The use of epidermal strips led to the hypothesis that there may be a $CO_2$ sensing mechanisms within the guard cells (Assmann 1999). Assmann (1999) showed the potential of the $CO_2$ sensing mechanisms such as cytosolic pH and malate levels, cytosolic $Ca^{2+}$ levels, chloroplastic zeaxanthin levels, or plasma-membrane anion channel regulation being present in guard cells. Hu et al. (2010) showed the $CO_2$-binding carbonic anhydrase proteins that catalyse the reversible reaction, $CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$ might function early in $CO_2$ signalling due the involvement of $CO_2$-/$HCO_3^-$-binding proteins RuBisCO (ribulose-1,5-bisphosphate carboxylase oxygenase) and PEPC (phosphoenolpyruvate carboxylase) in the photosynthesis pathway. Transcriptome analyses of mesophyll and guard cells showed that the β-carbonic anhydrase genes βCA1, βCA4, and βCA6 are highly expressed in guard cells and mesophyll cells (Hu, Boisson-Dernier et al. 2010). Hu et al. (2010) went on to show *ca1 ca4* double and *ca1 ca4 ca6* triple mutants in *A. thaliana* showed strong insensitivities to $CO_2$-induced stomatal conductance changes. However, *ca1 ca6* and *ca4 ca6* mutants did not show an altered $CO_2$response, indicating no major role for the more distantly related βCA6. It has been shown that $CO_2$-regulated stomatal conductance is independent of RuBisCO activity (Allen, Kuchitsu *et al.* 1999), and that PEPC levels have no direct effect on high $CO_2$ concentration-triggered stomatal closure (Catala, Ouyang *et al.* 2007), and so the mechanism these genes employ to sense $CO_2$ is not to do with these proteins. As no other $CO_2$-binding proteins have been identified in genetic screens of $CO_2$-regulated stomatal signalling, Hu *et al.* (2010) postulated that $CO_2$-binding proteins that mediate this response may be encoded by a gene family with overlapping gene functions, explaining why differences in RuBisCO and PEPC do not affect carbonic-anhydrase regulation of $CO_2$ sensing. This explanation both explains how $CO_2$ is sensed in the mesophyll, but also provides an explanation when no change in $C_i$ was present but stomatal conductance changed, as the guard cells also contain the genes to sense the $CO_2$ causing the plant to adapt/acclimate.

It is not only elevated $CO_2$ that can cause stomatal closing; drought also causes stomatal closure through abscisic acid (ABA) to retain as much water as possible. This response occurs using ABA to activate the same mechanisms as the response to elevated $CO_2$ (see Figure 1.14), and can also use ABA to inhibit stomatal opening in cases of severe drought. ABA releases more $Ca^{2+}$ which causes

depolarisation favouring $K^+$ efflux, instead of intake and so causing the stomata to remain closed (Schwartz, Wu *et al.* 1994). It has been shown that the $CO_2$ and ABA mechanisms of stomatal closure can interact with each other and that the pathways overlap (Young, Mehta *et al.* 2006), and so it is important to differentiate between the two pathways where possible, so genes can be used to identify the response causing the stomata to close. *Arabidopsis* mutants have been used to identify parts of the ABA or $CO_2$ mediated pathways which are unique to either pathway. βCA1, βCA4, and βCA6 as mentioned before, are unique to the sensing of $CO_2$ and therefore the $CO_2$ response pathway (Hu, Boisson-Dernier *et al.* 2010). Other genes which have been identified include the HT1 (High Leaf Temperature 1) protein kinase (Hashimoto, Negi *et al.* 2006), which is the first molecular component that has been identified as a major negative regulator of the $CO_2$ induced stomatal closure pathway. Stomatal responses to $CO_2$ changes in the leaf epidermis and in intact leaf gas-exchange analyses showed that the recessive *ht1–2* mutation caused a constitutive elevated $CO_2$ stomatal closure. Although HT1 protein kinase activity is greatly reduced in *ht1–1* and *ht1–2* mutants, they still retained responsiveness to ABA and blue light, indicating that HT1 may function upstream of the convergence of the $CO_2$ and ABA induced stomatal closure pathways (Hashimoto, Negi *et al.* 2006). Other genes which have been identified have not come before the convergence of the $CO_2$ and ABA, but are still important in the overall pathway of stomatal aperture. ABCB14 (ABC Transporter B Family Member 14) is a plasma membrane ABC malate uptake transporter, and a negative regulator of stomatal closure (Lee, Choi *et al.* 2008). $CO_2$ induced stomatal closure in detached leaves was slightly accelerated in *abcb14* mutants and decreased in *abdb14* overexpressing plants (Lee, Choi *et al.* 2008), suggesting that malate uptake into guard cells by AtABCB14 plays a role in the $CO_2$-induced regulation of stomatal closure, although this happens downstream of the convergence of ABA in the pathway.

Positive regulators of the $CO_2$ induced pathway have also been identified. The ABA-insensitive mutant *gca2* (growth controlled by abscisic acid 2) is strongly impaired in $CO_2$ induced stomatal closure in response to elevated $CO_2$ (Himmelbach 1998, Young, Mehta *et al.* 2006). Changes in $CO_2$ concentrations did not elicit significant changes in cytosolic $Ca^{2+}$ transient rate in *gca2* mutant guard cells, indicating an impairment in $CO_2$ induced depolarization of the membrane potential (Young, Mehta *et al.* 2006). Other research has also shown that *gca2* mutant plants are impaired in ABA-induced stomatal closure (Allen, Chu *et al.* 2001), indicating that GCA2 likely functions downstream of the convergence point of $CO_2$ and ABA signalling. SLAC1 (Slow Anion Channel-Associated 1) was recently identified as another $CO_2$ insensitive mutant in *Arabidopsis*, with *slac1* mutant plants having highly impaired stomatal closure response under elevated $CO_2$ (Negi,

Matsuda *et al.* 2008). This protein interacts with the S-type anion channel previously mentioned, and mutants impair activation of this anion channel. The anion channels are activated downstream of any ABA interaction, and so this gene functions post convergence of the $CO_2$ and ABA pathways. The R-type anion channel activity and ABA activated $Ca^{2+}$ channel activity are retained in *slac1* mutants, providing genetic evidence that the S-type anion channel functions as a central control mechanism for ABA and $CO_2$ induced stomatal closure (Vahisalu, Kollist *et al.* 2008). A summary of these genes has since been provided, along with a hypothesise where they interact in the $CO_2$ induced stomatal closure pathway (Kim, Böhmer *et al.* 2010), along with indication where convergence with the ABA induced pathway occurs (Figure 1.13).

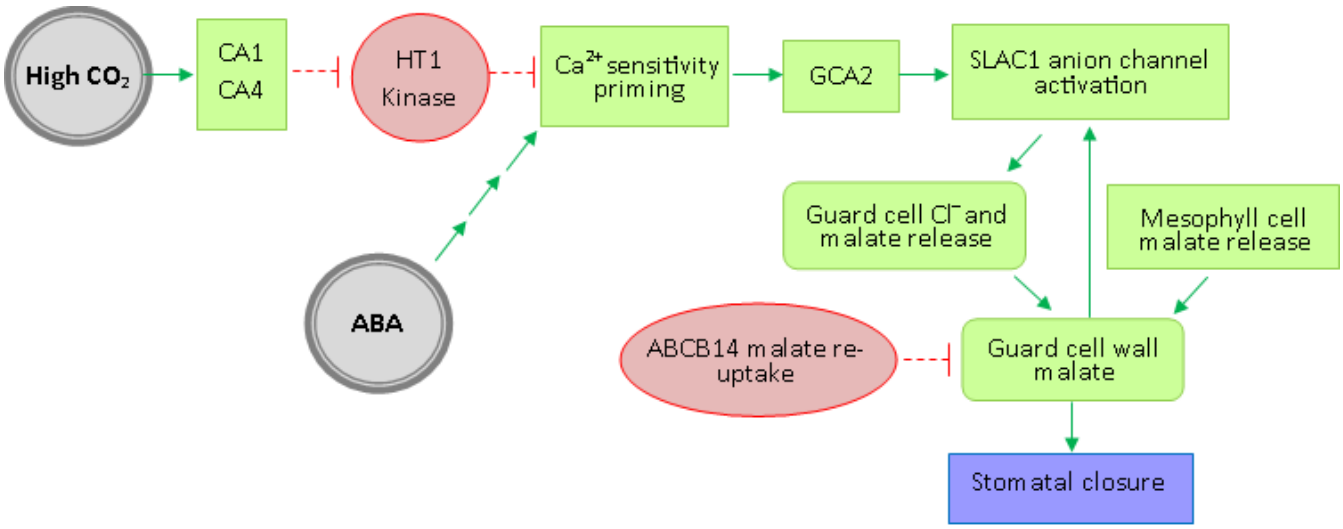



Figure 1.13 - A simplified model illustrating the functions of recently identified genes and mechanisms in guard cells mediating $CO_2$ control of stomatal movements. In this model, the HT1 protein kinase and ABCB14 proteins function as negative regulators (*red*), and CA1 and CA4, GCA2, and SLAC1 function as positive mediators (*green*) of high $CO_2$-induced stomatal closing. Convergence with abscisic acid (ABA) signalling is also indicated (adapted from Kim *et. al.* (2010)).

Although ABA is the key hormone involved in stomatal aperture, other hormones have been shown to have an interaction with stomatal opening and closing (Acharya and Assmann 2009). As $CO_2$ acts early in the pathway it is important to note changes in these other factors which can affect stomatal aperture, as change in $CO_2$ concentration could be causing these effects. Auxins are a class of plant hormones consisting of indole-3-acetic acid (IAA), and are generally involved in processes such as promoting cell division, cell elongation, vascular tissue differentiation and are essential components in tropic responses (Kepinski 2007). Auxin typically plays a positive role in stomatal opening, with low concentrations promoting the inward activity of $K^+$ during stomatal

opening, although higher concentrations of auxin have been shown to inhibit stomatal opening by promoting activity of outward $K^+$ channels (Lohse and Hedrich 1992). Antagonistic stomatal regulation has also been observed between ABA and auxin. Auxin was shown to repress stomatal closure in response to ABA in epidermal peels of *Commelina communis* (Snaith and Mansfield 1982) and a similar oppositional role of auxins was observed when 1-Naphthaleneacetic acid (NAA) was applied in combination with ABA to opened stomata in epidermal peels of *Arabidopsis* (Tanaka, Sano *et al.* 2006).

Although gibberellins (GAs) have a common interaction with ABA in seed germination, they have been shown to not have a major role in stomatal aperture, although it has been shown to be able to cause transient stomatal opening (Göring, Koshuchowa *et al.* 1990) and promote stomatal opening in isolated epidermal strips showing it can have a role (Santakumari and Fletcher 1987). The gaseous plant hormone ethylene has been shown to have contrasting effects to stomatal aperture, when acting alone the hormone stimulates stomatal closure (Desikan, Last *et al.* 2006), but when in cohort with other hormones (such as ABA) it opposes stomatal closure via those hormones (Tanaka, Sano *et al.* 2006). Little is understood about the mechanism ethylene uses to control stomatal aperture under these multi-hormone pathways due to the complex nature of how they must interact.

Cytokinins are adenine-derivatives which have positive roles in germination, root and shoot development, as well as having a positive role in stomatal opening. Both synthetic and natural cytokinins have been shown to cause stomatal opening (Jewer and Incoll 1980), and inhibition of closure has been shown in the over producing cytokinin mutant *amp1-1* (Tanaka, Sano *et al.* 2006). Recently it has been shown that in darkness, cytokinin induces stomatal opening by decreasing $H_2O_2$ and NO levels within guard cells (Xiao-Ping and Xi-Gui 2006), and both these ROS compounds have been shown to have a role in the stomatal closure pathway (Zhang, Zhang *et al.* 2001).

Brassinosteroids are growth-promoting polyhydoxylated steroidal plant hormones which have been shown to have a regulation in stomatal aperture in similar ways to ABA in response to water stress, and appears to cross-talk with ABA in stomatal regulation to drought (Xu, Shida *et al.* 1994). Brassinosteroids have been shown to not be vital in plants response to water stress (Jager, Symons *et al.* 2008), but they have been shown to induce water stress tolerance such as stomatal closure in plant species (Rajasekaran and Blake 1999). Another hormone which has been linked to stomatal closure due to drought is the lipid-derived plant hormones Jasmonates (JAs). It has been proposed that JAs could be an important player for stomatal closure during drought stress based on its accumulation during drought (Creelman and Mullet 1995) and its positive regulatory role in

stomatal closure (Munemasa, Oda *et al.* 2007). JA levels also increase more than 50-fold in response to pathogens introduction (Penninckx, Eggermont *et al.* 1996), but the stomatal physiology response of pathogen induced JA accumulation is not known, although it could indicate a response to stomatal closure in the presence of pathogens. The hormone that is more known for its role in plant defence is the phenolic compound salicylic acid (SCA), which has been shown to play a positive role in stomatal closure. Stomatal closure in response to bacterial pathogens is compromised in transgenic NahG plants (deficient in SCA) and in the SAC biosynthetic mutant eds16-2, indicating that SCA is required for stomatal defence (Melotto, Underwood et al. 2008). It is proposed that SCA and ABA together mediate stomatal closure in response to bacterial pathogens (Melotto, Underwood *et al.* 2006).

Acharya and Assmann (2009) provided an illustrative overview of how all of these hormones interact with ABA during stomatal opening and closing (Figure 1.14).

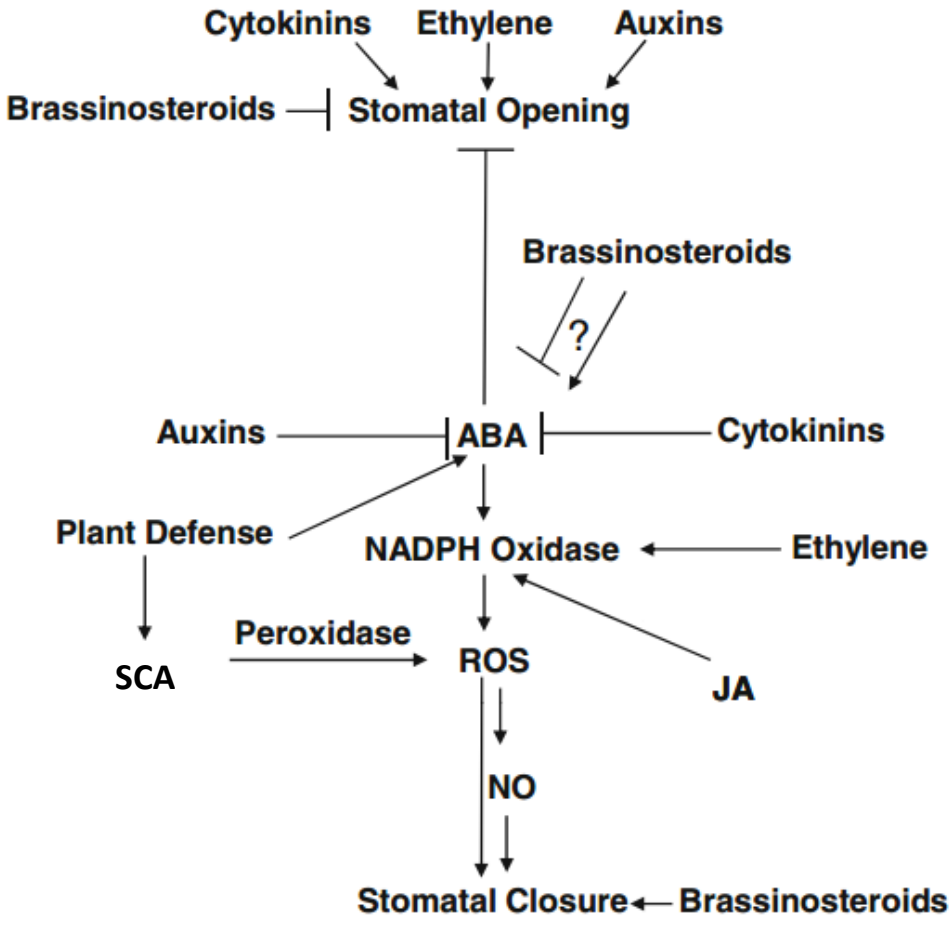



Figure 1.14 - A proposed model of hormonal interaction in stomatal regulation (Acharya and Assmann 2009).

Even though these stomatal responses aim to provide advantageous conditions for the plant, such as increased water use efficiency, it has been found that negative secondary effects can occur. Reduced stomatal conductance has been shown to increase leaf temperature, which has been predicted to contribute to heat stress in plants which can reduce crop yield (Cardon, Berry *et al.* 1994, Battisti and Naylor 2009), and plants with higher stomatal conductance and lower leaf temperatures obtain higher yields (Lu, Radin *et al.* 1994). This highlights the importance of understanding these stomatal response mechanisms if we hope to sustain yields under future high $CO_2$ concentrations.

## 1.7 Epigenetics

Epigenetics is one of the most rapidly expanding fields in life sciences, but is still not completely understood, leaving the aspects which define epigenetics blurred (Meloni and Testa 2014). This being said, it is still a field which requires consideration when determining conclusions, even if work does not directly involve it. Epigenetics involves genetic factors other than changes in DNA sequence, such as DNA methylation, histone modification and RNA associated silencing (Bird 2007). Studies have shown that these DNA and histone modifications play a key role in gene expression and plant development under stress (Thomashow 1999, Iba 2002). Most of these stress-induced modifications are reset to the basal level once the stress is relieved, while some of the modifications have been identified as stable and are carried forward as 'stress memory' (Chinnusamy and Zhu 2009). This stress memory can be inherited across mitotic or even meiotic cell divisions (Bonasio, Tu *et al.* 2010), and may help plants more effectively cope with subsequent stresses.

DNA methylation occurs at the cytosine bases of eukaryotic DNA, which are converted to 5-methylcytosine by DNA methyltransferase (DNMT) enzymes. How this effects gene expression is not well understood but it is suggested that methylation occurs in the promoter regions of genes (Suzuki and Bird 2008). This modification alters the recognition of the DNA, leading to transcriptional silencing of any genes encoded in the methylated section of DNA (Kass, Pruss *et al.* 1997). There are different members of the DNMT family which have two distinct roles (Phillips 2008). The first are known as '*de novo*' DNMT's, which complete the initial methylation pattern onto the DNA. The second are maintenance DNMT's which copy the existing methylation to new strands after replication. Currently the mechanisms in which the *de novo* DNMT's select sites are again not well understood and so it is difficult to anticipate where methylation will occur, but as plants are the most highly methylated eukaryotes (Phillips 2008), with up to 50% of their cytosine residues exhibiting methylation, it has to be considered.

Histones are highly alkaline proteins found in eukaryotic cells which are involved in packaging DNA into structural units called nucleosomes, allowing the DNA to form chromosomes within the nucleus (Peterson and Laniel 2004). Methylation can also occur to histones (both lysine and arginine methylation), amongst several other modifications including lysine acetylation, serine and threonine phosphorylation, and lysine ubiquitination and sumoylation (Vaquero, Loyola et al. 2003, Bannister and Kouzarides 2011). Histone modifications can exert their effects via two main mechanisms (Bannister and Kouzarides 2011). The first involves the modifications that directly influence the overall structure of chromatin, which can be either over short or long distances. This

is achieved via histone acetylation and phosphorylation which effectively reduce the positive charge of histones, disrupting the electrostatic interactions between histones and DNA leading to a less compact chromatin structure. This can facilitate DNA access by protein machineries involved in transcription to potentially increase expression. Multiple histone acetylation's are also enriched at enhancer elements and particularly in gene promoters, where they presumably again facilitate the transcription factor access (Wang, Zang *et al.* 2008). Ubiquitination adds a large molecule to a histone which is highly likely to induce a change in the overall conformation of the nucleosome, which in turn will affect intra-nucleosomal interactions and/or interactions with other chromatin-bound complexes. The second involves the modification regulating (either positively or negatively) the binding of effector molecules (Sun and Allis 2002). Numerous chromatin associated factors have been shown to interact with modified histones via many distinct domains including plant homeodomain (PHD) fingers (Champagne and Kutateladze 2009) and the so called Tudor 'royal' family of domains (Maurer-Stroh, Dickens *et al.* 2003). These modifications not only function by providing dynamic binding platforms for various factors potentially increasing gene expression, but can also disrupt an interaction between the histone and a chromatin factor disrupting binding and repressing genes.

RNA associated silencing, otherwise known as RNA interference (RNAi) is a group of epigenetic phenomena directed by small RNAs which cause post-transcriptional repression (Castel and Martienssen 2013). There are two types of RNAi that affect plants, these are microRNAs (miRNAs), which are hairpin-derived RNAs with imperfect complementarity to targets and that cause translational repression, and small interfering RNAs (siRNAs), which have perfect complementarity to targets and cause transcript degradation. miRNAs have recently been noted to respond to elevated $CO_2$ concentrations (May, Liao *et al.* 2013), making them very interesting in terms of this thesis. May *et al.* (2013) found that doubling $CO_2$ concentration significantly altered the expression of miRNAs, affecting *Arabidopsis* growth and development. Using sequencing, the authors also identified two miRNAs (miR156/157 and miR172) which they suggest could be responsible for the regulation of transcriptional networks under elevated $CO_2$ for the onset of early flowering. This highlights the importance of identifying any potential miRNAs of genes of interest when possible, to comprehensively explain differences in gene expression.

## 1.8 Using Natural Springs to Study Plant Acclimation and Adaptation to Elevated Atmospheric $CO_2$

Although studies using controlled environments, such as open top chambers and FACE, provide a good insight into short term acclimation, they provide only limited information about the multi-generational effects of elevated $CO_2$. Few studies published to date have considered the trans-generational effects of elevated $CO_2$, or taken the unique advantage of using naturally emitting $CO_2$ springs to study species that have endured elevated $CO_2$ over many generations. A number of studies have highlighted the importance of the data these springs provide (Cook, Oechel *et al.* 1997, Grace, Gardingen *et al.* 1997, Woodward, Beerling *et al.* 1997). In particular, Woodward *et al.* (1997) explored the potential of $CO_2$ controlled selection in higher plants due to the variations of $CO_2$ concentrations over millions of years, and using the natural springs to give insight and information about potential selection. A range of phenotypic and physiological measurements have been made in relation to plant responses to elevated $CO_2$ using natural springs, including:

- Stomatal numbers (Paoletti, A. Raschi *et al.* 1997)
- Photosynthesis (Jones, Brown *et al.* 1995)
- Water relations (Tognetti, Longobucco *et al.* 1999)
- Growth (Miglietta, Raschi *et al.* 1993b)
- Isoprene emission (Scholefield, Doick *et al.* 2004)
- Photoinhibition (Stylinski, Oechel *et al.* 2000)
- Non-structural carbohydrates (Körner and Miglietta 1994)

The location which laid the foundations for this exciting area of research is a natural spring in central Italy, in a region called Bossoleto (Figure 1.15; Lat. 43˚17', Long. 11˚35'). This site is also the location which will be used during this study. Infra-red gas analysis (IRGA) has been used at the site for four months, to provide exact data about the $CO_2$ concentration of the spring (Bettarini, Grifoni *et al.* 1999). The results showed a marked enrichment of $CO_2$. During the day the average concentration ranged between 600 and 1200 µmol $mol^{-1}$, while un-stable nocturnal conditions were often over 7500 µmol $mol^{-1}$ (Bettarini, Grifoni *et al.* 1999).

Figure 1.15 –The Bossoleto spring site. Photo taken at 10:00am, before the $CO_2$ boundary layer had completely dissipated.

The characteristics of the site have been well documented (Van Gardingen, Grace *et al.* 1997, Bettarini, Grifoni *et al.* 1999). The origin of the $CO_2$ comes from several vents around the spring, but emissions are dominated by a large vent, in a cave at the bottom of the bowl of the spring. The $CO_2$ from the vents is 99% pure, and has a very low levels of other pollutants, such as $H_2S$ and $SO_2$, often found to be undetectable (Van Gardingen, Grace *et al.* 1997). The levels of these pollutants can fluctuate higher with weather conditions, and have been shown to have the potential to be considerably high (up to 0.8 µmol $mol^{-1}$) (Schulte, Raiesi *et al.* 1997). A study was carried out using $H_2S$-sensitive soybean using another spring, with magnitudes higher $H_2S$ concentrations compared to the Bossoleto spring, to ensure this would not cause an effect. It was shown that there was no evidence of a direct toxic effect of the sulphur gases (Miglietta, Raschi *et al.* 1993). Although no toxic effect was shown, later Grill *et al.* (2004) showed the levels of sulphur gases at the Bossoleto spring were having a physiological effect. Acorns collected from the spring of *Quercus ilex* and *Quercus Punescens* showed significantly higher concentrations of sulphur and glutathione (a compound synthesised from $H_2S$ and $SO_2$) compared to a nearby control site (Grill, Müller *et al*. 2004). This research was furthered by Herschbach *et al.* (2012), who agreed the

sulphur gas emissions, could pose a physiological effect. Again using *Q. ilex* and *Q. pubescens* samples collected from three different springs (including the Bossoleto spring); the activity of enzymes associated with sulphate assimilations was measured. Decreased activity of sulphur assimilation enzymes were found, coupled with increased concentrations of glutathione, in plants originating from the spring compared to a nearby control (Herschbach, Schulte *et al.* 2012). The levels of glutathione were comparably lowest at the Bossoleto spring, but it was still shown that the sulphur gases were causing physiological effects, which could override consequences of $CO_2$ concentration. The effects of particularly $H_2S$, can range from causing extinction when exposed to large amounts (30-100 µmol $mol^{-1}$) (Knoll, Bambach *et al.* 2007), to significantly increasing biomass and fruit yield when exposed to extremely low concentrations (0.01 µmol $mol^{-1}$ or less), compared to plants exposed to no $H_2S$ (Dooley, Nair *et al.* 2013). $H_2S$ has also emerged as a potential messenger molecule involved in modulation of numerous physiological processes, including photosynthesis via promoting chloroplast biogenesis (Chen, Wu *et al.* 2011). The product of increased $H_2S$ and $SO_2$ concentrations, glutathione, is also a compound which can reduce oxidative stress (Ding, Lu *et al.* 2009). These responses are similar to responses which we could expect to see under elevated $CO_2$, such as increased growth and modification of the photosynthesis pathway, leading to potential difficulty in differentiating between the two responses.

Natural springs in Japan have been shown to contain below detection levels of $H_2S$ (<0.1 µmol $mol^{-1}$) (Onoda, Hirose *et al.* 2007), suggesting theses may provide better experimental sites, but unfortunately logistics dictate that these sites are harder to obtain samples from. Also even though levels are below detection, they could still be having a physiological effect within the Japanese springs, as concentrations lower than 0.01 µmol $mol^{-1}$ have still been shown to have a physiological effect (Dooley, Nair *et al.* 2013).Using the spring with the lowest concentrations of sulphur gases of those available is the best way to attempt to alleviate any potential effects as much as possible. Out of the Italian springs this is the Bossoleto site (Schulte, Raiesi *et al.* 1997), and is why this spring site is most often used, but the effects still have to be considered when drawing conclusions.

A refractive boundary layer within the Bossoleto spring can be seen in the mornings after $CO_2$ concentrations have been at their highest, and forms towards the bottom of the spring (approximately 2 metres from the bottom). This forms as a result of the optical densities between a higher (more dense) concentration of $CO_2$ at the bottom of the spring, and a lower (less dense) concentration of $CO_2$ above. This creates a mini greenhouse gas effect within the spring. As the

day proceeds, this layer becomes unstable due to lower $CO_2$ concentrations, and direct solar radiation. This layer creates both a $CO_2$, and temperature gradient, originating from the bottom of the spring. Temperatures can decrease as much as 20˚C from the bottom compared to the top of the spring (Van Gardingen, Grace *et al.* 1997). $CO_2$ decreased by more than a 3000 µmol $mol^{-1}$ when comparing the bottom to the top of the spring, but interestingly in certain conditions (low wind), these differences were negated (Van Gardingen, Grace *et al.* 1997). The water table is close to the surface of the soil, and plant roots are in contact with a supersaturated salt solution when temperatures reach 38˚C. The water consists of mainly cations $Ca^{2+}$, $Mg^{2+}$, $Na^{+}$ and the anions $HCO^{-}_{3}$, $SO^{2-}_{4}$ and $Cl^{-}$ (Van Gardingen, Grace *et al.* 1997). The solid residue of these ions can reach over 5 g $1^{-1}$.

These factors make it difficult to determine which factors are causative of certain phenotypes occurring in springs, as they are not only uncontrollable, but also very variable within themselves. This being said, the Bossoleto is the most used site for multi-generational $CO_2$ studies in Europe, due to a lower degree of environmental factors compared to other sites in Europe (Miglietta, Raschi *et al.* 1993). Studies which have been carried out using the Bossoleto spring are summarised in Table 1.4.

| *Study* | *Species* | *Study Focus* | *Findings* |
|---|---|---|---|
| Tognetti, Longobucco *et al.* (1999) | *Quercus Ilex*<br>*Quercus Pubescens* | Variations in water relations between a naturally high $CO_2$ site and a control site. | Leaf conductance and sap flow was both decreased in plants originating from the naturally high $CO_2$ site. |
| Bettarini, Vaccari *et al*. (1999) | *Allium sphaerocephalon*<br>*Convolvolus cantabrica*<br>*Globularia punctata*<br>*Planatgo lanceolata*<br>*Scabiosa columbaria*<br>*Silene vulgaris*<br>*Stachys recta* | Stomatal physiology and morphology of calcareous grasslands in a $CO_2$ enriched environment. | Long term exposure to elevated $CO_2$ had no significant effect on stomatal density or index. Although not significant *P. lanceolata* showed a slight decrease in stomatal density but a non-significant increase in stomatal index in plants originating from the natural spring, compared to the control site. |
| Johnson, Michelozzi *et al.* (1997) | *Quercus pubescens* | The impact of exposure to naturally elevated $CO_2$ concentrations on the leaf carbon economy. | Plants grown in sites of natural elevated $CO_2$ had similar rates of photosynthesis, whist retaining significantly higher water use efficiencies compared to a control. They also exhibited higher isoprene emissions which may relate to the use of the excess $CO_2$ which they cannot use for growth. |
| Paoletti, Raschi *et al*. (1997) | *Quercus ilex* | To investigate the stomatal morphology of leaves of holm oak trees grown in a naturally and artificial $CO_2$ enriched environment. | An increase in stomatal density in both natural and artificial elevated $CO_2$ environments was observed, compared to ambient controls. |
| Ineson, Cotrufo *et al*. (1997) | *Quercus ilex*<br>*Quercus pubescens* | The impacts of elevated $CO_2$ on the decomposition process. | Elevated $CO_2$ decreases litter Nitrogen concentrations with a resulting increase in Carbon/Nitrogen and lignin/Nitrogen ratios, leading to a slowdown in litter decomposition rates. |
| Van Gardingen, Grace *et al*. (1997) | *Phragmites australis*<br>*Quercus Pubescens* | The long term effects of enhanced $CO_2$ concentrations on leaf gas exchange. | *P. australis* showed reductions in stomatal density, stomatal conductance, and maximum photosynthetic rates when compared with nearby control sites. In contrast, work with *Q. pubescens* showed no evidence of photosynthetic acclimation. |
| Woodward (1999) | *Planatgo lanceolata* | Morphological and stomatal adaptations to an elevated $CO_2$ environment, using seed collected from the spring and a nearby control site grown in controlled environment chambers. | *P. lanceolata* showed stomatal initiation decreased with $CO_2$ enrichment when originated from ambient $CO_2$ conditions, whereas plants which originated from elevated $CO_2$ showed a non-significant increase in stomatal index when grown in elevated conditions. Both biomass and leaf size decreased in plants originating from elevated $CO_2$, and increased in plants originating from ambient $CO_2$, when exposed to elevated $CO_2$. |
| Marchi, Tognetti *et al*. (2004) | *Plantago lanceolata*<br>*Arabis irsuta*<br>*Scabiosa columbaria*<br>*Silene Vulgaris*<br>*Hypocrepis comasa* | Identify changes in stomatal density and index using a controlled $CO_2$ environment inside a natural $CO_2$ spring. | *P. lanceolata, A irsuta* and S. *columabia* all showed a significant increase in stomatal density under elevated $CO_2$ compared to ambient. Other species showed no significant difference for density, and no species showed a significant difference for index. |

| | | | |
|---|---|---|---|
| Balaguer, Manrique *et al.* (1999) | *Parmelia caperata* (Green-algal lichen) | The effect of long term elevated $CO_2$ exposure on photosynthetic capacity of the green-algal lichen. | No down regulation of photosynthetic capacity was found, but a decrease in RuBisCO was found in the pyrenoid of the algae. Light saturated photosynthesis was found to be similar in algae originating from elevated $CO_2$ concentration compared to plants originating from ambient $CO_2$ concentrations. |
| Hattenschwiler, Miglietta *et al*. (1997) | *Quercus ilex* | How 30 years exposure to an elevated $CO_2$ environment affects trees as they come to maturity. | Trees grown in elevated $CO_2$ conditions has a 12% greater final radial stem width compared to trees grown in a control site. The annual difference between the elevated $CO_2$ and control site was its greatest before 25 years, suggesting forest growth stands to be accelerated in the regeneration phase. |
| Miglietta, Raschi *et al*. (1998) | Grassland species<br>Ruderal species<br>Tree species<br>*Plantogo lanceolata* | How carbon isotope discrimination differs between plants grown in multi –generational elevated $CO_2$ compared to a control site. How the photosynthetic capacity of *Plantago lanceolata* is affected by multi-generational elevated $CO_2$ concentrations. | Carbon isotope discrimination was significantly lower in grassland species from the spring compared to the control site, but not significant in ruderal species. Tree species showed both significantly lower values (*Q. pubescens*) and values with no significant difference (*Q. ilex*). This indicates photosynthesis may be affected differently depending on the species. Photosynthesis was significantly down regulated in *P. lanceolata*, as well as lower stomatal conductance in plants originated from the spring compared to the control site. |
| Andalo, Godelle *et al*. (1999) | *Arabidopsis thaliana* | Investigate whether populations of *A.thaliana* originating from an elevated $CO_2$ spring have adapted to the conditions, compared to populations from a nearby ambient control sites. | All populations responded in exactly the same way to elevated $CO_2$, apart from seed characteristics, which were lighter, produced a higher number, with lower germination rates, in plants originating from the elevated $CO_2$ spring. |
| Scholefield, Doick *et al*. (2004) | *Phragmites australis* | Investigate the isoprene emissions of plants originating within the spring, compared to a nearby control site. | Isoprene emissions were lower in plants in plants growing at elevated $CO_2$ in the natural spring, then plants growing at ambient $CO_2$ at the nearby control site. |

Table 1.4 – A summary of studies related to $CO_2$, conducted using the Bossoleto spring, Italy.

Due to the variables identified at the spring site in Bossoleto, it is important to relate any findings to other springs in an attempt to find consistent results. There are the springs across the rest of Italy (Fordham, A. Raschi *et al.* 1997), but also Asia (Onoda, Hirose *et al.* 2007), America (Ehleringer J.R., Sandquist D.R. *et al.* 1997), Africa (M. Mousseau, Sabroux *et al.* 1997, Woodward, Beerling *et al.* 1997) and the rest of Europe (Cook, Oechel *et al.* 1997) in which studies on a range of the effects of elevated $CO_2$ levels plant species have also been carried out. A selection of these studies are summarised in Table 1.5.

| *Study* | *Species* | *Study Focus* | *Findings* |
|---|---|---|---|
| Fordham, Raschi *et al*. (1997) Solfatara (Italy) | *Agrostis canina* *Plantago major* | The study looked at the growth and dry matter partitioning under both elevated and ambient $CO_2$ using seed from naturally elevated $CO_2$ spring, and a nearby control site. | Plants grown in elevated $CO_2$ produced a greater biomass, and exhibited higher initial relative growth rates. Seeds which originated from the elevated $CO_2$ regions of the natural springs were larger, and larger seeds were associated with the high relative growth rates exhibited by the plants grown under elevated $CO_2$. |
| Bettarini, Raschi *et al*. (1997) Viterbo, Rome, and Sienna (Italy) | *Scirpus lacustris* | A study of adaptive traits in relation to elevated $CO_2$ was conducted on genetically isolated populations of *S. lacustris* collected from a $CO_2$ spring, compared to a nearby ambient control site. | Plants originating from the $CO_2$ spring, showed no significant difference in photosynthetic rate, nitrogen content or Ci/Ca ratio, but did exhibit a downward regulation in stomatal density for the spring verses control plants. |
| Woodward, Beerling *et al*. (1997) Florida (America) and Egypt | *Boehemeria cylindrica* | The response of stomatal density to variations in $CO_2$ concentrations, independent of the climate experienced by plants from different regions of the globe. | The two sites from Florida and the tombs in Egypt showed no significant difference when exposed to the both ambient and elevated $CO_2$ concentrations. |
| Marchi, Tognetti *et al*. (2004) Strmec (Slovenia) | *Tanacetum vulgaris* *Rumex crispus* *Plantago lanceolata* *Polygonum hydropiper* *Trifolium pratense* | Identify changes in stomatal density and index using a naturally elevated $CO_2$ environment, and an ambient $CO_2$ control site. | *P. lanceolata* and *T. pratense* both showed a significant increase in stomatal density and index in the elevated $CO_2$ site compared to the ambient $CO_2$ site. *P. hydropiper* showed a significant decrease in both density and index in the same comparison, whilst the rest showed no significant differences. |
| Onoda, Hirose *et al*. (2007) Ryuzin-numa, Yuno-kawa and Nyuu (Japan) | *Hydrangea paniculata* *Polygonum sachalinerse* *Sasa kurilensis* *Tiarella polyphylla* *Phragmites australis* *Plantago asiatica* | The response of monocotyledonous and dicotyledonous species to multi-generational elevated $CO_2$, with respect to nutrient uptake, photosynthesis, and iWUE. | Dicotyledonous species from the elevated $CO_2$ spring contained more starch and less nitrogen than ambient controls, whereas monocotyledonous species showed no starch or nitrogen differences. Photosynthesis was up regulated in plants from the elevated $CO_2$ spring compared to controls when grown at their respective $CO_2$ concentrations, but when both were measured at ambient $CO_2$, the spring plants had lower photosynthetic rates than the controls. |

| | | | |
|---|---|---|---|
| Onoda, Hirose *et al.* (2009) Ryuzin-numa, Yuno-kawa and Nyuu (Japan) | *Plantago asiatica*<br>*Polygonum sachalinense* | Plants from an elevated $CO_2$ spring, and a nearby ambient $CO_2$ control site, were both grown within the spring and at the control site. Photosynthesis, iWUE and nutrient uptake were measured and compared between treatments and origin of the plants. | There was an increase in photosynthetic capacity, iWUE and starch accumulation in response to elevated $CO_2$ for both sites of origin, but no significant difference between the origins of the plants. There was a significant origin x $CO_2$ effect for iWUE, suggesting this aspect could be adaptive. |
| Tognetti, Raschi *et al.* (2000) Lajatico (Italy) | *Erica arborea*<br>*Myrtus communis*<br>*Juniperus communis* | Identify differences in water relations between species from a naturally elevated $CO_2$ spring and a nearby ambient $CO_2$ control site. | Plants originating from the spring exhibited an increase in turgor pressure, particularly during the summer months. Species dependant reactions to osmotic potentials were observed. Plants from the spring showed increased osmotic potentials in *E. arborea* under drought, while the opposite was the case for *J. communis*. This indicates species can have different strategies of water relations in response to elevated $CO_2$. |
| Fordham and Barnes (1999) Solfatara (Italy) | *Agrosstis canina*<br>*Plantago major* | Characterise the impacts of $CO_2$-enrichment on growth and photosynthesis on plants adapted to long term elevated $CO_2$ and a control site, grown at both elevated and ambient $CO_2$. | Plants originating from the spring showed an increase in biomass at both ambient and elevated $CO_2$, largely due to an initial stimulation in relative growth rate. No significant differences in photosynthetic characteristics were found between populations or $CO_2$ treatments. |

Table 1.5 – A summary of a selection of studies utilising springs from around the world.

Studies using natural springs have provided an interesting insight into plant responses to multi-generational elevated $CO_2$, but no study has been able to give a definitive response. Variability in photosynthesis, stomatal numbers, starch content, water relations and carbon isotope discrimination have all been found in a relatively small amount of studies. This is partly because variability between species is rife, but may also be due to the different conditions within the range of springs used. It is important to understand these potential adaptations with more clarity, understanding why we see the variation we do, and which responses yield the most positive results, or even negative results. The variability in stomatal response (Chapter 2) and photosynthesis (Chapter 4) are two areas which will be extensively examined here in order to hope to elucidate these responses, and provide a clearer conclusion in terms of multi-generational responses to elevated $CO_2$.

## 1.9 Species of Interest

The flora within the Bossoleta spring has a highly diversified taxonomic range (Selvi 1999), more so than the other natural springs in Italy. It contains a vascular flora of about 90 species of grass, herbs, shrubs and trees belonging to 39 families, along with several bryophytes and lichens (Selvi 1999).

### 1.9.1 *Plantago lanceolata*

*Plantago lanceolata* (Figure 1.16) is a perennial species found in the spring which is part of the *Plantago* genus and is classed as a weed due to its invasive nature. It is widespread over the world, especially in the Americas, Australia and parts of Europe. In the UK it is known as ribwort plantain and is often harvested as its leaves are palatable and seeds provide valuable nutritious food for grazing stock and wildlife. Field longevity estimates for *P. lanceolata* are highly variable (Roach 2003), and individuals may live to be more than 5 years old, while in other cases there may be 99% mortality of a cohort in the first year.

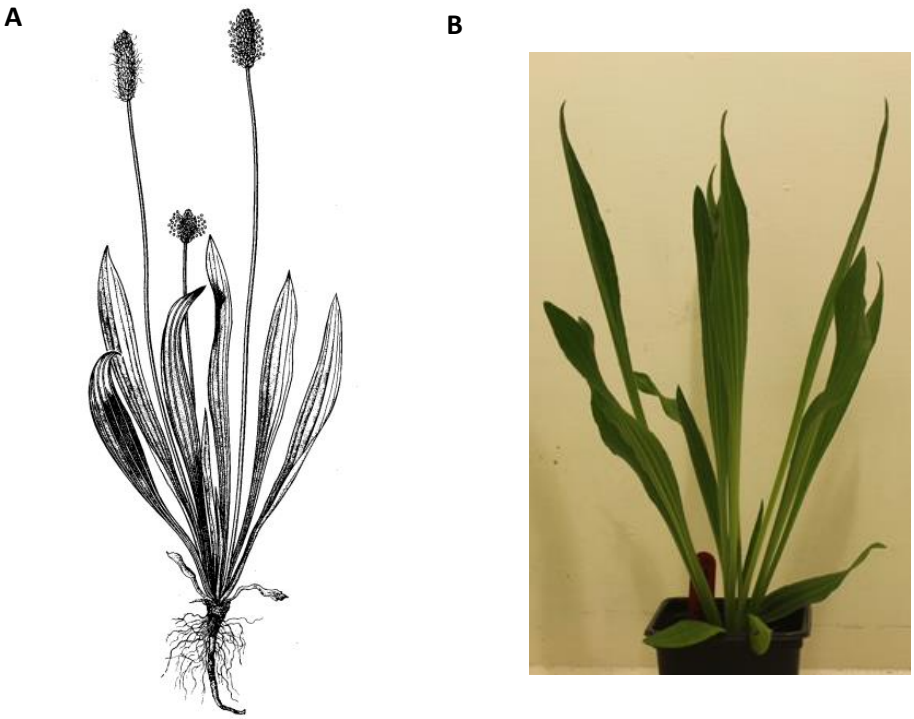



Figure 1.16—(**A**) A drawn image of *P. lanceolata* (Cavers, Bassett *et al.* 1980), and (**B**) a 6 week old *P. lanceolata.*

*P. lanceolata* has restricted seed dispersal and the seed of *P. lanceolata* is wind-dispersed, mainly over relatively short distances of around 1.5m (Tonsor 1985). It also has a gametophytic self-incompatibility system and a protogynous growth syndrome making it an obligate out crosser (Bos and Van der Haring 1988). This can lead to population sub-structuring in wind pollinated species with low seed dispersal (Bos, Harmens *et al.* 1986). Clonal reproduction can take place via the production of side-rosettes (Cavers, Bassett *et al.* 1980), and this incidence of side-rosettes is considered to be quite high. The genetic structure within populations of a species like *P. lanceolata* can still be explained by restricted gene flow, in spite of the presence of outcrossing mechanisms like self-incompatibility and protogyny.

The chromosome number for *P. lanceolata* has been reported as 2n = 12 from widely different locations, Canada (Bassett and Crompton 1968), United States and Eurasia (Sagar and Harper 1964). The genome for *P. lanceolata* has not yet been sequenced which makes certain genetic analyses of the species very difficult. The nearest fully sequenced genomes are the tomato, potato and tobacco species, but these are still fairly distant taxonomically. The olive genome is currently being sequenced (Barghini, Natali *et al.* 2014) and when this is released will be a better reference genome for *P. lanceolata* as the olive is in the same order as *Plantago* (Figure 1.17). *Mimulus guttatus* has partial sequencing available (Wu, Lowry *et al.* 2007) and that is closer taxonomically than olive to *P. lanceolata* (Figure 1.19), and so could be used when looking for specific genes, but as only partial sequence is available cannot yet be used as a reference genome.

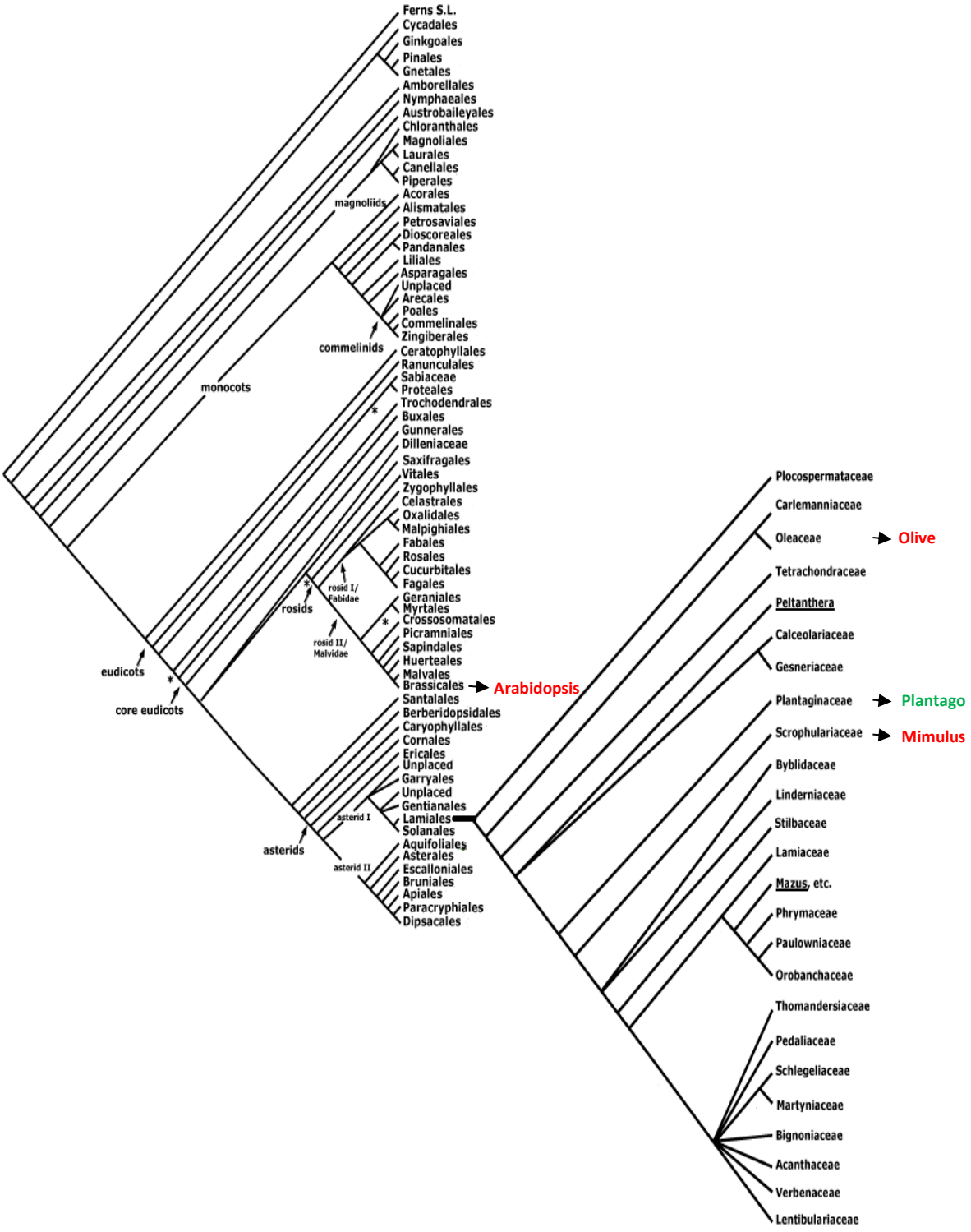



Figure 1.17- The taxonomy of Plantago (green), Arabidopsis, Mimulus and the Olive (red), taken from 'http://www.mobot.org/MOBOT/research/APweb/welcome.html.'

*P. lanceolata* is of interest as it appears not to follow the previously documented common stomatal adaptations/acclimations to elevated $CO_2$ concentrations. Samples of the species collected from the Bossoleta spring showed an up-regulation in stomatal numbers compared to samples grown in ambient $CO_2$. This has been documented previously, once significantly (Marchi, Tognetti *et al.* 2004), and once where the increase was not statistically significant (Woodward 1999). Marchi *et al.* (2004) created a controlled environment within the spring using polycarbonate film, where the $CO_2$ concentration was controlled so the concentration didn't fluctuate, whilst retaining other aspects of the spring. They showed stomatal density was significantly higher in the plants grown in elevated $CO_2$ compared to ambient $CO_2$, whereas stomatal index was not significantly different. Although interesting, this comparison shows a short term acclamatory response to $CO_2$, but is interesting in terms of how plants grown in elevated $CO_2$ over generations respond to ambient $CO_2$. Historic studies would suggest that an increase in $CO_2$ results in a decrease in stomatal numbers, whereas this study suggests a decrease in $CO_2$ results in fewer stomata. Marchi *et al.* (2004) also used a second spring in Strmec (Slovenia), where they compared samples of *P. lanceolata* from inside the spring and a nearby control site (similar to the Woodward (1999) study). Here they found both a significant increase in stomatal density and index in *P. lanceolata* grown inside the spring compared to the ambient control site. This shows the potential variation that could be occurring between springs, and why it is important to use controlled environments as well as in-situ studies to help identify the drivers of phenotypes.

There have been previous studies which have investigated *P. lanceolata* using a controlled environment experiments in relation to $CO_2$ (Klus, Kalisz *et al.* 2001). Klus *et al.* (2001) used several families from two populations of *P. lanceolata* from normal ambient $CO_2$ conditions, and then placed them into open top chambers at two partial pressures of $CO_2$, 35 Pa and 71 Pa to ascertain differences in family inheritance of responses to changing atmospheric $CO_2$. Significant effects for $CO_2$ × family were observed for nitrogen content, carbon : nitrogen ratio, photosynthetic rate, intercellular $CO_2$ concentration, transpiration rate, and water use efficiency (Klus, Kalisz *et al.* 2001). Interestingly stomatal conductance did not have a significant response to elevated $CO_2$. It was also found that families within populations differed significantly in their average trait values, regardless of the $CO_2$ environment in which they were examined. This shows that inherited variation plays a role in how this species will respond to elevated $CO_2$, and how variable the species is across families and populations. Other studies have documented this variation among populations in *P. lanceolata* (Wulff and Alexander 1985). Wulff and Alexander (1985) grew four genotypes of *P. lanceolata* to maturity at combinations of two levels of atmospheric $CO_2$ concentrations and two temperature conditions. The results of their study

showed a significant difference between families for several characteristics including seed weight and leaf area in response to elevated $CO_2$, indicating a potential for long term selection.

The genetic composition of *P. lanceolata* has been investigated in relation to ozone ($O_3$) concentration. Kölliker, Bassin *et al.* (2008) used amplified fragment length polymorphisms (AFLP) and simple sequence repeat markers (SSR) to test for differences in *P. lanceolata* under elevated and ambient $O_3$ after 5 years of exposure. Genetic diversity was shown to be higher in populations exposed to elevated $O_3$ than in populations sampled from control plots. This was based on a significant effect for both the AFLP-based measures of diversity and the SSR markers based on observed heterozygosity. A small but significant difference in genetic composition between $O_3$ treatments was also detected by analysis of molecular variance and redundancy analysis. These results show that *P. lanceolata* show micro-evolutionary processes that could in this case, take place in response to long-term elevated $O_3$ exposure in highly diverse populations of outbreeding plant species. This also shows the potential that the same micro-evolution could occur for other components such as $CO_2$.

### 1.9.2 *Silene vulgaris* and *Sanguisorba minor*

Although *P. lanceolata* is the main species of interest, some comparative analyses were carried out on two other species from the spring, *Silene vulgaris* and *Sanguisorba minor*. *S. vulgaris* is a semi-sprawling grassland perennial of the genus *Silene* of the Pink Family (Caryophyllaceae), and is also known as Bladder campion. It is found on a wide range of well drained soils, usually on sites that have been subject to a degree of disturbance like rough pasture, roadside verges, waste land and the edge of arable fields. *S. minor* is an evergreen perennial of the family Rosaceae native to most of Europe and the Americas. It is typically found in dry grassy meadows, often on limestone soils, and is drought-tolerant, growing all year around.

## 1.10 Aims

-The over-arching aim of this PhD is to understand the multi-generational plant response to elevated $CO_2$ here termed 'adaptation', comparing this response to the 'acclimation' or short term phenotypic plasticity that has been observed earlier.

-To identify the underlying gene expression changes linked to adaptation and acclimation observed using RNA-Seq.

-Test the physiological mechanisms underlying the phenotypic/gene expression changes observed to identify the how the function of these mechanisms are affected.

-Use population genetics to further understand the origins of the variation between populations and individual genes to elucidate where selection pressures are occurring in response to elevated $CO_2$.

# Morphological Responses to Multi-Generational Exposure to Elevated $CO_2$

## 2.1 Introduction

Using natural springs has produced some interesting multi- generational responses to elevated $CO_2$, in a range of species, in springs across the world. Many of these studies have originated from the spring site in Bossoleto (Table 1.4), where previous multi-generational effects of *P. lanceolata*, alongside several other species, have been investigated. Bettarini *et al.* (1999) found no significant difference in stomatal index between a population of *P. lanceolata* originating within the naturally elevated $CO_2$ spring, and a nearby ambient $CO_2$ control site. Woodward (1999) collected seed *of P. lanceolata* from both the Bossoleto spring and a nearby control site, and exposed both to ambient and elevated $CO_2$ using controlled environments. Using this method he was able to identify how the two populations respond to elevated $CO_2$. The control population exhibited a decrease in stomatal index in response to elevated $CO_2$, whereas the spring population showed no significant response. Although not significant, the trend moved towards an increase in stomatal index in spring populations, in response to elevated $CO_2$. Marchi *et al.* (2004) later found a significant increase in stomatal density for *P. lanceolata* in response to elevated $CO_2$; again this comparison was between plants which had both originated from multi-generational elevated $CO_2$ environments. A significant increase in stomatal index and density was noted for *P. lanceolata* in samples taken from a spring in Strmec (Croatia), when compared to a nearby control site. This agrees with the trend found in the Bossoleta spring, but is the first instance where the difference has been shown to be significant when compared to a control population. This response is not unique to *P. lanceolata*, but has also been found in almost all species studied within natural springs, showing either an increase or no significant difference in stomatal density or index in response to elevated $CO_2$ (Paoletti, Raschi *et al.* 1997, Andalo, Godelle *et al.* 1999, Bettarini, Vaccari *et al.* 1999, Marchi, Tognetti *et al.* 2004). An increase in stomatal numbers was noted in one other species within the Bossoleto spring, Paoletti *et al.* (1997) showed stomatal density increased in *Q. ilex* originating in the spring in response to elevated $CO_2$. A decrease in stomatal density was shown in *P. australis* (Van Gardingen, Grace *et al.* 1997), but interestingly this was the only case. This is a shift in the previous findings of stomatal response to elevated $CO_2$ in controlled environments, open top chambers and FACE studies, where the most common phenotype was decreased stomata (Woodward and Kelly 1995). The majority of species using natural springs show no significant stomatal response to elevated $CO_2$ (Bettarini, Vaccari *et al.* 1999), and may represent the true plant adaptation over multiple generations. This gives an insight into why we see an increase in stomata in *P. lanceolata* originating from an elevated $CO_2$ environment, compared to an ambient $CO_2$ environment, as they are governed by a new common phenotype.

Other morphological traits which have been measured within natural springs include biomass, and seed characteristics. An increase in biomass is expected under elevated $CO_2$ (Curtis and Wang 1998), due to the additional assimilates available, but research has shown that the increase is often lower than expected (Ainsworth and Long 2005). *P. lanceolata* originating from the control site have been shown to significantly increase biomass when grown at elevated $CO_2$, whereas *P. lanceolata* originating from naturally elevated $CO_2$ springs show no significant difference in biomass, and even a reduction in leaf area when grown at elevated $CO_2$ (Woodward 1999). The root: shoot ratio of *P. lanceolata* originating from the spring was also increased (Woodward 1999), in contrast to a decrease in populations from the control site, suggesting the spring plants are relocating assimilates for growth. Several other studies have noted a significant increase in biomass in various other species originating from naturally elevated $CO_2$ springs, when compared to nearby control sites (Fordham, Raschi *et al.* 1997, Hattenschwiler, Miglietta *et al.* 1997, Fordham and Barnes 1999). Further to this, unlike *P. lanceolata*, when seed originating from a naturally elevated spring, and seed from an ambient control site, have both been grown at elevated $CO_2$, the former has exhibited increased biomass compared to the control (Fordham and Barnes 1999). These findings suggest that the *P. lanceolata* is showing an irregular response, and the majority agree with previous hypotheses surrounding biomass under elevated $CO_2$, and also suggest a potential adaptation/acclimation, allowing $CO_2$ to be utilised more efficiently in plants originating from naturally elevated $CO_2$ springs. This increase in growth has been identified as occurring predominantly in the initial growth phase, and appears to not be as apparent as the plants age. Previous exposure to elevated $CO_2$ appears to alleviate this transitional period to the conditions which control plants exhibit, whilst growth is stimulated in spring plants, and highlights an important stage of development when looking to elucidate this area.

Analyses of *A. canina* and *P. major* seeds collected from within natural springs, compared to a nearby control site, showed seeds originating within the spring were larger (Fordham, Raschi *et al.* 1997), which may be associated with the increased growth. This was also investigated in *A. thaliana*, where lighter seeds, in higher numbers, were identified inside the spring, along with lower germination rates (Andalo, Godelle *et al.* 1999). These findings appear to contradict, but may just represent different adaptations. The species represent populations from different springs, and so may require different characteristics to thrive in their separate environments. The literature shows a variety of seed responses to elevated $CO_2$ (Jablonski, Wang *et al.* 2002), which helps explain why there may be variation surrounding this aspect.

As previously mentioned, there are a number of variables present when using natural springs, and as much as these variables can be selected against, there will always be variables present when using such a diverse environment. One way to combat this is to take plants from the spring and grow them in controlled environment systems. Investigations using controlled environments are not new, and have been extensively used in research to elucidate responses to elevated $CO_2$ (Warwick and Taylor 1995, Kellomäki, Wang *et al.* 2000). Using seed from natural springs paired with controlled environments negates the variables as much as possible, and the multi-generational effects of $CO_2$ are not lost. Studies using natural spring sites have utilised this method due to the advantages surrounding it (Woodward 1999, Marchi, Tognetti *et al.* 2004). Adopting this method, the morphology of species within the spring site in Bossoleto can be further investigated, especially with the hope to elucidate the morphological phenotypes of *P. lanceolata*.

### 2.1.1 Aims

- Using seeds collected from the Bossoleto spring and a nearby control site, grow *P. lanceolata* in both ambient and elevated $CO_2$ in controlled environments to investigate both acclimation and adaptation to elevated $CO_2$.

-In particular to understand the physiological, growth and morphological responses to ambient and elevated $CO_2$, to confirm previous findings enabling these to be linked to underlying global gene expression and DNA variation, thus providing an entirely novel dataset for this well-studied site.

-To study, in detail, stomatal responses to elevated $CO_2$, unravelling the inconsistencies in previous spring studies, for *Plantago lanceolata* such that these may be investigated further using functional genomic and genetic approaches at the molecular level and previously uncharacterised.

## 2.2 Materials and Method

### 2.2.1 Plant Material

*Plantago lanceolata* seeds were collected from a natural carbon dioxide spring (Lat. 43°17', Long. 11°35') called Bossoleto, located in Rapolano Terme, near Siena in Central Italy. The site has been extensively described and studied (Grace, Gardingen *et al.* 1997, Selvi 1999). Seeds samples were taken from two sites. *Sanguisorba minor* and *Silene vulgaris* seeds were also collected from both sites. The first site is located in the spring where the $CO_2$ concentrations were classed as elevated (spring plants(S)), ranging from 600 to 1200 µmol $mol^{-1}$ due to emissions from gas vents, from historic data (Figure 2.1). The second is a control site 200m from the spring where the concentrations were classed as ambient (control plants (C)), around 390 µmol $mol^{-1}$. Seeds were collected from nine randomly selected maternal plants at both sites on 12 May 2008 and stored in the University of Southampton's cold room (0˚C).

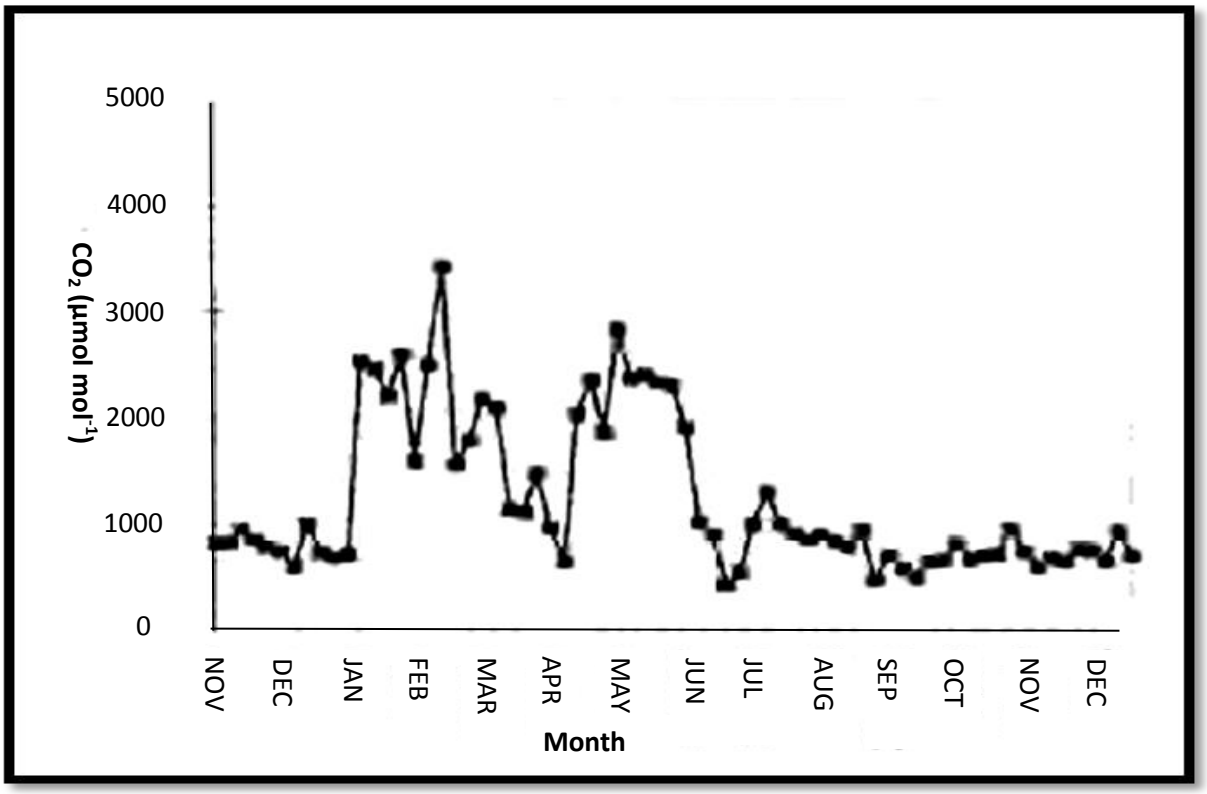



Figure 2.1 – The daily recording of the $CO_2$ concentration (µmol $mol^{-1}$) within the spring in Bossoleto measured from November 1992 to December 1993 using an infrared gas analyser and omitting values off scale (≤340 and ≥7500 µmol $mol^{-1}$) (Bettarini, Grifoni *et al.* 1999).

## 2.2.2 Chamber experiment

In the experiments described here, the first chamber experiment was conducted by Yunan Lin, as part of her PhD. However, data presented here for growth have been re-analysed, samples of all leaf epidermal imprints have been re-sampled, measured and analysed and provide a new insight into the original study. The seeds were potted using compost (John Innes potting compost No.2, John Innes manufactures association, UK), covered by 1 cm of fine sand. They were established on the 21st Sep 2009 at 23°C/ 20°C and 16h day length. Three weeks after germination the seedlings were transferred into small pots filled with sterilized vermiculite and moved into $CO_2$ chambers, so the $CO_2$ concentration could be regulated. Eight chambers were used, four chambers at an ambient $CO_2$ concentration (Average = 410.63 ± 33.74 µmol $mol^{-1}$) and four at an elevated $CO_2$ concentration (Average = 718 ± 46.81 µmol $mol^{-1}$) (22°C /17°C, 16h day length, flow rate 3.4 $m/s^2$ and light intensity around 104-134 µmol/$m^2 \cdot s^2$). The position of the pots was randomised within and between each chamber every week and treatments were swapped between chambers to remove chamber effects (Warwick and Taylor 1995). One plant from each of the nine maternal plants from each site was grown in each chamber. This gave four sample groups;

A. Control Ambient (CA) –Progeny of each of the maternal plants from the control site, grown in a chamber at a $CO_2$ of 390 µmol $mol^{-1}$.

B. Spring Ambient (SA) - Progeny of each of the maternal plants from the spring site, grown in a chamber at a $CO_2$ of 390 µmol $mol^{-1}$.

C. Control Elevated (CE) - Progeny of each of the maternal plants from the control site, grown in a chamber at a $CO_2$ of 700 µmol $mol^{-1}$.

D. Spring Elevated (SE)- Progeny of each of the maternal plants from the spring site, grown in a chamber at a $CO_2$ of 700 µmol $mol^{-1}$.

Every two weeks, additional nitrogen (all-purpose soluble plant food, The Scotts Miracle-Gro Company, USA) was added. At 123-124 days after establishment, cell imprints were taken from each plant for analysis. Imprints were originally taken by Prof. G. Taylor, but re-sampled and analysed here. The cell imprints were taken from the abaxial surface of the leaves. They were taken by coating the mid-section of the first fully mature leaf of the abaxial leaf surface with nail varnish. Once the nail polish was dried the coated area of the leaf was removed using sellotape. This left an imprint on the nail varnish, which was then placed on a glass slide and labelled.

The slides were analysed using a Zeiss upright light microscope. Five images of each slide were taken using a x10 magnification lens, ensuring each image had a calibration bar. These images

were then analysed using ImageJ (Schneider, Rasband *et al.* 2012). Five random 400 µm x 400 µm boxes were then selected from each image and five random epidermal and guard cells were selected from each box and measured for area. The number of epidermal cells and stomata were also counted in each box.

Mean values were calculated to give one value for each image, thus avoiding pseudo replication. Using these measurements the following calculations were made;

- Stomatal density (See Section 1.6.1)
- Stomatal index (See Section 1.6.2)
- Epidermal cell size
- Epidermal cell number

The plants were then harvested and additional measurements were taken:

- Single leaf dry mass
- Specific leaf area
- Single leaf area
- Above ground biomass

The single leaf dry mass, single leaf area and specific leaf area was taken on a single young mature leaf picked from each plant.

### 2.2.3 *In-Situ* Measurements

To investigate the stomatal phenotypes inside and outside the spring stomatal imprints were taken from the plants *in-situ* from the spring and the control site. Forty samples of *P. lanceolata* (20 spring and 20 control) were collected. Stomatal imprints of *Sanguisorba minor* and *Silene vulgaris* were also taken at the same time to see how the phenotype contrasted across different species. 20 samples of both *S. minor* and *S. vulgaris* were taken (10 spring and 10 control). The measurements carried out on the imprints were:

- Stomatal density (See Section 1.6.1)
- Stomatal index (See Section 1.6.2)
- Epidermal cell size
- Guard cell length

## 2.2.4 Statistics and Experimental Design

The morphology measurements were tested using a Generalized Linear Model (GLM) in SPSS 16 (SPSS Inc.) with the model: Response = Location + Treatment + Treatment | Location + Treatment | Treatment | Family (Location) + Location | Chamber (Treatment) (Doncaster & Davey, 2007) (Table 2.1). If normalisation tests were rejected (P<0.05) with the *Kolmogorov-Smirnov* test traits were log-transformed before analysis. *P. lanceolata* are not capable of self-fertilization, plus the plants did not flower and were moved between and within each chamber every week, so family and chamber can both be regarded as random factors. The degrees of freedom of component Chamber (Treatment) | Family (Location) is 0 so there is no statistic test possible on this component. If a post-hoc test was required to be carried out to further investigate a value of significance, a Tukey's honestly significant difference (HSD) test was carried out to show which means differed.

| Y=S'(A)\|C'(B) | | Control Site Plants (A1) | | | | | | | | | Spring site plants (A2) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S11 | S12 | S13 | S14 | S15 | S16 | S17 | S18 | S19 |
| a[$CO_2$] (B1) | C1 | Q1 | Q3 | Q3 | Q4 | Q5 | Q6 | … | … | … | … | … | … | … | … | … | … | … | … |
| | C2 | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| | C3 | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| | C4 | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| e[$CO_2$] (B2) | C5 | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| | C6 | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| | C7 | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … | … |
| | C8 | … | … | … | … | … | … | … | … | … | … | … | … | … | … | Q141 | Q142 | Q143 | Q144 |

Table 2.1 –Table representing the statistical model used. 'A' represents original site (location), 'B' represents $CO_2$ concentration (treatment), 'C' represents chamber number (chamber), 'S' represents P. lanceolata's maternal family (family) and 'Q' represents individual plants.

For the samples taken in-situ a one way ANOVA was used to analyse the data as there was only fixed factor, Location, with two levels, spring and control.

## 2.3 Results

### 2.3.1 Controlled Environment Experiment

From the stomatal imprints differences are apparent between elevated and ambient $CO_2$ treatments, especially in the spring plants where there are clearly more stomata in the spring elevated imprint than the spring ambient imprint. Full analysis was conducted on all of these imprints along with the other morphological measurements (Figure 2.2).

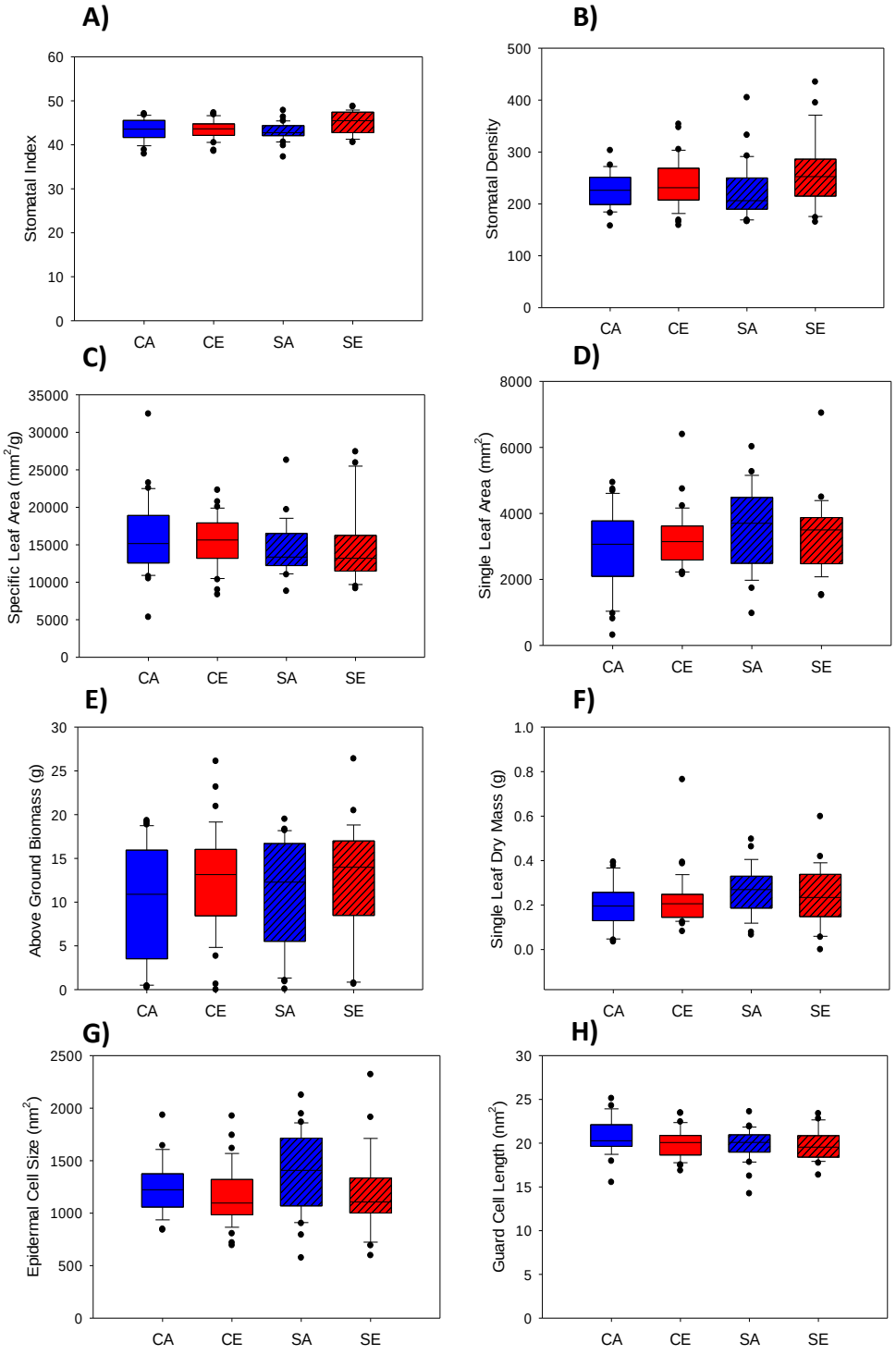



Figure 2.2 – The impact of elevated $CO_2$ on *P. lanceolata*, originating from either a naturally high $CO_2$ spring (S) or a nearby ambient $CO_2$, control (C), site exposed to either ambient (A) or elevated $CO_2$ (E) at a target concentration of 700 umol $mol^{-1}$ .Results for growth and morphology traits are shown. CA= Control Ambient, CE = Control Elevated, SA = Spring Ambient, SE = Spring Elevated. The central line in each boxplot shows the median. Whiskers indicated the $5^{th}$/$95^{th}$ percentiles. Each dot indicates the observation of an outlier.

Representative images of the stomatal imprints of the *P. lanceolata* taken from the chamber experiment can be seen in Figure 2.3.

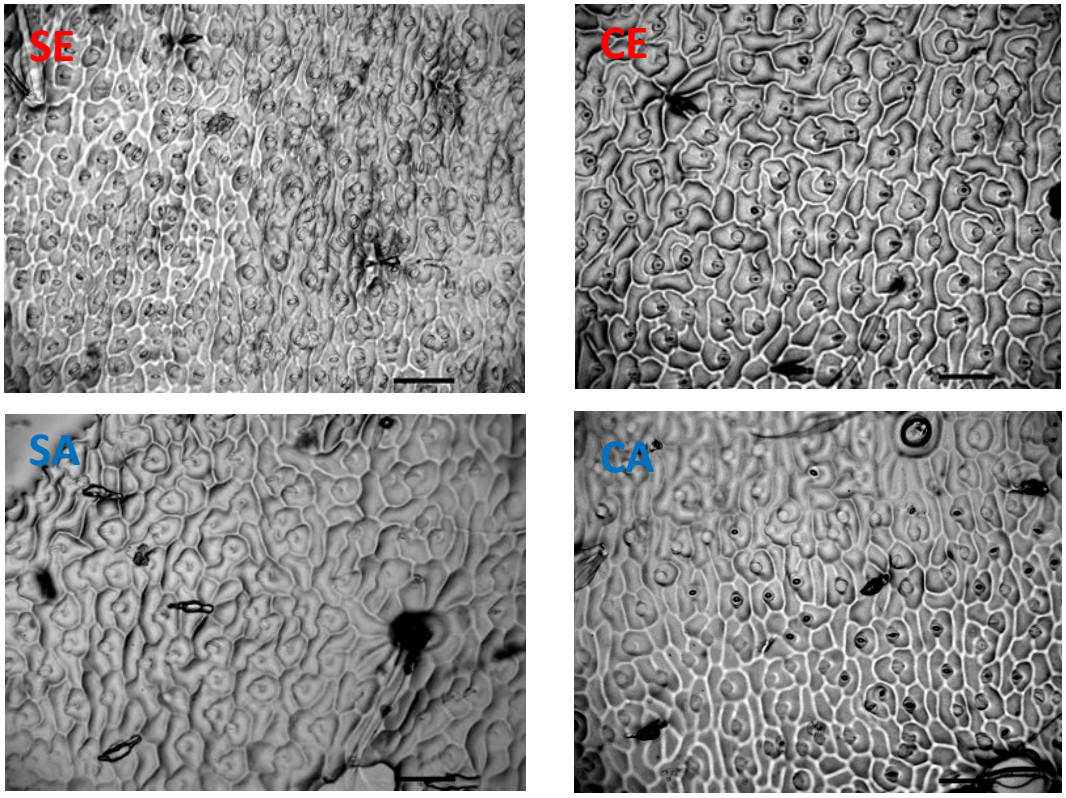



Figure 2.3 - Stomatal imprints from *P. lanceolata*. SE = Spring Elevated, SA = Spring Ambient, CE = Control Elevated, CA= Control Ambient, 100µm scale bar.

The percentage change of all of the traits measured in the original set of measurements was calculated (Figure 2.4). The percentage change was calculated using the formula below;

$$\text{Percentage Change (\%)} = \frac{(\text{Elevated} - \text{Ambient})}{\text{Ambient}} * 100$$

Formula 2.1 – The formula used to calculate percentage change between elevated and ambient samples.

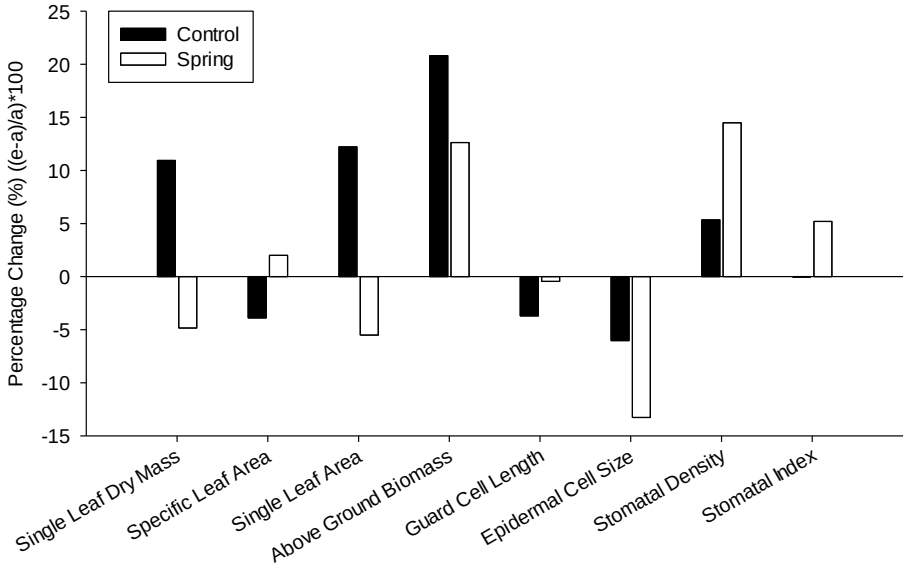



Figure 2.4 – The impact of elevated $CO_2$ on *P. lanceolata*, originating from either a naturally high $CO_2$ spring (open bars) or a nearby ambient $CO_2$, control (solid bars). The percentage change was calculated for each morphological trait using the calculation ((E-A)/A)*100), where E = plants grown at elevated $CO_2$ (E) at a target concentration of 700 umol $mol^{-1}$, and A = plants grown at ambient $CO_2$ concentration.

A generalised linear model (see section 2.2.4) was used to identify any statistical differences between the groups (Table 2.2).

| Source | d.f. | Stomatal Index | | Stomatal Density | | Guard Cell Length | | Epidermal Cell Size | |
|---|---|---|---|---|---|---|---|---|---|
| | | *T* | *P* | *T* | *P* | *T* | *P* | *T* | *P* |
| **Location** | 1 | 7.087 | **0.008**** | 1.294 | 0.255 | 1.891 | 0.169 | 3.156 | 0.076 |
| **Treatment** | 1 | 13.950 | **0.001**** | 12.296 | **0.001**** | 1.544 | 0.214 | 13.203 | **0.001**** |
| **Location*Treatment** | 1 | 9.335 | **0.002**** | 2.891 | 0.089 | 0.406 | 0.524 | 3.314 | 0.069 |
| **Treatment*Family(Location)** | 16 | 32.242 | **0.009**** | 16.686 | 0.406 | 23.408 | 0.103 | 13.808 | 0.631 |
| **Location*Chamber (Treatment)** | 6 | 27.206 | **0.001**** | 28.283 | **0.001**** | 13.586 | **0.035*** | 36.603 | **0.001**** |



| Source | d.f. | Above Ground Biomass | | Single Leaf Dry Mass | | Single Leaf Area | | Specific Leaf Area | |
|---|---|---|---|---|---|---|---|---|---|
| | | *T* | *P* | *T* | *P* | *T* | *P* | *T* | *P* |
| **Location** | 1 | 0.002 | 0.963 | 6.126 | **0.013*** | 4.426 | **0.035*** | 1.263 | 0.261 |
| **Treatment** | 1 | 3.698 | 0.054 | 0.021 | 0.886 | 0.024 | 0.876 | 0.006 | 0.941 |
| **Location*Treatment** | 1 | 2.128 | 0.145 | 2.731 | 0.098 | 4.898 | **0.027*** | 3.135 | 0.077 |
| **Treatment*Family(Location)** | 16 | 24.262 | 0.084 | 40.611 | **0.001**** | 27.473 | **0.037*** | 29.914 | **0.018*** |
| **Location*Chamber (Treatment)** | 6 | 1.274 | 0.973 | 1.601 | 0.952 | 1.956 | 0.924 | 3.161 | 0.788 |

Table 2.2 – Statistical analysis of morphological data from *P. lanceolata*. A generalized Linear Model was used (See section 2.2.4). P = significance value, T = wald chi-square value. Significance level: * P≤0.05; ** P≤0.01.

## 2.3.2. Measurements of Stomatal Patterning at the Spring Site

Stomatal imprints were collected from inside the spring at Bossoletto, and at a nearby control site (see section 2.2.1). The results from the stomatal imprints can be seen below, Figure 2.5 shows the results for *P. lanceolata*, Figure 2.6 shows the results for *S. vulgaris* and Figure 2.7 shows the results for *S. minor*.

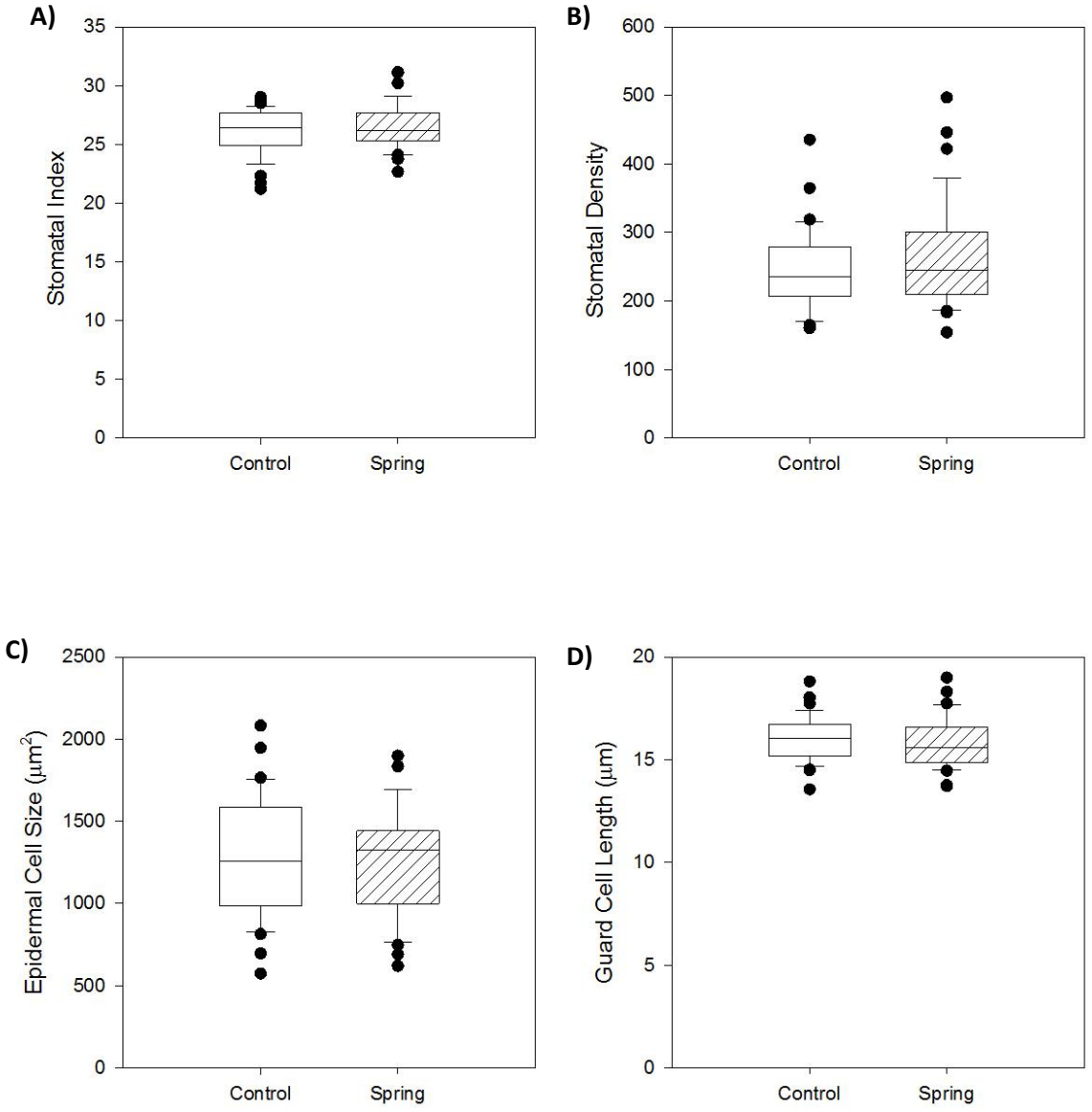



Figure 2.5 – The data from the stomatal imprints taken *in-situ* from the spring and the control site in Bossoleta for *P. lanceolata*. **A** is the stomatal index, **B** is the stomatal density, **C** is the epidermal cell size and **D** is the guard cell length. The central line in each boxplot shows the median. Whiskers indicated the $5^{th}$/$95^{th}$ percentiles. Each dot indicates the observation of an outlier.

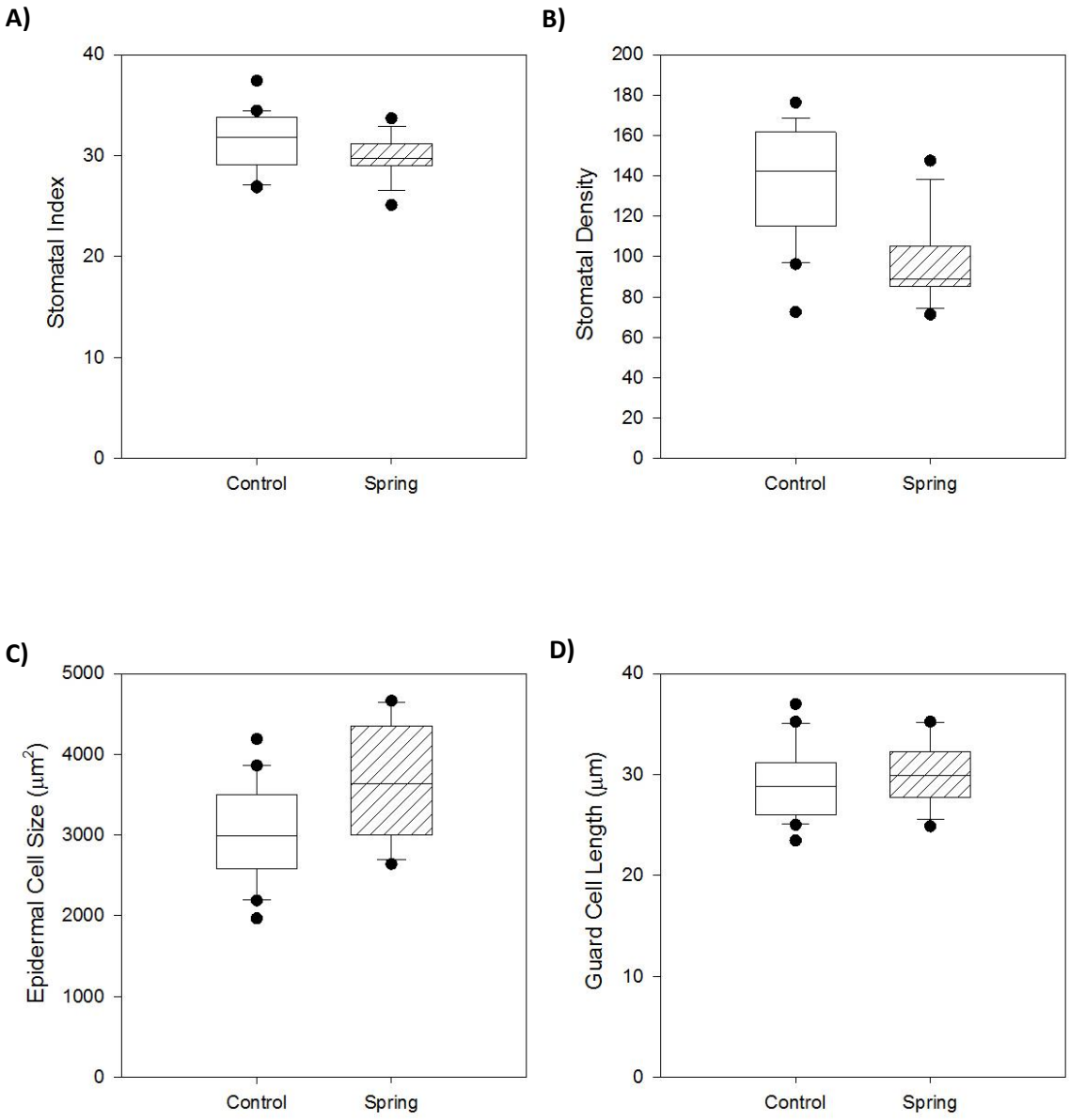



Figure 2.6 – The data from the stomatal imprints taken *in-situ* from the spring and the control site in Bossoleta for *S. vulgaris*. **A** is the stomatal index, **B** is the stomatal density, **C** is the epidermal cell size and **D** is the guard cell length. The central line in each boxplot shows the median. Whiskers indicated the $5^{th}$/$95^{th}$ percentiles. Each dot indicates the observation of an outlier.

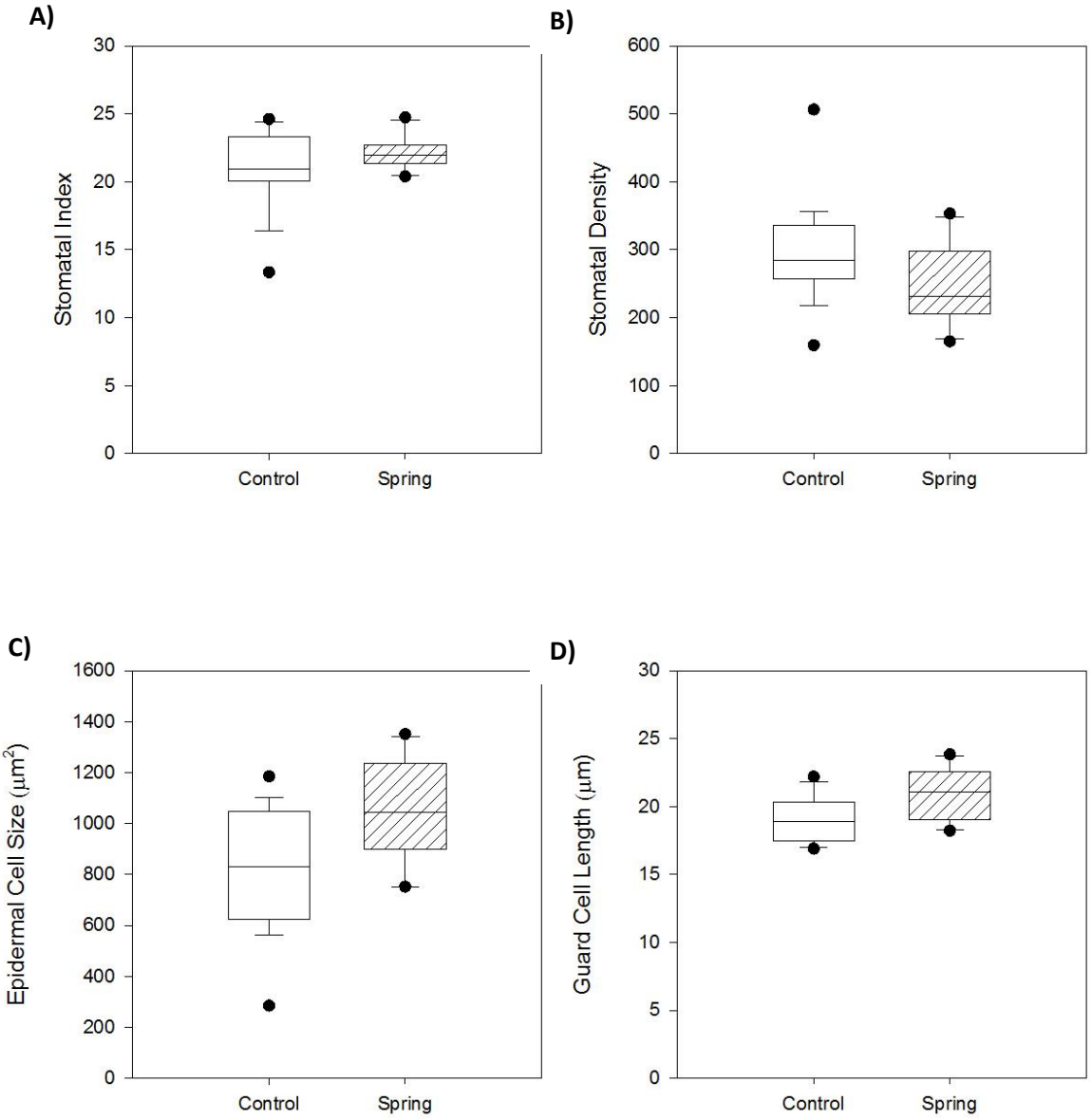



Figure 2.7 – The data from the stomatal imprints taken *in-situ* from the spring and the control site in Bossoleta for *S. minor*. **A** is the stomatal index, **B** is the stomatal density, **C** is the epidermal cell size and **D** is the guard cell length. The central line in each boxplot shows the median. Whiskers indicated the $5^{th}$/$95^{th}$ percentiles. Each dot indicates the observation of an outlier.

Statistics were carried out for all of the traits, for each species. The results can be seen in Table 2.3.

| ***Species*** | ***Stomatal index*** | ***Stomatal density*** | ***Epidermal cell area (µm$^2$)*** | ***Guard cell length (µm)*** |
|---|---|---|---|---|
| *Silene Vulgaris* | | | | |
| Control | 31.396 | 137.14 | 3015.8 | 29.099 |
| Spring | 29.874 | 96.340** | 3664.600** | 30.165 |
| *Sanguisorba minor* | | | | |
| Control | 20.905 | 298.11 | 824.5 | 19.061 |
| Spring | 22.129 | 250.3 | 1046.1* | 20.907* |
| *Plantago lanceolata* | | | | |
| Control | 25.977 | 241.994 | 1278.1 | 16.02 |
| Spring | 26.521 | 264.525 | 1248.4 | 15.865 |

Table 2.3 – Table of the mean values of each trait for each species measured from stomatal imprints taken *in-situ*, and also if the spring plants were significantly different from the control plants for that trait using a one way ANOVA. * P≤0.05; ** P≤0.01.

## 2.4 Discussion

Stomatal density and index often decline in response to elevated $CO_2$, and this is recognised as the most common plant response to elevated $CO_2$ concentrations (Woodward and Bazzaz 1988). Advantages that can arise from a decrease in stomatal numbers include a decrease in transpiration, with no loss of $CO_2$ intake, due to the higher concentrations available (Field, Jackson *et al.* 1995). The phenotypic data presented here indicates a unique counter intuitive adaptation/acclimation in response to elevated $CO_2$ concentrations, which may represent a genuine adaptation to multi-generational exposure to elevated $CO_2$. Stomatal imprints revealed a percentage change increase of 5.21% in stomatal index (SI), and 14.50% in stomatal density (SD) in spring plants in elevated, compared to ambient $CO_2$. There was a significant location (spring or control plant), treatment (ambient or elevated $CO_2$ concentration) and interaction effect for SI ($P < 0.01$), whereas SD only exhibited a significant treatment effect ($P < 0.01$). Tukey's HSD test confirmed the spring elevated plants had increased stomatal index compared to the spring ambient plants ($P < 0.05$). This agrees with the responses of *P. lanceolata* originating from natural springs previously shown (Bettarini I., Vaccari F.P. *et al.* 1999, Woodward 1999, Marchi, Tognetti *et al.* 2004), where no significant difference (trending to an increase) or a significant increase in stomatal index was found in response to elevated $CO_2$. There was no significant difference in stomatal size, and so the increase in stomatal numbers, appears to not be counter-set via stomatal size. The increase in stomatal index in spring plants under elevated $CO_2$ concentrations appears not to provide any immediate advantages, without corresponding disadvantages, yet there must be a reason for the adaptation/acclimation. *P. lanceolata* is not the only species to exhibit this phenotype when sampled from a naturally elevated $CO_2$ spring (Paoletti, A. Raschi *et al.* 1997), supporting that this response is not an anomaly, and what we are seeing here is a true multi-generational adaptation to elevated $CO_2$.

Single leaf area and dry mass both showed a significant effect ($P <0.05$) for location, and single leaf area also had a significant interaction effect ($P <0.05$). The percentage change in single leaf area for CA/CE was a 12.23% increase, whereas the spring plants decreased by 5.51% in response to elevated $CO_2$. There were however, no significant changes in specific leaf area. There were no significant changes in above ground biomass either, but the control plants did double the percentage increase in response to elevated $CO_2$ produced by the spring plants. An increase in biomass in the control plants would be expected, as it is the 'predictable response' to elevated $CO_2$ concentration (Curtis and Wang 1998). It has also been previously shown in *P. lanceolata* originating from ambient conditions, under elevated $CO_2$ (Klus, Kalisz *et al.* 2001). On the contrary, a reduction is biomass has previously been identified by Woodward (1999) in plants originating

from natural springs when exposed to elevated $CO_2$, compared to ambient, including *P. lanceolata.* On the whole, the results have shown that the spring plants exhibit the opposite leaf phenotypic response to elevated $CO_2$ to the control plants, with the spring plants decreasing their leaf size and increasing stomatal numbers, while the control plants increase their leaf size and decrease stomatal numbers. Although this trend was not predicted, results indicate an adaptation has occurred within the spring plants.

The *in-situ* measurements differed in stomatal density and index in *P. lanceolata* in the spring compared to the control site (a 9% and 2% increase respectively), but the differences were not significant. This could be due to the low number of replications, but due to restrictions in number of plants obtainable the low number of replications could not be avoided. The results from the *S. minor* and *S. vulgaris* both showed a decrease in stomatal density and index in the spring site compared to the control site, with the *S. vulgaris* showing a significant decrease in stomatal density ($P < 0.01$). From previous experiments from the spring, all three of the species showed increases in stomatal index, but decrease in stomatal density, however no changes were significant (Bettarini, Vaccari *et al.* 1999). The *P. lanceolata* changes agree with the stomatal index response, but from both the chamber experiment and in-situ experiment we see increases in stomatal density as well. The changes in *S. vulgaris* tie in with the previous research perfectly (Bettarini, Vaccari et al. 1999), but the *S. minor* shows a decrease in both density and index. As the changes were non-significant not many conclusions can be drawn from these similarities.

The changes in *P. lanceolata* could be explained by an exhaustion of adaptation to elevated $CO_2$, which has left only plants that have lost their ability to respond to elevated $CO_2$ concentrations by one or multiple selective sweeps experienced by the plants as they inhabited the spring site. This would cause the plant to create more stomata as it is making more assimilates from photosynthesis, driving cell division at a higher rate with no control of stomatal numbers due to a lack of response, resulting in high stomatal density and index. This explanation could have occurred from an original selective sweep selecting for plants which adapted to elevated $CO_2$ concentrations, leaving a population of plants adapted to live in those conditions. Then the plants began to compete within themselves, adapting in new ways, subsequently losing their ability to further adapt to elevated $CO_2$ concentrations as that had already being exhausted and so not necessary. The next selective sweep then selects for the plants which have found improved ways to survive and without the unnecessary ability to adapt to elevated $CO_2$, leaving the population we see now, which no longer has the ability to adapt to elevated $CO_2$ concentrations. Studies have shown that after only a short time (in comparison to the time spent in the $CO_2$ spring), plants can

exhibit a lack of photosynthetic or stomatal regulation after prolonged exposure to elevated $CO_2$. After only 9 years, this lack of regulation was shown in alpine trees *Larix decidua* and *Pinus mugo* (Streit, Siegwolf *et al.* 2014). Stomatal conductance was insensitive to changes in $CO_2$ in the tree needles, thus transpiration rates remained unchanged and intrinsic water-use efficiency increased due to a higher photosynthetic rate. Similar results have been found in relation to morphology (Steinger, Stephan *et al.* 2007, Newingham, Vanier *et al.* 2013). Newingham *et al.* (2013) found that after 10 years of exposure to elevated $CO_2$, species in a desert ecosystem showed no significant increase in biomass, compared to ambient $CO_2$ controls. It was suggested that this could be due to lack of water in the desert environment, but wet seasons did occur and increases in biomass were still not apparent (Newingham, Vanier et al. 2013). Steinger *et al.* (2007) again found no increase in biomass using *Bromus erectus*, after 7 years of exposure to elevated $CO_2$, suggesting this is a genuine response. If a lack of regulation can happen after such a short periods of time, it is feasible that after thousands of years of exposure to elevated $CO_2$ this could have occurred, to a greater extent, in *P. lanceolata*.

It could also be due to a maximisation of photosynthesis, as increased stomata can lead to a much greater intake of $CO_2$, leading to higher levels of photosynthesis if water loss is not an issue for the plants. The high $CO_2$ spring is a fairly humid environment with reasonable soil moisture, and these plants could therefore be adapted for maximum $CO_2$ gain rather than water use efficiency. An interesting test would be to consider the phenotype in water stressed conditions. If this was the case, you would expect greater biomass in the spring plants, but no significant difference in above ground biomass was observed. The root : shoot ratio has been shown to be affected under elevated $CO_2$ (Rogers, Peterson *et al.* 1992, Woodward 1999), with plants originating from natural springs shifting more biomass to roots under elevated $CO_2$. Larger roots give potential access to greater nutrients in the soil if carbon is in excess, and could be where the spring plants are locating any extra resources. Woodward (1999) showed a shift in the root : shoot ratio towards increased root size in *P. lanceolata* originating from the spring compared to a control site, but also showed that total biomass was decreased. If this was an example of extra resources were being utilised, you would expect an increase in total biomass.

Although the results appear interesting we cannot just assume our data is correct. The phenotypic data appears to be a good representation, but the statistics show some issues, such a significant chamber effect across several measurements. This should not occur as the plants were randomised between the chambers throughout the course of the experiment. This needs to be evaluated to ensure the effect that is causing this significance does not influence the results in a

way that would negate the statistical power of the other results from the model. A significant family effect was also noted in several of the traits. Previous work with *P. lanceolata* has noted family affects during their experiments (Wulff and Alexander 1985, Klus, Kalisz *et al.* 2001), and it appears to be uncontrollable, as using only one maternal plant would negate the variation and give biased results.

It is impossible to conclude just from the phenotypic data the reasoning behind the phenotypic differences noted. The pathways responsible, and the pathways that the phenotype may be a consequence of are of interest as the phenotype appears counter-intuitive, and so investigations into the expression of the pathways could enlighten us to the reasoning behind the phenotype. Although it would appear to be adaptive or acclamatory, there is no way to rule out response being epigenetic. Research into the genetic code, and expression of genes related to the phenotype could elucidate these aspects to give a more comprehensive understanding of the phenotypes observed.

# Elucidating Adaptive Responses to Elevated $CO_2$ Using RNA-Seq

# 3.1 Introduction

The previous chapter highlighted some interesting responses to elevated $CO_2$ on a phenotypic level, but understating the genetic basis at both sequence and expression level of these responses is vital to elucidate them. Using RNA-Seq both of these aspects can be investigated as a profile of gene expression across the genome and the sequences of these genes can be produced.

When using transcriptomics careful experimental technique and planning are essential as RNA can be lowly expressed and is easily degraded. The field of transcriptomics has become very advanced, and there are techniques which allow for large amounts of RNA to be successfully extracted from plant tissue (Logemann, Schell *et al.* 1987). When extracting RNA, conditions must be considered, such as extracting in two different conditions such as light and dark, or stressing one plant more than another can have massive effects on RNA expression. This can lead to noise as there may be large expression differences due to poor experimental planning, and could also lead to certain differences being missed. Thus consistency in time of day for sampling fresh material, time between sampling and freezing must be strictly adhered to. After extraction, analysis of the RNA needs to be carried out to assess the integrity of the RNA. This is vital in obtaining the best possible data, as low-quality RNA can strongly compromise results of downstream experiments (Fleige and Pfaffl 2006). Techniques such as RNA quantification ensure RNA has been extracted and ascertain the quality of the sample. This can be done using analysis systems such as a NanoDrop or Bioanalyser. These issues are very easy to overcome, so to be aware of them usually means they are accounted for.

## 3.1.1 RNA-Seq

### *3.1.1.1RNA-Seq Methods*

A number of RNA-Seq technique can be used, all based on clonal template sequencing and multi-parallelization. Some techniques use sequencing by synthesis (SBS), such as Illumina and 454 sequencing, whereas others use sequencing by ligation (SBL), such as SOLID sequencing. Table 3.1 shows examples of the different platforms for each technique.

| Platform | Read Length (bp) | Max Number of Reads per Run | Sequencing Output per Run | Run Time |
|---|---|---|---|---|
| Ilumina HiSeq 2000 | 100-200 | $6x10^9$ | ≤ 540-600 Gb | 11 days |
| Illumina MiSeq | 100-150 | $7x10^6$ | ≤ 1-2 Gb | 19-27 hours |
| Roche 454 GS FLX+ | 600-800 | $1x10^6$ | ≤ 700 Mb | 23 hours |
| Roche 454 GS FLX | 400-500 | $1x10^6$ | ≤ 500 Mb | 10 hours |
| AB SOLiD 5500 system | 35-75 | $2.4x10^9$ | ~ 100 Gb | 4 days |
| AB SOLiD 5500 xl system | 35-75 | $6x10^6$ | ~ 250Gb | 7-8 days |

Table 3.1 – The range of platforms that can be used when using RNA-Seq

Illumina sequencing data is highly replicable, with relatively little technical variation. The information gathered from a single lane ('lanes' of Illumina can run completely independently for smaller experiments or in combination for larger genomes) of Illumina sequencing data appears comparable to that in a single microarray which suggests it may suffice to sequence each mRNA sample only once. A single lane of Illumina enables identification of differentially expressed genes, while still allowing for additional analyses such as detection of low-expressed genes, alternative splice variants, and novel transcripts (Marioni, Mason *et al.* 2008).

The process of Illumina sequencing uses multiple single molecules attached to a glass surface which is classed as the 'library.' For all RNA-Seq techniques the construction of this library requires isolation of RNA, random fragmentation of the transcripts into smaller pieces, conversion of the RNA into DNA by reverse transcription, ligation of adapter sequences for amplification, fragment selection size, and priming the sequences reaction (Young, McCarthy *et al.* 2012). The library is amplified to give multiple identical copies of the molecules. This amplification is necessary as it provides a high signal : background ratio. When using Illumina sequencing, the sample provided is broken up using focussed acoustic waves. The sample then undergoes conversion from RNA to DNA and Illumina adapters are ligated onto each end of the DNA fragments. The DNA is then washed across a flow cell. The flow cell is a glass slide on which the library fragments will be sequenced in a multi-parallel fashion. As the fragments are washed across they randomly graft onto the flow cell. Instruments are then used to create clusters on the flow cell via bridge amplification (Figure 3.1). Once the fragments are attached to the flow cell they are extended by polymerases. The double-stranded molecule is then denatured and the original template is washed away leaving the newly synthesised fragment covalently attached to

the flow cell surface. These single-strands flip over to hybridise to adjacent primers to form a bridge, hence the term “bridge amplification.” The hybridised primer is then extended, again using polymerases and a double-stranded bridge is formed. The bridge is then denatured to leave two copies of covalently bound single-stranded templates. The two new strands then undergo the same process, and it is repeated until multiple bridges are formed and eventually leaving a cluster with forward strands only. This bridging technique allows for paired-end sequence reads, which refers to both ends of the same DNA molecule being sequenced. In addition to the more comprehensive sequence information reading from both ends gives, both reads also contain long range positional information, allowing for highly precise alignment of reads in downstream analysis. Once the clusters are formed sequencing by synthesis is used. Sequencing primers are then annealed to the clusters and free bases with attached fluorophores are added. Each base (A, C, G and T) are labelled with different fluorophores so they give off a different coloured signal. Bases attach one at a time onto each DNA fragment, and then all the unincorporated bases are washed away. The flow cell is imaged after each base is added. The amount each of the sequences occurs gives a measurement of their expression (Figure 3.1).

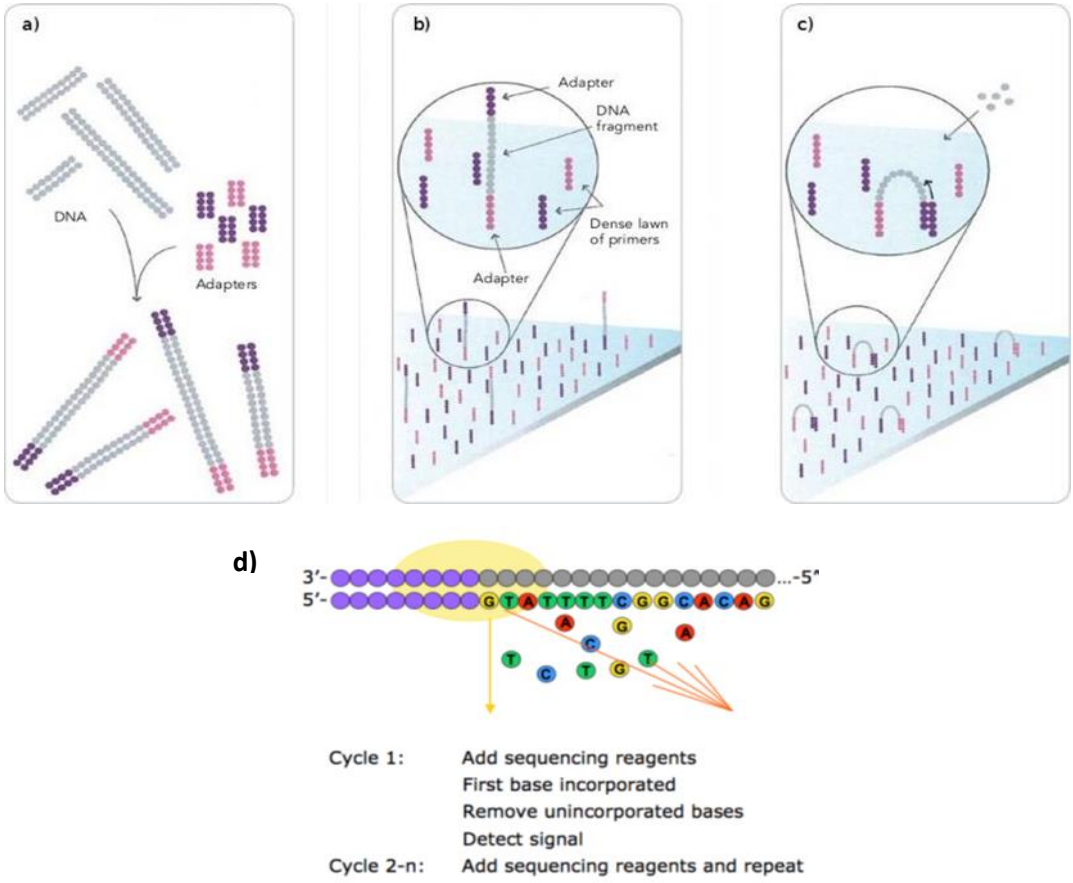



Figure 3.1 - Overview of the Illumina RNA-Seq process . (**a**) Represents the Illumina adaptors been ligated onto the DNA. (**b**) Represents the DNA grafting onto the flow cells by attaching to the lawn of primers. (**c**) Represents the start of bridge amplification, a process that happens multiple times to create the clusters on the flow cell. (**d**) Represents the bases attaching to the clusters and the light signals being read one by one as each base attaches. (Images taken from presentation given by IGA, http://www.igatechnology.com/about-us/)

#### *3.1.1.2 Sequence Assembly*

Sequence assembly refers to alignment and merger of fragments of a much longer DNA sequence in order to reconstruct an original sequence. This is needed as sequencing technology such as RNA-Seq reads small pieces of between 20 and 1000 bases which then need to be assembled. When a reference genome is available sequence assembly is made simpler, as fragments are mapped onto the reference genome. Even with a reference genome this is still not a straightforward task as sequence errors and polymorphisms still need to be taken into account, which can lead to large mistakes in the assembly process. Additionally, alternative splicing can make these issues more prominent when using RNA-Seq. Sequence assembly becomes substantially more difficult when no reference genome is available, but the ability to assemble without a reference genome is still a massive advantage over all other methods. Without a reference genome, short reads of sequence for the RNA-Seq data have to be assembled using *de novo* assembly. *De novo* assembly uses algorithms to assemble the short reads, generally based around either two basic algorithms: overlap graphs and de Bruijn graphs.

Overlap graph-based assemblers are the most established as they were developed for Sanger sequencing reads. They compute all pair-wise overlaps between the reads and capture the information in a graph. The overlap graph is used to compute a layout of reads producing a consensus sequence of *contigs*: a contiguous sequence reconstructed from reads. It is possible to assemble RNA-Seq data using overlap graphs, but even simple organisms produce millions of short reads, and therefore creating a graph for a large number a reads can be computationally difficult. Even so, there are still a number of whole genome assembly programmes which use overlap graphs for alignment including, ATLAS (Havlak, Chen *et al.* 2004), ARACHNE (Batzoglou, Jaffe *et al.* 2002) and PCAP (Huang, Wang *et al.* 2003). Due to the low efficiency of overlap graphs with a high number of short reads, most assemblers for RNA-Seq use de Bruijn graphs. De Bruijn graphs reduce the computational effort by breaking reads in to smaller sequences of DNA called k-mers. K-mers use the parameter k to denote the length in bases of the sequences. They can use a number of methods within de Bruijn assembly, including using either Hamilton or Eulerian cycles, but it is the Eulerian method which is the most efficient and is used by most assemblers using de Bruijn graphs (Figure 3.2). Using de Bruijn graphs has a number of other advantages. By reducing the entire data set down to k-mer overlaps, the redundancy is reduced in short-read data sets. De Bruijn graphs also allow repeats in the genome to be collapsed in the graph leading to less spurious overlaps, although this does not mean they can be more easily bridged or resolved.

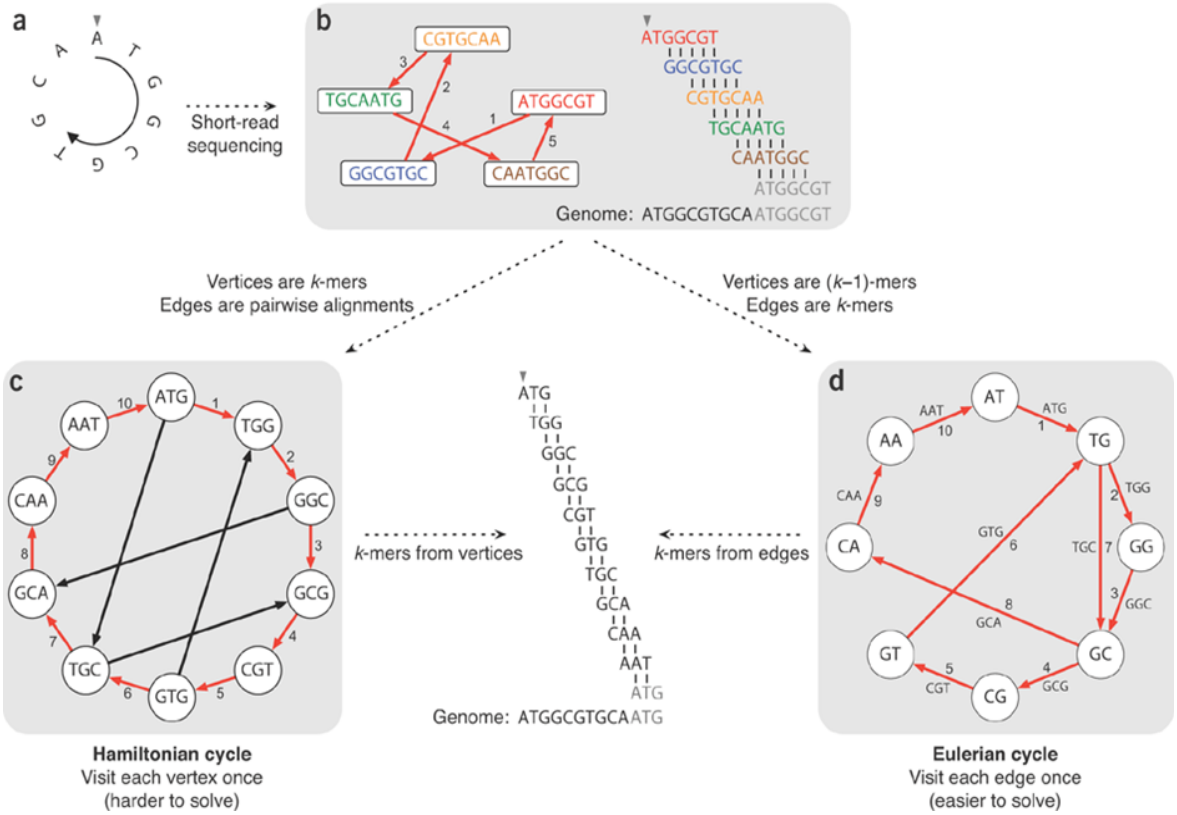



Figure 3.2 – (**a**) An example small circular genome. (**b**) Reads are represented as nodes in a graph, and edges represented alignments between reads. Following the edges around in a numerical order allows you to reconstruct a circular genome by combining alignments between each successive read. The repeated part of the sequence is grayed out in the alignment diagram. (**c**) An alternative de Bruijn assembly technique first splits reads into all possible k-mers: with k = 3, ATGGCGT comprises ATG, TGG, GGC, GCG and CGT. If you the follow this with a Hamilton cycle (Following a Hamiltonian cycle (a path in an undirected graph that visits each vertex exactly once indicated by red edges) allows one to reconstruct the genome by forming an alignment in which each successive k-mer (from successive nodes) is shifted by one position. This method recovers the genome but issues arise when using a large number of reads as it does not scale well to large graphs. (**d**) Modern short-read assembly algorithms construct a de Bruijn graph by representing all k-mer prefixes and suffixes as nodes and then drawing edges that represent k-mers having a particular prefix and suffix. For example, the k-mer edge ATG has prefix AT and suffix TG. Finding an Eulerian cycle (a trail in a graph which visits every edge exactly once) allows one to reconstruct the genome by forming an alignment in which each successive k-mer (from successive edges) is shifted by one position. This generates the same cyclic genome sequence without performing the computationally expensive task of finding a Hamiltonian cycle (Compeau, Pevzner *et al.* 2011).

The field of short read *de novo* assembly have been developed from pioneering work on de Bruijn graphs by Pevzner *et al.* (Pevzner and Tang 2001, Pevzner, Tang *et al.* 2001). They adapted a fundamentally different approach using de Bruijn graphs called the EULER assembler (Pevzner, Tang *et al.* 2001).This de Bruijn graph representation is prevalent in many current short read assemblers, such as EULER+ (Pevzner, Tang *et al.* 2004), Velvet (Zerbino and Birney 2008), ALLPATH (Butler, MacCallum *et al.* 2008) SOAP (Li, Fan *et al.* 2010) and ABySS (Simpson, Wong *et al.* 2009). ABySS was developed to assemble the very large data sets produced by sequencing the human genome which shows the assembling capacities these programmes have. The innovation of AbySS in particular is it provides a distributed representation of a de Bruijn graph, therefore allowing parallel computation of the assembly algorithm across a network of commodity computers (commodity computing is the use of large numbers of already available computing components for parallel computing to get the largest amount of useful computation at low cost). ABySS is particularly useful when sequencing organisms for which no reference genome is available which is the case with *P. lanceolata*. CLCBio is another *de novo* assembler algorithm that has more recently emerged and improved the standards of *de novo* assembly of large genomes, by enabling *de novo* assembly of gigabase-size genomes on a single desktop computer, in a reasonable time period. Another more recent advancement in *de novo* assembling technology was the release of a programme called Trinity (Grabherr, Haas *et al.* 2011). Compared with other *de novo* transcriptome assemblers, Trinity recovers more full-length transcripts across a broad range of expression levels, with sensitivity similar to methods that rely on genome alignments but without a reference genome. The programme claims to provide a unified solution for transcriptome reconstruction in any sample, especially in the absence of a reference genome. Trinity is presented as a method for the efficient and robust *de novo* reconstruction of transcriptomes, consisting of three software modules: Inchworm, Chrysalis and Butterfly, applied sequentially to process large volumes of RNA-Seq reads (Figure 3.3).

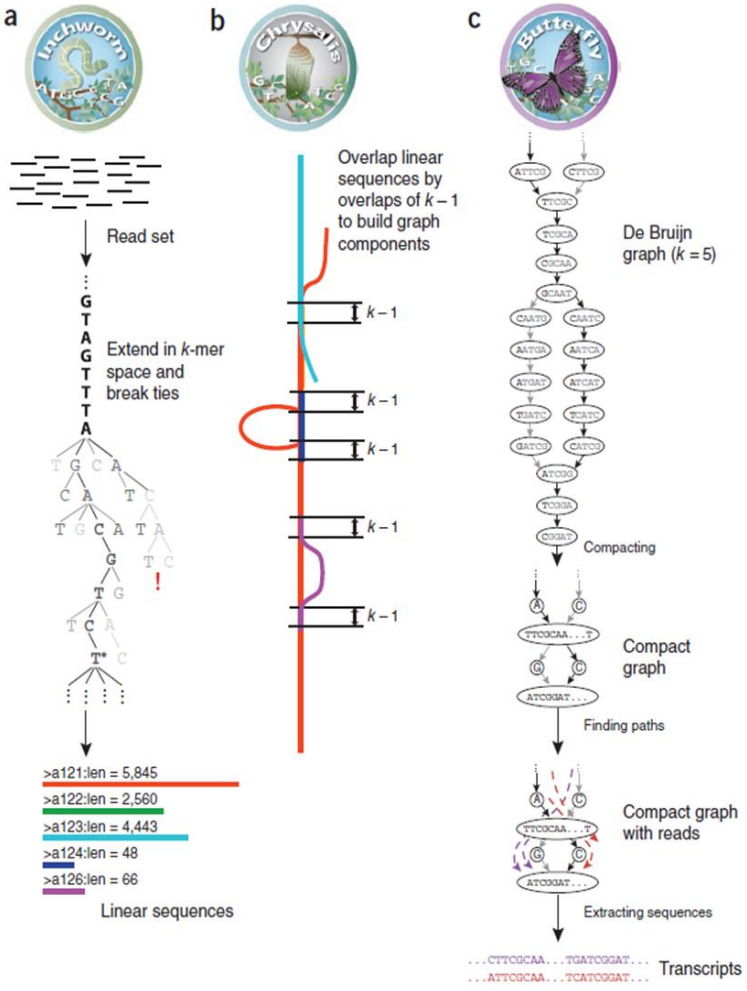



Figure 3.3 - Overview of Trinity (Grabherr, Haas *et al.* 2011) - (**a**) Inchworm assembles the read data set (short black lines, top) by greedily searching for paths in a k-mer graph (middle), resulting in a collection of linear contigs (colour lines, bottom), with each k-mer present only once in the contigs. (**b**) Chrysalis pools contigs (coloured lines) if they share at least one k – 1-mer and if reads span the junction between contigs, and then it builds individual de Bruijn graphs from each pool. (**c**) Butterfly takes each de Bruijn graph from Chrysalis (top), and trims spurious edges and compacts linear paths (middle). It then reconciles the graph with reads (dashed coloured arrows, bottom) and pairs (not shown), and outputs one linear sequence for each splice form and/or paralogous transcript represented in the graph (bottom, coloured sequences).

Grabherr *et al*. (2011) compared Trinity to other *de novo* assemblers currently available using a number of example organisms. The example of most interest due to using *P. lanceolata* is the assembly of the white fly transcriptome. This is due to the lack of a reference genome of the white fly, therefore having the same potential difficulties as *P. lanceolata*. The whitefly (*B. tabaci*) genome is not sequenced, and the RNA-Seq samples are genetically polymorphic, as they are

derived from a mixture of individuals from an outbred population adding complications such as sequence variation between samples. The figure below shows the results (Figure 3.4);

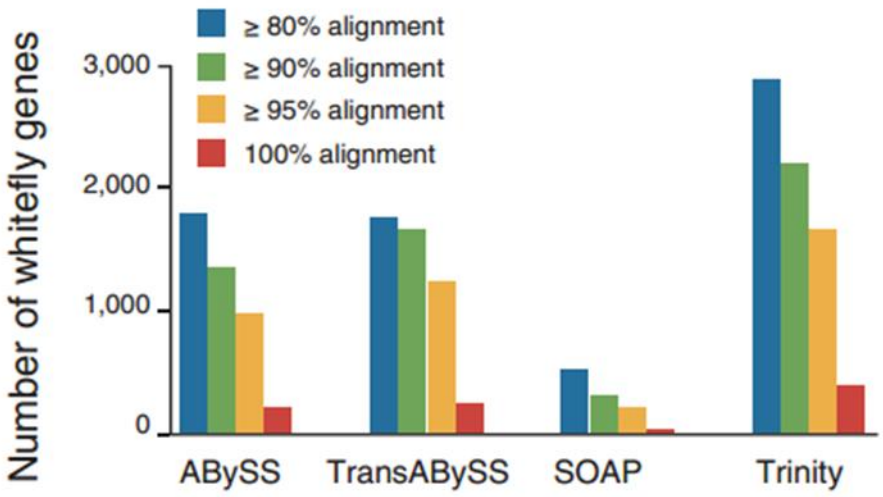



Figure 3.4 - Figure taken from Grabherr *et al*. (2011). The y axis is a count of the unique top-matching (BLASTX) protein sequences aligned Trinity transcripts across a minimal percent of their length. This shows Trinity's performance is considerably higher, not only in assembling with 100% alignment, but with its overall performance with almost 1000 more matching genes being aligned.

It is clearly shown that Trinitys performance is considerably better than ABySS, TransABySS (an analysis pipeline for post-processing ABySS assemblies of transcriptome sequencing data) and SOAP. Figure 3.4 shows the number of approximately full-length Trinity-assembled transcripts is substantially higher than achieved by other *de novo* assemblers. This makes Trinity appear to be the number one choice when using *de novo* assembly.

#### *3.1.1.3 Data Analysis*

As yet, a rigorous RNA-Seq data analysis protocol has not been developed and we are still in the stages of exploring the features of the data. Due to this, there are many published papers which use various methods of data analysis, including different methods of normalisation and different pieces of software. The processed reads from RNA-Seq provide information on contigs, which are sets of overlapping DNA segments that together represent a consensus region of DNA. The

sequence and expression levels of these contigs are provided in the raw data which can be used for differential expression analysis. Issues have arisen with the raw data as a number of biases have been identified producing irregularities in results that do not take them into account.

One bias is the number of extra reads produced by longer genes, which may skew the expression value when comparing to shorter genes. To overcome this, current RNA-seq analysis methods typically standardise data within samples by scaling the number of reads in a given lane or library to a common value across all sequenced libraries in the experiment. This reduces values if sequences have been read multiple times so all expression values are taken from a similar number of reads allowing alleviation of the bias. One method which is widely used, is to normalise by adjusting counts using the reads per kilobase of exon per million mapped (RPKM) model (Mortazavi, Williams *et al.* 2008). The formula that defines RPKM is;

$$\text{RPKM} = \frac{\text{Total Exon Reads}}{\text{Mapped Reads (Millions)} \times \text{Exon Length (KB)}}$$

Formula 3.1 – Used to calculate the RPKM value for normalisation of NGS data.

RPKM can also be referred to as fragments per kilobase of exon per million mapped (FPKM). Scaling to library size as a form of normalization makes intuitive sense, given it is expected that sequencing a sample to half the depth (on average) will give half the number of reads mapping to each gene. Due to this it is believed this is appropriate for normalizing within replicate samples of an RNA population. However, issues arise as library size scaling is too simple for many biological applications such as comparisons between two different conditions (Wagner, Kin *et al.* 2012). The number of tags expected to map to a gene is not only dependent on the expression level and length of the gene, but also the RNA composition of the population that is being sampled. This means if a large number of genes are unique to, or highly expressed in, one experimental condition, the sequencing 'real estate' available for the remaining genes in that sample is decreased. If this is not adjusted for, this approach can force the differential expression analysis to be skewed towards one experimental condition. A slight modification to RPKM has been suggested which increases the effectiveness of RPKM with an additional calculation which takes into account the total expression of each sample to remove the bias, and allow for comparisons

between conditions (Wagner, Kin *et al.* 2012). This modification is known as TPM (transcripts per million).

Transcript length, paralogous gene families, low-complexity sequence and high sequence similarity between alternatively spliced isoforms are all factors that can lead to *multireads*: reads that have high-scoring alignments to multiple positions in a reference genome or transcript set (Mortazavi, Williams *et al.* 2008). Multireads have implications for the ranking of differentially expressed genes, and, in particular, may introduce bias in gene set testing for pathway analysis and other multi-gene systems biology analyses (Robinson and Oshlack 2010). A previous study showed 8-19% of the genes with multiple known transcript isoforms, expressed at least two isoforms in the same blastomere or oocyte, which unambiguously demonstrates the complexity of the transcript variants at whole-genome scale in individual cells (Tang, Barbacioru *et al.* 2009). At first the strategy to deal with these multireads was to discard them, keeping only uniquely mapped reads for expression estimation (Nagalakshmi, Wang *et al.* 2008). This method was wasteful as the multireads could contain interesting expression data. A strategy used to overcome this is to uniformly divide each multi-mapped read to all of the positions it maps to, so a read mapping to 10 positions will count as 10% of a read at each position. This alone can create biases, so these values are then normalised to the abundances, dividing each multi-mapped read probabilistically based on the initial abundance estimation of the genes it maps to, the inferred fragment length, and fragment bias (Trapnell, Williams *et al.* 2010).

The most common method to deal with both multiread and library scaling bias is to use specific software packages that take the bias into account as part of the algorithm they use to analyse the data. A package which has had good reviews dealing with the issue that library scaling provides is *DEGSeq* (Wang, Feng *et al.* 2010). *DEGSeq* does this by considering the number of total reads that map to the gene exon regions when the sequence depth between samples is not the same. *DEGSeq* is run using R, and also employs another R based package *samr* (Li and Tibshirani 2011) that runs within *DEGSeq*. This allows for comparisons of two sets of samples with multiple replicates or two groups of samples, from different individuals (e.g. disease sample verses control sample). The package implements the method described in Tusher *et al.*(2001), who created methods to determine the significance of differences in expression between different biological states in microarrays. The package *samr* assigns a score to each gene on the basis of change in gene expression relative to the standard deviation of repeated measurements and uses permutations of the repeated measurement to estimate FDR (false discovery rate). Another package available is *edgeR* (Robinson, McCarthy *et al.* 2010), which is also run using R and

considers the same issues in its algorithm and so gives similar results to *DEGSeq*. The difference between the two packages is that *DEGseq* is based on a Poisson distribution while *edgeR* is based on negative binomial distribution. The use of Poisson distribution has been criticised, as it has been noted that the assumption of Poisson distribution is too restrictive: it predicts smaller variations than what is seen in the data (Anders and Huber 2010). Therefore, the resulting statistical test does not control type-I error (the probability of false discoveries) as advertised. This gives the presence of greater variability in a data set than would be expected based on a given simple statistical model, and is referred to as over dispersion. A basic property of Poisson distribution is the equality of mean and variance. If variance is larger than mean, then the data are said to be over dispersed, and the Poisson assumption is inappropriate. The use of negative binomial distributions addresses the so-called over dispersion problem making *edgeR* a more suited choice for data analysis. Although this is true, studies have shown it to not be true in every situation. If apparent over dispersion results from specification errors in the systematic part of the Poisson regression model, resorting to negative binomial distribution could potentially make the situation worse by giving a false sense of security when the fundamental errors in the model remain (Berk and MacDonald 2008).

Another issue which needs to be dealt with is fragment bias. Current technological limitations of NGS require that the cDNA molecules represent only partial fragments of the RNA being probed (Roberts, Trapnell *et al.* 2011). Recent careful analysis of this has revealed both positional and sequence-specific biases in sequenced fragments. Positional bias refers to a local effect in which fragments are preferentially located towards either the beginning or end of transcripts. Sequence-specific bias is a global effect where the sequence surrounding the beginning or end of potential fragments affects their likelihood of being selected for sequencing. These biases can affect expression estimates, and it is therefore important to correct for them during RNA-Seq analysis if possible (Li, Jiang *et al.* 2010, Roberts, Trapnell *et al.* 2011). Again there is a piece of software which deals with this bias, called Cufflinks (Trapnell, Williams *et al.* 2010). The software deals with this bias by 'learning' which sequences are being selected for, and taking this into account when calculating abundance. As the bias is usually caused by primers used either in the PCR or reverse transcription step, it will appear near the ends of the sequenced fragments as suggested before. They suggest they have developed a method to correct the bias by 'learning' what sequences are being both selected for and ignored in a given experiment, and including these measurements in the abundance estimation.

To do this they first of all generate initial abundance estimations without using bias correction. Since different transcripts contain different sequences, it is possible to approximate abundance to

weight reads by the expression level of the transcript from which they arise. This also helps avoid over-counting sequences that may be common in the mapping data due to high expression rather than a bias, and deals with potential multiread and library scaling biases. Next they revisit each fragment in the alignment file and apply the abundance weighting as they 'learn' features of the sequence in a window surrounding the 5' and 3' end of the transcript using a graphical model of the statistical dependencies between bases in the window. A separate model is kept for each end of the read since the biases in the first and second strand synthesis of the fragment are not always the same.

Finally, the abundances are re-estimated using a new likelihood function that has been adjusted to take the sequence bias into account, based on the parameters of the graphical model computed in the previous step. The result is a new set of FPKMs that are less affected by sequence-specific bias (Roberts, Trapnell *et al.* 2011). Cufflinks demonstrates its ability to deal with both fragment and library scaling biases, and also has accompanying programmes such as *Cuffcompare* and *Tophat* which can be used to perform extra analysis pre and post *Cufflinks*, making it a very capable piece of software when dealing with RNA-Seq data. Issues with cufflinks arise when there is a lack of a reference genome for the software to map onto. *De novo* assembled sequences can be a lot harder to map on to than a complete reference genome, as *de novo* assembly does not give a complete sequence, and this can cause problems for the software. The assembly programmes do advise programmes which could be used with the processed data. Trinity suggests *edgeR* as a suitable programme for analysing the data, meaning *edgeR* is compatible with *de novo* assembled data.

Attempts have been made to put *de novo* assembled data in to programmes such as Cufflinks. Using paired-read sequencing data it is possible to assess the order, distance and orientation of contigs and combine them into so-called 'scaffolds' (Boetzer, Henkel *et al.* 2011). These scaffolds can be used with programmes such as cufflinks as they are much longer, around 20 Kbp. Illumina sequencing produces pair end reads which can be put into programmes such as *SSPACE* (Boetzer, Henkel *et al.* 2011) to scaffold the pair end contigs into scaffolds. There have been limited attempts at this in the literature as it would appear to still be very difficult due to complications with inputting the data into programmes. This is because scaffolds are still not fully designed to be inputted into programmes such as *Cufflinks*, and so numerous modifications have to be made to the SAM (Sequence Alignment/Map Format) files before accurate results can be obtained.

Another method some people have used to overcome the lack of a reference genome is to use a homologous genome. If a genome has a high homology it is possible to map onto that genome to

build sequences and identify genes (Adhikari, Wall *et al.* 2009). This is useful as potential genes of interest can be mapped to give a better idea of the role of the gene. As previously mentioned, when the olive genome is sequenced this should have a good homology with *Plantago* due to its close taxonomy. Also the *Mimulus guttatus* is currently being sequenced, and is already assembled to a decent degree which is also close to the Plantago genome and so can be used to look at specific sequences that have already been sequenced in the *M. guttatus* genome. As the model plant organism, *Arabidopsis thaliana* provides a fully annotated sequence, but due to its distance taxonomically only genes with conserved regions are likely to map on.

NOISeq (Tarazona, García-Alcalde *et al.* 2011) is a piece of software which differs from existing methods in that it is data-adaptive and nonparametric. Most existing methodologies suffer from a strong dependency on sequencing depth for their differential expression calls, and this causes a considerable number of false positives that increases, as the number of reads grows. In contrast, NOISeq models the noise distribution from the actual data, and can therefore better adapt to the size of the data set, and is more effective in controlling the rate of false discoveries. (Tarazona, García-Alcalde *et al.* 2011). The stringency of this method will in turn lead to less significant results and may results in some interesting results being missed, but this method along with another is a good way to confirm any significant differences found, and to give the results more power.

Figure 3.5 is an overview of all the NGS options covered, with a suggested protocol to follow when using NGS to analyse mRNA data.

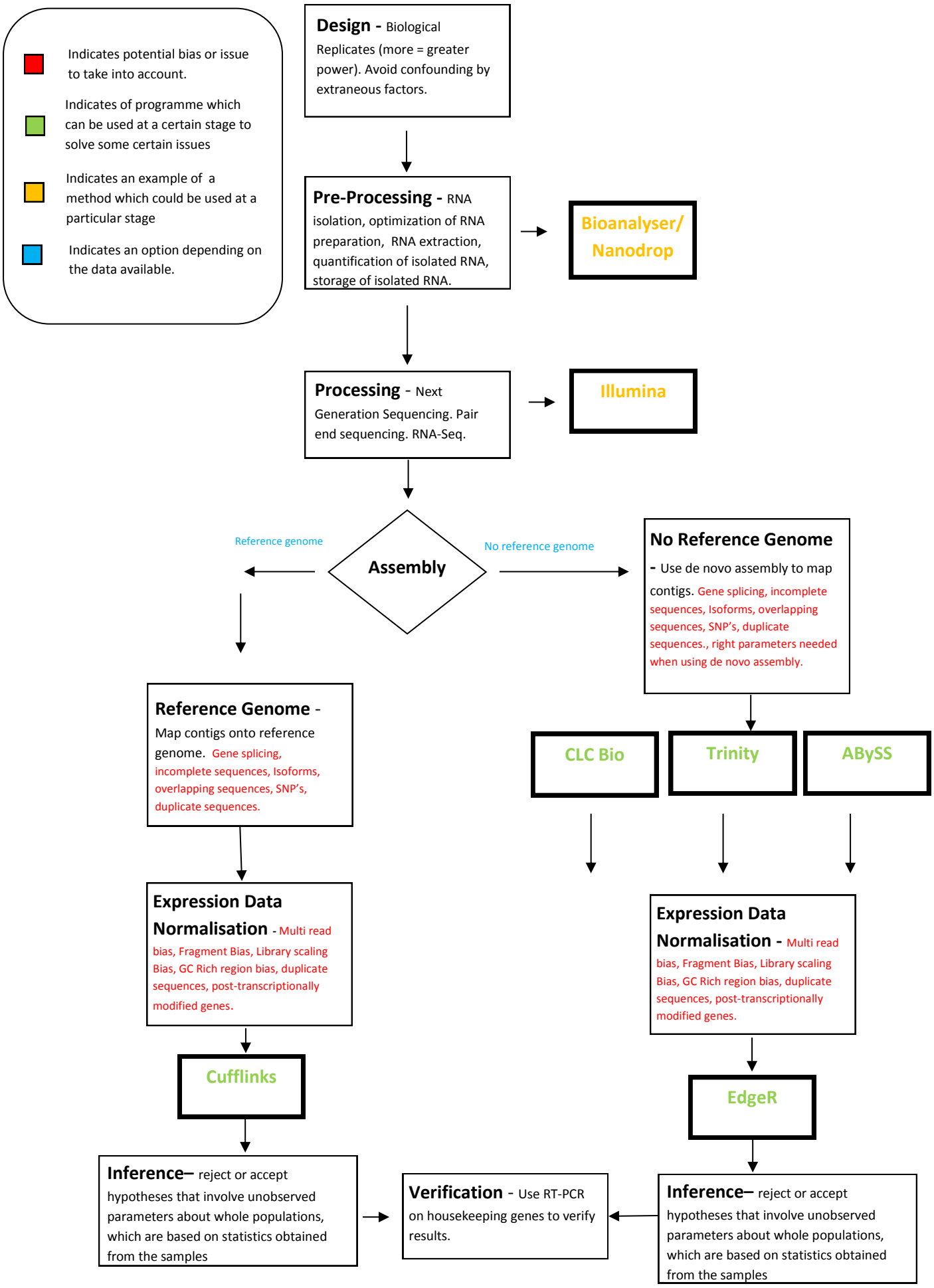



Figure 3.5 - A flow chart of the steps of mRNA analysis when using next generation sequencing.

### 3.1.2 Aims

- Use RNA-Seq to identify the genomic impacts of multi-generational exposure to elevated $CO_2$ on *P. lanceolata*.

- Compare how the transcriptomic profile of plants adapted to elevated $CO_2$ compares to the profile of plants acclimated to elevated $CO_2$.

- Investigate atmospheric $CO_2$ as a selective agent driving directional micro-evolution.

## 3.2 Methods

RNA was extracted from all samples from the chamber experiment (Chapter 2) from *P. lanceolata* leaf tissue 58 days after establishment by Dr Yunan Lin (previous PhD student at the University of Southampton) using a CTAB-based protocol modified from (Chang, Puryear *et al.* 1993). 900 µl of pre-warmed (65 ºC) CTAB ( 2% hexadecyltrimethylammonium bromide (weight (w)/v), 2% polyvinylpyrrolidone (w/v), 100 mM Tris-HCl (hydrochloric acid), 25 mM EDTA Ethylenediaminetetraacetic acid and 2 mole (M) NaCl (sodium chloride)) and 2 µl β-nercaptoethanol was added to each Eppendorf tube, and tubes were incubated at 65 ºC for 5 minutes. After incubation, 900 µl CHISAM (Chloroform/Isoamyl alcohol) was added into each tube and tubes were centrifuged for 10 minutes at 12 x gravity (g). The upper layer (aqueous phase) was transferred to a new tube and 280 µl of 10M lithium chloride was added. After incubation at 4 ºC for 30 minutes, the samples were centrifuged for 15 minutes at 12x g at 4 ºC to form a pellet. The pellet was then dissolved in 700 µl of SSTE ( 11.7g 1M NaCl (w/v), 0.5% SDS, 10mM Tris-HCl (pH 8.0) and 1mM EDTA) which was pre-warmed to 60 ºC and incubated at 60 ºC for 5 minutes. 700 µl of CHISAM was then added and centrifuged at 12x g for 10 minutes at room temperature. The upper layer was transferred to a fresh Eppendorf and 700 µl of 100% ethanol was added. The tubes were incubated at -20ºC for 10 minutes and then centrifuged for another 10 minutes at 4 ºC. The liquid phase was discarded and the pellet was washed with 1 ml of cold 70% ethanol. The samples were centrifuged again and the remaining liquid was discarded. The pellet was air-dried, then re-suspended in 20 µl DEPC (diethylpyrocarbonate) treated water. The RNA concentration was measured using the Nanodrop spectrophotometer (ND100, NanoDrop Technologes, Deleware, USA). For genetic analysis we carried out RNA-Seq on RNA from twenty four samples, six from each sample group. The sequencing was carried out by IGA (Istituto Di Genominca Applicata) who used an Illumina Hi Seq 2000 sequencer. Paired end read sequencing was used, and the sequencing produced and average of 12,831,481 (± 963,285) reads per sample.

### 3.2.1 CLC Bio Assembly

The RNA-Seq data was processed by IGA, the company who carried out the sequencing , who *de novo* assembled the data using CLCbio (See section 1.5.1.2) and also carried out normalisation on the data using RPKM (See section 1.5.1.3). On analysing the total contig expression the control ambient group clearly showed a higher total contig expression than other three groups. This has been identified as a problem in other research (Wagner, Kin *et al.* 2012), and has been identified as a variable that RPKM doesn't necessarily account for. To correct for this each of the RPKM values were divided by the summed expression value of the corresponding group which represented a proportion of the total expression value which is similar to the TPM correction. This corrected for the difference in total contig expression.

Analysis was carried out on the whole data set to check the validity of the data. A principal component analysis (PCA) was carried out to see if the samples correlated with each other across the two main components. Principal component analysis is one of the oldest and best known of the techniques of multivariate analysis, and allows for efficient analysis of variation in a data set, while reducing the dimensionality of that data set if there are a large number of interrelated variables (Jolliffe 2005). This reduction is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables. This analysis (Figure 3.6) showed that the samples separated as you would expect for plants in such different environments, but it was also noted that certain samples separated from the rest within treatment groups.

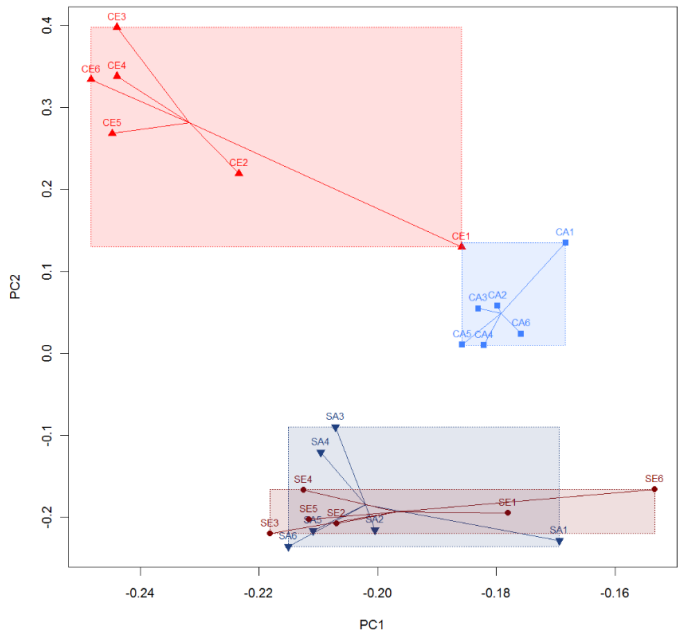



Figure 3.6 – Principal component analysis on the CLC Bio assembled data. PC1 separated the ambient and elevated $CO_2$ treatments whereas PC2 separated the control and spring growth sites. CA is control ambient. SA is spring ambient. CE is control elevated. SE is spring elevated.

The data shows samples SA1, SE1, CA1, CE1 and SE6 separated away from the other samples. This is a concern as the de novo assembly was carried out using samples SA1, SE1, CA1 and CE1 and so could indicate a bias in these samples. To see if this was impacting individual contigs random samples were looked at to see if there were any major differences between the outlying samples and the rest. Figure 3.7 shows an example where the outliers show a clear difference to the rest of the samples.

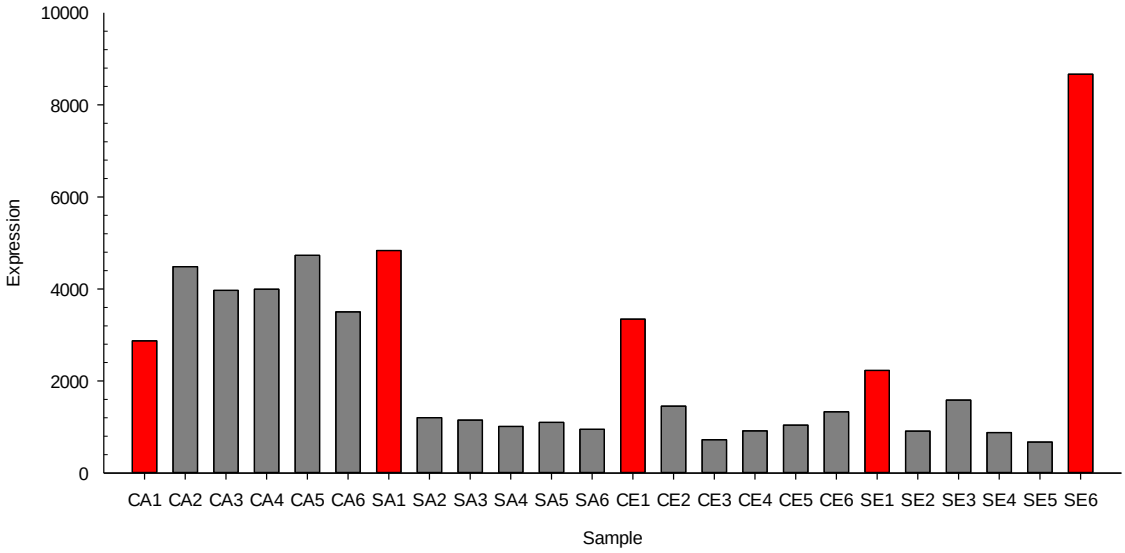



Figure 3.7 – An example of the expression of one of the contigs after the CLC Bio *de novo* assembly. The columns in red are the columns highlighted as potential outliers from the rest of the data.

Some of the samples showed over an 8 fold difference from the average of the rest of the samples. As the *de novo* assembly was made from these samples, you would expect the reads to map back better to these samples, and this could explain the differences in expression in these samples. Sample SE6 was not involved in the assembly it's harder to identify the reason for differences in expression in this sample. It could be a technical fault with the lane during sequencing which caused irregularities. Due to the low number of replicates, to name over 1/6$^{th}$ of the samples as outliers would be cause problems when it came to statistical analysis, and so a second *de novo* assembly was carried out using Trinity to try and improve the assembly.

### 3.2.2 Trinity Assembly

Due to improvements in *de novo* assembly software, Trinity was chosen to carry out the second assembly due to its proven enhanced performance (Grabherr, Haas *et al.* 2011). Ideally all samples would be used for the *de novo* assembly to remove any bias, but due to computing power this was not possible, but an increase to 8 samples was used. The samples chosen were CA1, CA2, SA1, SA2, CE1, CE2, SE1 and SE2. First of all these samples were normalised within Trinity for kmer coverage using the command *normalize_by_kmer_coverage.pl* with maximum coverage set to 30. The resultant libraries were assembled into a *de novo* transcriptome using *Trinity.pl* with minimum kmer coverage set to 2 to reduce the proportion of error-containing kmers, and taking into account orientation of paired reads. This resulted in a transcriptome comprising 154,179 transcripts from 43,908 components. Transcripts (comparable to contigs) represent different forms of the same component and components are loosely comparable to genes, however two components can represent different portions of the same gene. Individual libraries were mapped back to the transcriptome using *run_RSEM_align_n_estimate.pl* and then converted to .bam files for sequence analysis. Because of the high number transcripts, most of which had very low expression, we carried out DE analysis on the sum of reads that mapped to each component (gene). This entailed converting the mapped reads to counts and then to FPKM to standardise by gene length and by total number of reads per sample. edgeR (Robinson, McCarthy *et al.* 2010) was utilised in *run_DE_analysis.pl* and then differential expression was determined between the four location/treatments using *analyze_diff_expr.pl*. A 5% FDR cut-off was used for all statistical analysis.

Transcript sequence sorted .bam files from Trinity were parsed through *samtools* (Li, Handsaker *et al.* 2009) to (1) align the reads to the reference, (2) create consensus (with heterozygous bases encoded with IUPAC code) for each locus, and (3) to save sequences in fasta format, using the following settings: minimum base quality 13, minimum mapping quality 2 and minimum read depth 3. Resultant fasta files were imported into Proseq (Filatov 2009) and aligned to the reference.

Two more levels of normalisation were considered for the data set, the first is the base expression levels that each transcript needs to meet to reduce low quality data and sequencing errors, and the second a minimum percentage each isoform has to be of a gene to be named under that gene. A graph was made to see how many transcripts would be lost at numerous levels of normalisation across these two variables, Figure 3.8.

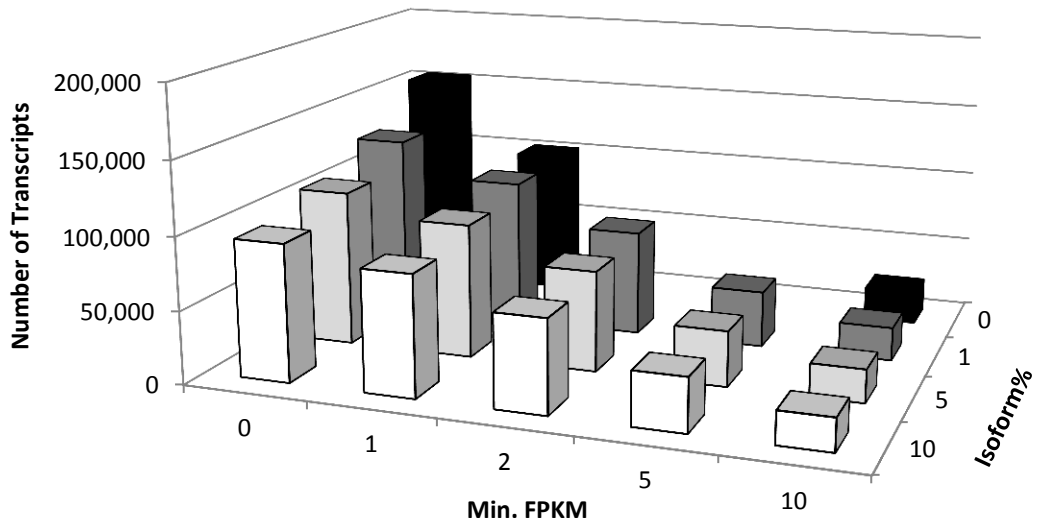



Figure 3.8 – A graph of the number of transcripts that would remain at numerous cut offs for minimum FPKM and minimum isoform coverage.

There is no standard cut off to use for FPKM, and so a range of cut offs have been used in the literature. Cut offs around 0.05 appear common (Vijay, Poelstra *et al.* 2013), with the lowest cut off noted as 0.03 (Rowley, Oler *et al.* 2011). Higher cut offs have also been used, with the highest cut off noted occurring at 10 (Jeukens, Renaut *et al.* 2010, Zemp, Minder *et al.* 2014). Other studies have used other methods, such as using a 95% confidence level and removing the 5% lowest expressed contigs (Zhang, Guo *et al.* 2010). Zhang *et al.* (2010) found this left them with a cut off of >0.78. Cut offs for FPKM are used to try and reduce noise as it is hard to determine real differences from noise in lowly expressed genes. This does also mean that real differences can be missed when removing data so achieving the right balance is essential. When looking at whole data set analysis we decided to use a cut off of 5 as when looking at the entire data set lowly expressed genes can create a lot of noise and misdirect the data set. As we are interested in stomatal patterning genes specifically which are known to be lowly expressed, and having a cut off of only 1 removed over 50,000 transcripts, we decided to use no cut off for FPKM for candidate gene analysis so no important transcripts would be missed, and to handle any lowly expressed genes with more caution when making conclusions.

The cut off for isoforms relates to the percentage coverage an isoform has to have of a transcript to be classed as part of that gene. So to use a cut off of 1% would require the isoform to cover at least 1% of the transcript to be classed as part of that transcript. Again two cut offs were used, for whole data sets a 5% cut off was used, and again when looking for specific genes no cut off was used to minimise removing any important data.

This gave two data sets to use, one with a 5 (min. FPKM) x 5 (min. isoform %) cut off for whole data set analysis, then a data set with no trimming when looking for specific genes.

As components are more comparable to genes than transcripts, it was decided to use components throughout the analysis. The could lead to differences in transcripts being missed, but as we are mostly concerned with differential expression at the gene level, using components will give us more of an insight into the data set.

### 3.2.3 Statistics and BLAST Parameters

A two way ANOVA was carried out on FPKM values for each component to give a significance value for Location (Spring or Control site), Treatment (elevated $CO_2$ or ambient $CO_2$) and an interaction term between the both Location and Treatment. A 5% false discovery rate (FDR) was implemented for all analyses. When sub sets of data were used (e.g. all stomatal patterning genes) an FDR of 5% was used on the P values of that subset. A t-test with a 5% FDR was used to show the number of significantly differentially expressed genes between all treatment (CA vs CE; SA vs SE; CA vs SA; CE vs SE). This analysis was then confirmed using EdgeR (Robinson, McCarthy *et al.* 2010), where the numbers showed the same pattern and there was a 69% match with the components from the t-test.

A BLAST search was carried out on all components to try and match them to an *Arabidopsis thaliana* gene to give functions to the components. The top hit was chose for each component with an E-value cut off of $10^{\wedge -5}$. If multiple components matched to a single AT number, then the total of those components was used for that gene and a new two way ANOVA was carried out for that gene.

### 3.2.4 Gene Expression Analysis Tools

An important aspect of dealing which such large quantities of data is to be able to find differences when there are so many within the data. Numerous tools exist to help visualise data to aid in the identification of genes and general areas of importance within a data set. MapMan (Thimm, Bläsing *et al.* 2004) takes where possible, expression across all known plant pathways to help visualise gene expression on both a single gene level and whole pathway level. MapMan uses the ratio between treatments to show whether a gene is being up regulated or down regulated via colour to give images where you can instantly identify where differences are occurring (Figure 3.9).

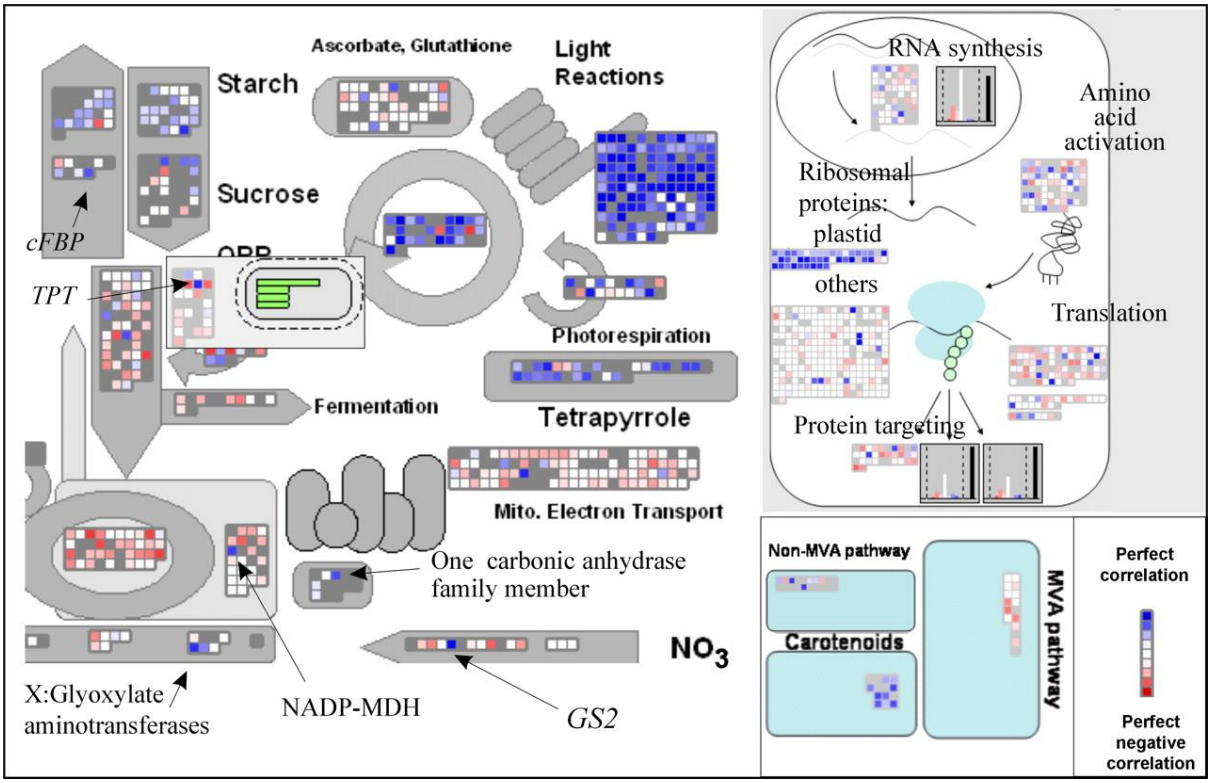



Figure 3.9 -.Example of a pathway analysis from MapMan (Thimm, Bläsing *et al.* 2004). Each box represents a gene, and the colour represents whether there is a negative or positive regulation between treatments, blue is an increase in expression and red is a decrease in expression.

A large number of studies have used MapMan to both identify and to present data as it gives an immediate striking impression of changes between data sets (Urbanczyk-Wochniak, Usadel *et al.* 2006, Sreenivasulu, Usadel *et al.* 2008, Rotter, Camps *et al.* 2009). A limitation of MapMan is each gene needs to be identified via the Arabidopsis database as each gene needs to have an ATG number, so this method can only be used where a match between a sequence of interest and an *Arabidopsis* sequence exists.

Another method is to use Gene Ontology (GO) analysis to give functional annotation and provide enrichment analysis between data sets. GO analysis simply takes all of the significantly different genes and places them into categories so you can visually see which categories the significant genes fall into. The piece of software most commonly used for this analysis in agricultural fields is AgriGO (Du, Zhou *et al.* 2010). The software can produce hierarchical tree graphs of overrepresented GO terms for biological processes, molecular function and cellular components. These elaborate graphical outputs can allow users to explore biological meaning in an intuitive way. The direct acyclic graph or tree structure graph based on the nature of GO can indicate which terms are over/under-represented and the inter-relationships between terms via colour so the differences are easily recognisable (Figure 3.10) .

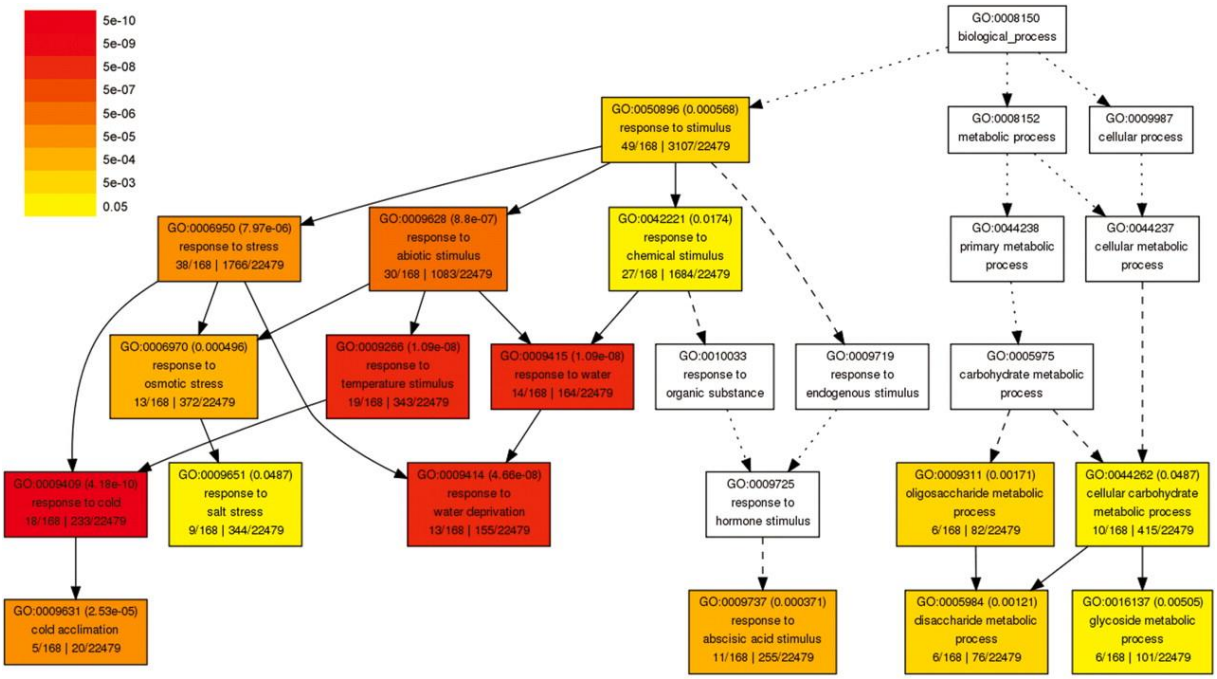



Figure 3.10 -Example of a GO analysis from AgriGO (Du, Zhou *et al.* 2010).

The software uses a Fishers-exact test to give a significance value for how significant the representation of each GO category is compared to the others. The colours are based on these significant values, so highly represented GO categories can be easily recognised. Again using these outputs is common in studies due to the encompassing of a whole data set in one image (Stamm, Ravindran *et al.* 2012, Song, Chen *et al.* 2013).

### 3.2.4 RT-PCR

An aliquot of RNA from the extraction at 58 days after establishment was kept for each sample for RT-qPCR analysis to allow for confirmation of the RNA-Seq results. To carry out RT-qPCR cDNA is required, so reverse transcription of the RNA was performed using the ImProm-II Reverse Transcription kit (Promega, UK). RNA concentration was assed using a spectrophotometer (NanoDrop ND-1000 Spectrophometer) and diluted with RNA-free water to ensure a constant concentration across samples. The volume of RNA and DEPC-treated water was 4µl for all samples. 1µl of oligo dT was then added to each sample. The samples were then incubated at 70˚C for 5 minutes and quickly chilled at 4˚C for 5 minutes. A reverse transcription mix was then added to the RNA/primer mix. The reverse transcription mix contained; 2.6 µl of nuclease free water, 4 µl of ImPromII X5 reaction buffer, 6.4 µl of $MgCl_2$ (25mM stock, 8mM final volume), 1 µl of dNTP (10mM), 1 µl of ImPromII RT. The samples then underwent a series of incubations which consisted of three steps; anneal at 25˚C for 5 minutes, extend for 60 minutes at 42˚C and heat-inactivated at 70˚C for 15 minutes. The PCR product was the cDNA and was stored at -20˚C.

For the RT-PCR reaction the cDNA was diluted 1:5 in DEPC-treated water. The RT-qPCR reaction mix was then made, consisting 5 µl of 2X Precision-SY Master Mix (PrimerDesign Ltd, UK), 5pmol of forward and reverse primers and 25ng diluted cDNA. Plates were run on a Chrome4 Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, USA). The programme used for the RT-qPCR reaction was; incubation at 95˚C for 10 minutes and then 40 cycles of 15 seconds at 95˚C, 1 minute at 60˚C and a plate read. An incubation at 72˚C for 10 minutes then followed. A melting curve was then performed from 60˚C to 95˚C with a read every 0.2˚C and a 1 second hold, in order to check for primer dimers, DNA contamination and secondary products.

Three replicates of each plate were completed to account for errors. Values were exported using the Opticon Monitor 3.1 software (Bio-Rad Laboratories). The data was then extracted from the software to a spread sheet. To do this a threshold of 0.01 was set which was just above the level of background fluorescence to avoid the background fluorescence affecting the values.

An average of the three technical replicates was taken to give one value for each sample. This gave a C(t) (cycle threshold) value. The C(t) is defined as the number of cycles required for the fluorescent signal to cross the threshold. C(t) levels are inversely proportional to the amount of target nucleic acid in the sample (the lower the C(t) level the greater the amount of target nucleic acid in the sample).

The efficiency (E) of each sample was calculated to take into account any differences in the efficiency of the primer amplification that could influence the expression level. To do this the Liu

and Saint method (Liu and Saint 2002) was used which calculates the amplification efficiency for each individual reaction from the kinetic curve, and the initial amount of gene transcript is derived and normalized, Formula 3.2. The two arbitrary thresholds are taken from the exponential phase of the kinetic curve, Figure 3.11.

$$E = (R_{n,A}/R_{n,B})^{[1/C_{T,A}-C_{T,B}]} + 1$$

Formula 3.2 – Amplification efficiency calculation where $R_{n,A}$ and $R_{n,B}$ are $R_n$ at arbitrary thresholds A and B in an individual curve, respectively, and $C_{T,A}$ and $C_{T,B}$ are the threshold cycles at these arbitrary thresholds.

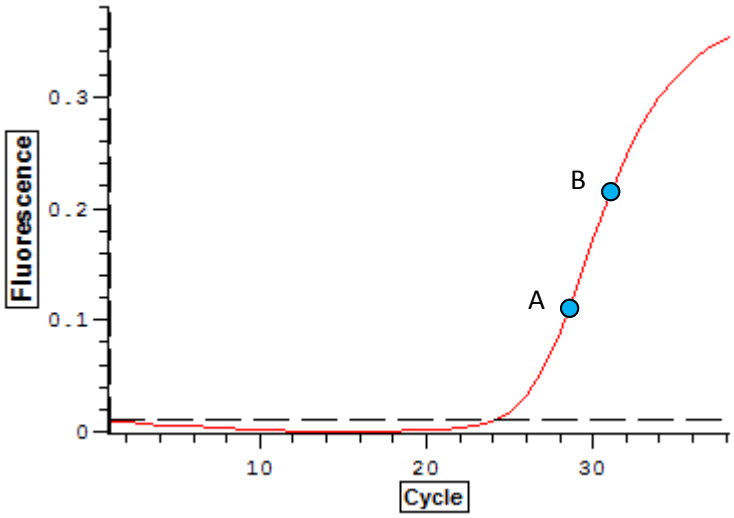



Figure 3.11 – Examples of where the two arbitrary thresholds (A and B) are taken on the kinetic curve to use in formula 3.2.

The results of the equation have been shown to be variable (Regier and Frey 2010), however they can be improved by applying the mean amplification efficiencies for each amplicon to the Liu and Saint method (Regier and Frey 2010). Once the E values were calculated they were placed into formula with the C(t) values for both the gene of interest and the reference gene. As the reporter fluorescence (R) for the gene of interest and reference gene were not equal the formula needed to incorporate two E values. Due to this Formula 3.3 was used to give a value for sample.

$$R_T/R_R = \frac{(1+E_R)^{Ct,R}}{(1+E_T)^{Ct,T}}$$

Formula 3.3 – Calculation used to give a normalised expression value for each sample. $R_T$ is the fluorescence of the target gene and $R_R$ is the fluorescence of the reference gene. $E_R$ is the efficiency of the reference gene and $E_T$ is the efficiency of the target gene. Ct,R is the C(t) value of the reference gene and Ct,T is the C(t) value of the target gene.

### 3.2.4.1 GeNORM

GeNORM is a piece of software which allows for rapid analysis of RT-PCR plates to identify valid housekeeping genes. It uses a calculation to measure the gene stability of a gene which it identifies as *M* (Vandesompele, De Preter *et al.* 2002). This measure is calculated by taking the Log of the ratio between a single potential housekeeping against all of the other potential housekeeping genes for all samples (gives $A_{jk}$ value, Formula 3.4). The standard deviation of the values for all samples for each gene is then taken, and an average of the standard deviations against each gene is then taken which gives the *M* value for that particular gene. This is carried out for each potential housekeeping gene so each gene has an *M* value. These values are then ranked, the lower value the higher the stability of the housekeeping gene. The software then uses a second calculation to determine how many housekeeping genes are required to give a stable expression to normalise to as often one gene alone is not stable enough. It does this by taking the geometric mean of the housekeeping genes with the lowest *M* values, then adding an extra gene into the calculation according to their *M* value rank. The differences in the geometric means are then calculated to give *V* value, and if the *V* value calculated is lower than 0.15 (values below this have been shown to have an non-significant difference) the difference is shown to be non-significant and the addition of further housekeeping genes does not further improve normalisation.

$$A_{jk} = \left\{ \log_2\left(\frac{a_{1j}}{a_{1k}}\right), \log_2\left(\frac{a_{2j}}{a_{2k}}\right), \ldots, \log_2\left(\frac{a_{mj}}{a_{mk}}\right) \right\} = \left\{ \log_2\left(\frac{a_{ij}}{a_{ik}}\right) \right\}_{i=1\rightarrow m}$$

Formula 3.4– The formula used to calculate an $A_{jk}$ value, for every combination of two internal control genes j and k, an array $A_{jk}$ of m elements is calculated which consist of $\log_2$-transformed expression ratios $a_{ij}/a_{ik}$ (Vandesompele, De Preter *et al.* 2002).

### 3.2.5 Chromosome number

Samples were sent to Plant Cytometry services for both genome size estimates and ploidy estimates to check that the plants from the spring and the control site had the same genetic structure, as auto ploidy has been known to occur in species like *P. lanceolata*. Two samples from the spring and two from the control were sent for genome size estimates, then 25 samples were sent for ploidy estimates (the samples sent for genome size also had been sent for ploidy estimates) as the ploidy estimates can indicate the genome size if the ploidy is the same as the samples sent for genome size. A 5 cm sample from a mature young leaf was sent for each sample. The results can be seen in Appendix Table 7.1.

## 3.3 Results

After normalisation a principal component analysis was carried out on the Trinity assembled data (Figure 3.12). First of all this showed that the samples that previously appeared to be outliers now correlated with the patterns of all the other samples, indicated the new assembly had improved the data set. Secondly it showed that a wide range of genes were being differentially expressed in the control plants when exposed to elevated $CO_2$ with rather small changes, in general, in gene expression when Spring plants were exposed to elevated $CO_2$.

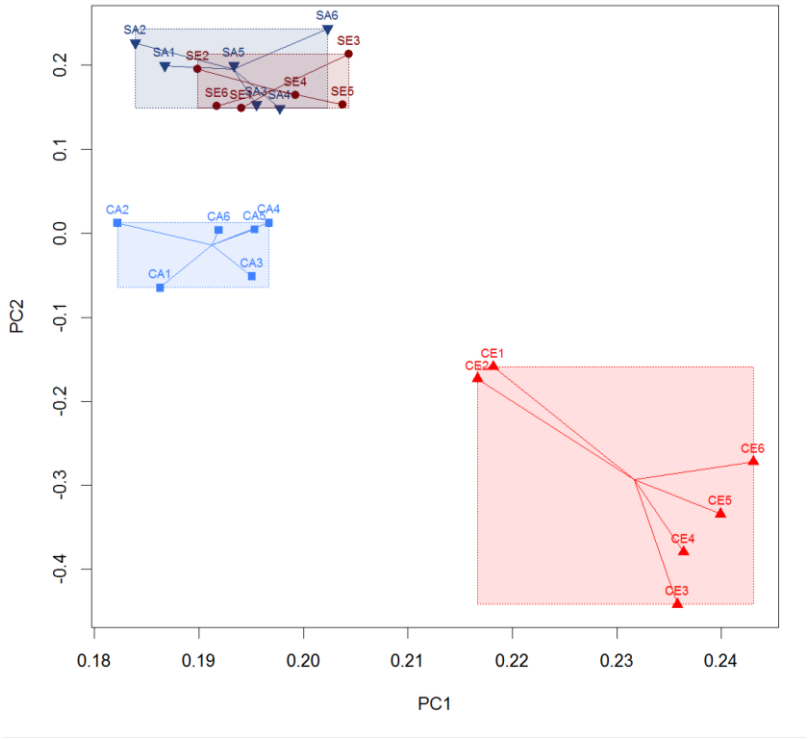



Figure 3.12 - Principal component analysis on contigs FPKM expression of the Trinity assembled data. PC1 separated the ambient and elevated $CO_2$ treatments whereas PC2 separated the control and spring growth sites. CA is control ambient. SA is spring ambient. CE is control elevated. SE is spring elevated.

From the two-way ANOVA and 5% FDR a list of significant differentially expressed genes between all groups was made (Figure 3.13) to give a representation of where the variation is occurring.

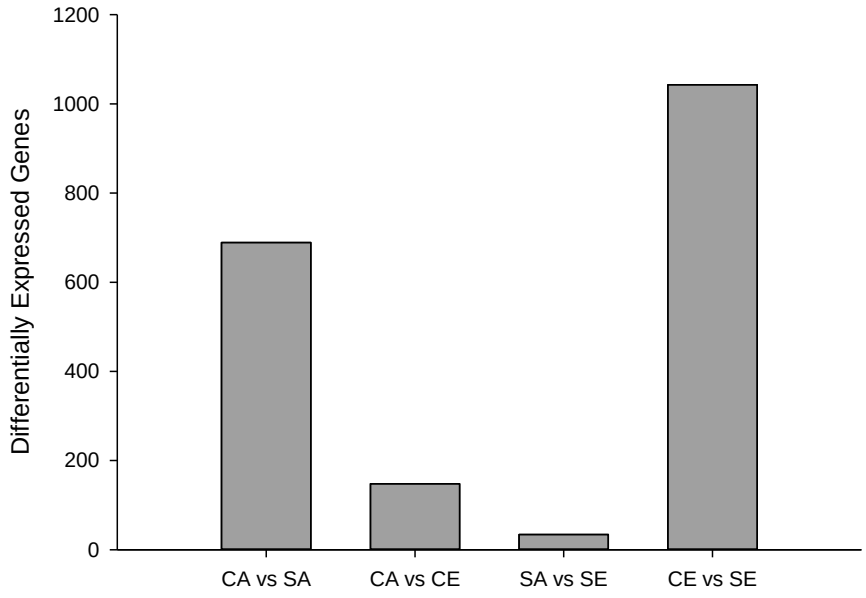


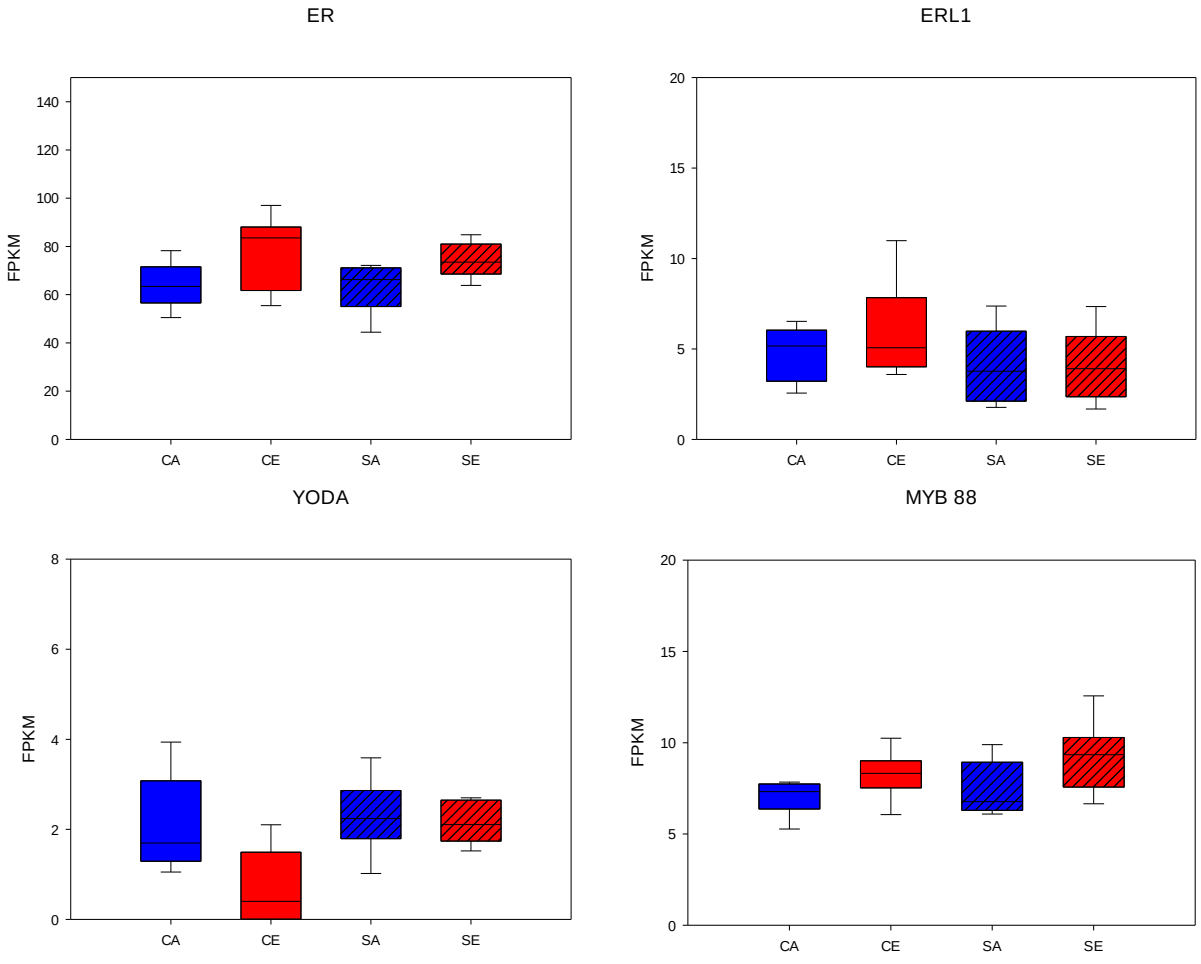

Figure 3.13 – A bar chart showing the number of significantly differentially expressed components there were between each group.

The total number of significantly differentially expressed genes between all Ambient vs. all Elevated was only 117, whereas the total number of differentially expressed genes between all Control vs. all Spring was 1868. To investigate if these significant changes were apparent in the stomatal patterning pathway we extracted any components which matched to a stomatal patterning gene from the BLAST search. The list of stomatal patterning genes can be seen in Table 3.2. The table shows only the stomatal patterning genes which had a match to a component.

| *Gene* | *Location* | *Treatment* | *Location*Treatment* |
|---|---|---|---|
| *ER* | 0.635 | 0.022* | 0.720 |
| *ERL1* | 0.240 | 0.462 | 0.659 |
| *YODA* | 0.033* | 0.039* | 0.462 |
| *MYB 88* | 0.279 | 0.026* | 0.673 |
| *FAMA* | 0.240 | 0.462 | 0.578 |
| *SPCH* | 0.279 | 0.226 | 0.673 |
| *CDKB1;1* | 0.022* | 0.097(10%) | 0.462 |
| *SCRM2* | 0.033* | 0.338 | 0.578 |
| *EPF1* | 0.279 | 0.462 | 0.659 |
| *BCA1* | 0.279 | 0.026* | 0.659 |
| *BCA4* | 0.878 | 0.097(10%) | 0.659 |

Table 3.2 – List of all stomatal patterning genes identified and the significance of the component they matched to. * P≤0.05; ** P≤0.01.

Box plots showing the expression of all of the stomatal patterning genes from Table 3.2 can be seen in Figure 3.14.



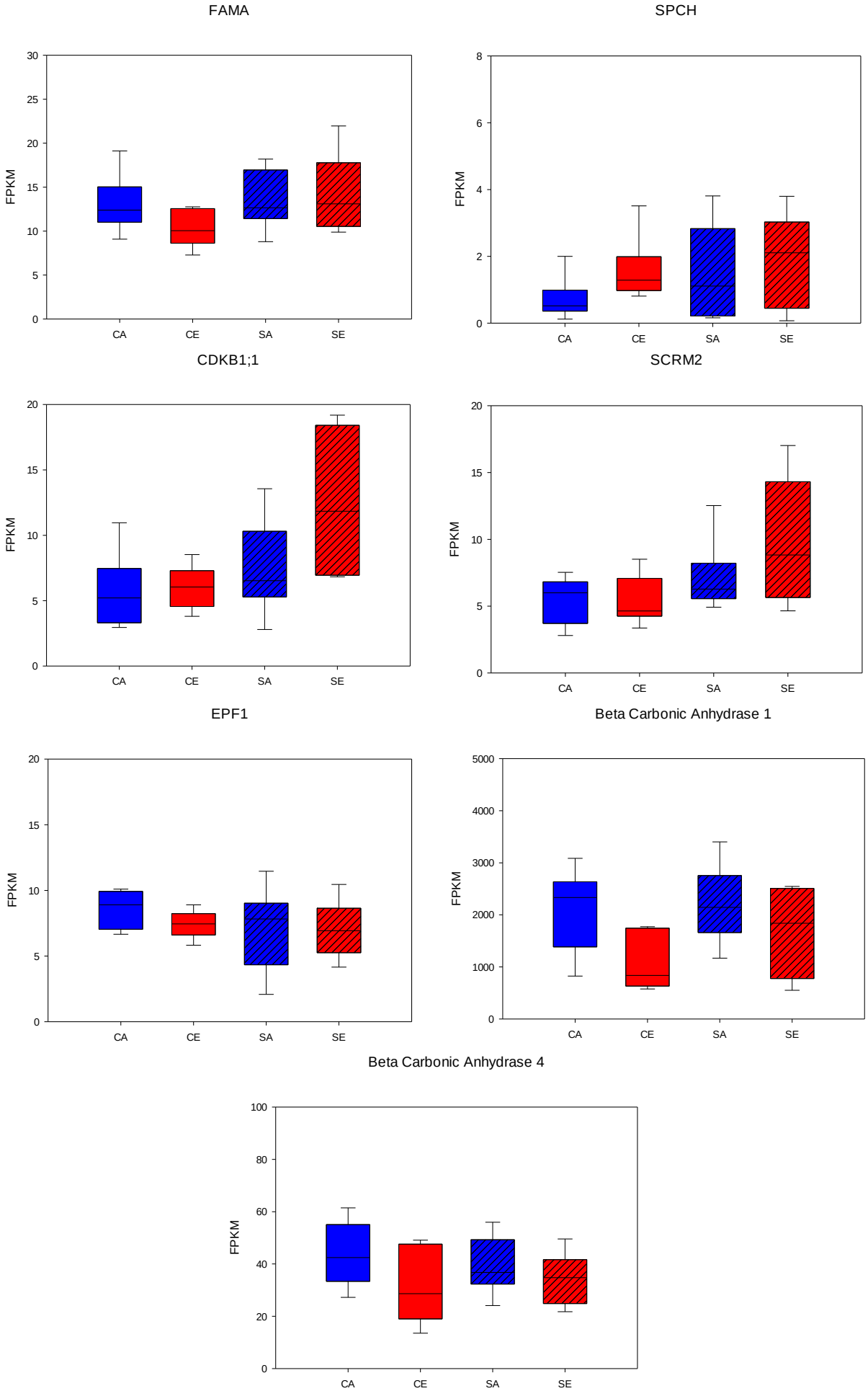



Figure 3.14 – The effect of elevated $CO_2$ on the stomatal patterning genes of *P. lanceolata*. Box plots of all of the stomatal patterning genes which matched to components. CA= Control Ambient, CE = Control Elevated, SA = Spring Ambient, SE = Spring Elevated. The central line in each boxplot shows the median. Whiskers indicated the $5^{th}$/$95^{th}$ percentiles, and boxes indicate the $25^{th}$/$75^{th}$ percentiles.

Several genes showed a significant difference on sub-group analysis of stomatal patterning genes of $P \leq 0.05$ for location or treatment. ER was down regulated in both control and spring plants in response to elevated $CO_2$, with a significant treatment effect ($P = 0.022$). Since this gene family is a negative regulator of stomatal initiation (Masle, Gilmore *et al.* 2005), with mutants having increased numbers of clumped stomata, any down-regulation might suggest this phenotype was prevalent in elevated $CO_2$ which was indeed observed.

βCA 1 and βCA 4 are the most recently identified genes in stomatal development. A double mutant in the βCA 1 and βCA 4 genes is associated with causing impaired $CO_2$-regulation of stomatal movements and increased stomatal density in response to elevated $CO_2$ (Hu, Boisson-Dernier *et al.* 2010). These genes showed the opposite expression pattern for control and spring plants than expected when exposed to elevated $CO_2$, with control plants showing a decrease in expression and spring plants showing an increase.

Two genes which have a large increase in expression in response to elevated $CO_2$ in the spring plants compared to the control plants are SCRM2 and CDKB1;1. Both of these genes are positive regulators (Boudolf, Barrôco *et al*. 2004, Hofmann 2008) and so this increase would indicate an increase in stomata, with both genes exhibiting a significant location effect ($P < 0.05$). Also CDKB1;1 has also been shown to be a positive regulator for smaller epidermal cell size (Boudolf, Barrôco *et al.* 2004) which was also noted in analysis of the chamber experiment.

The stomatal patterning genes were visualised using the stomatal patterning pathway to give a representation of the changes and where in the pathway they were occurring, Figure 3.15. Percentage change was calculated for each significantly differentially expressed stomatal patterning gene to give an overall representation of all the changes happening.

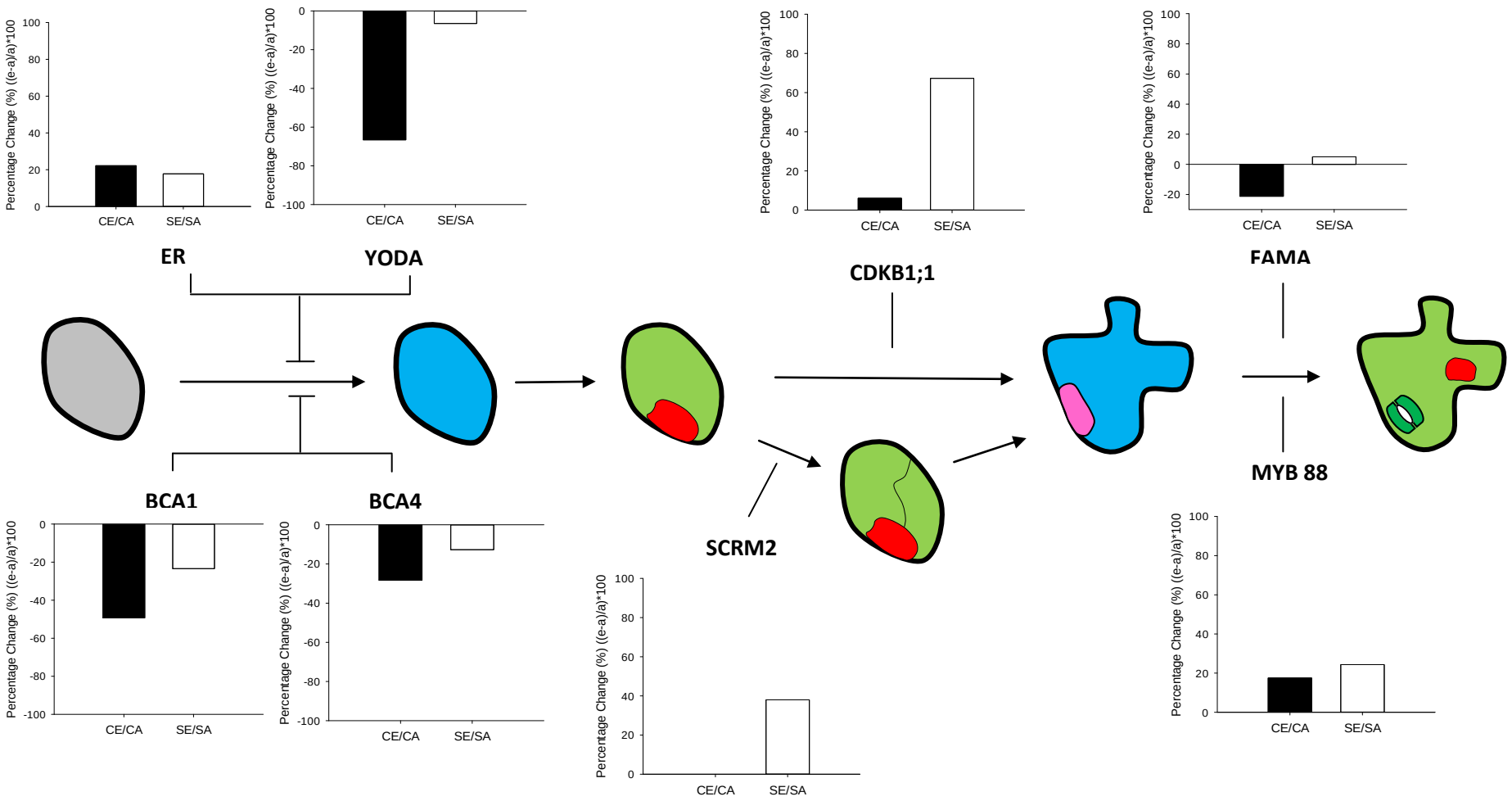



Figure 3.15 – The stomatal patterning pathway with the percentage change of the significantly differentially expressed stomatal patterning genes identified from the RNA-Seq. Negative regulation is indicated by T shaped lines; positive regulation is indicated with a I shaped line.

## 3.3.1 Gene Ontology

### 3.3.1.1 AgriGO

To see where the variation from the PCA was occurring in areas other than stomatal patterning, AgriGO was used to see which gene ontology categories the significant genes fell into. The analysis was first attempted using a 5% FDR significance cut off, but the number of samples that met this cut off and matched an AT number did not meet the minimum number of samples of AgriGO for some comparisons, so a significance cut off with no FDR was used when the minimum number wasn't met. Analysing data with no FDR is acceptable here, as we are only looking for trends in the data, rather than genes that are definitively controlling phenotypes. In AgriGO a Fisher statistical test was used with a significance level of 0.05, with a minimum number of 5 mapping entries to give a significance value to the representation of each GO category. Two lists were made, one of genes significantly different between control elevated and control ambient, and one list of significantly different genes between spring elevated and spring ambient. These lists were put into AgriGO. The cellular component analysis for spring elevated and spring ambient can be seen in Figure 3.16.

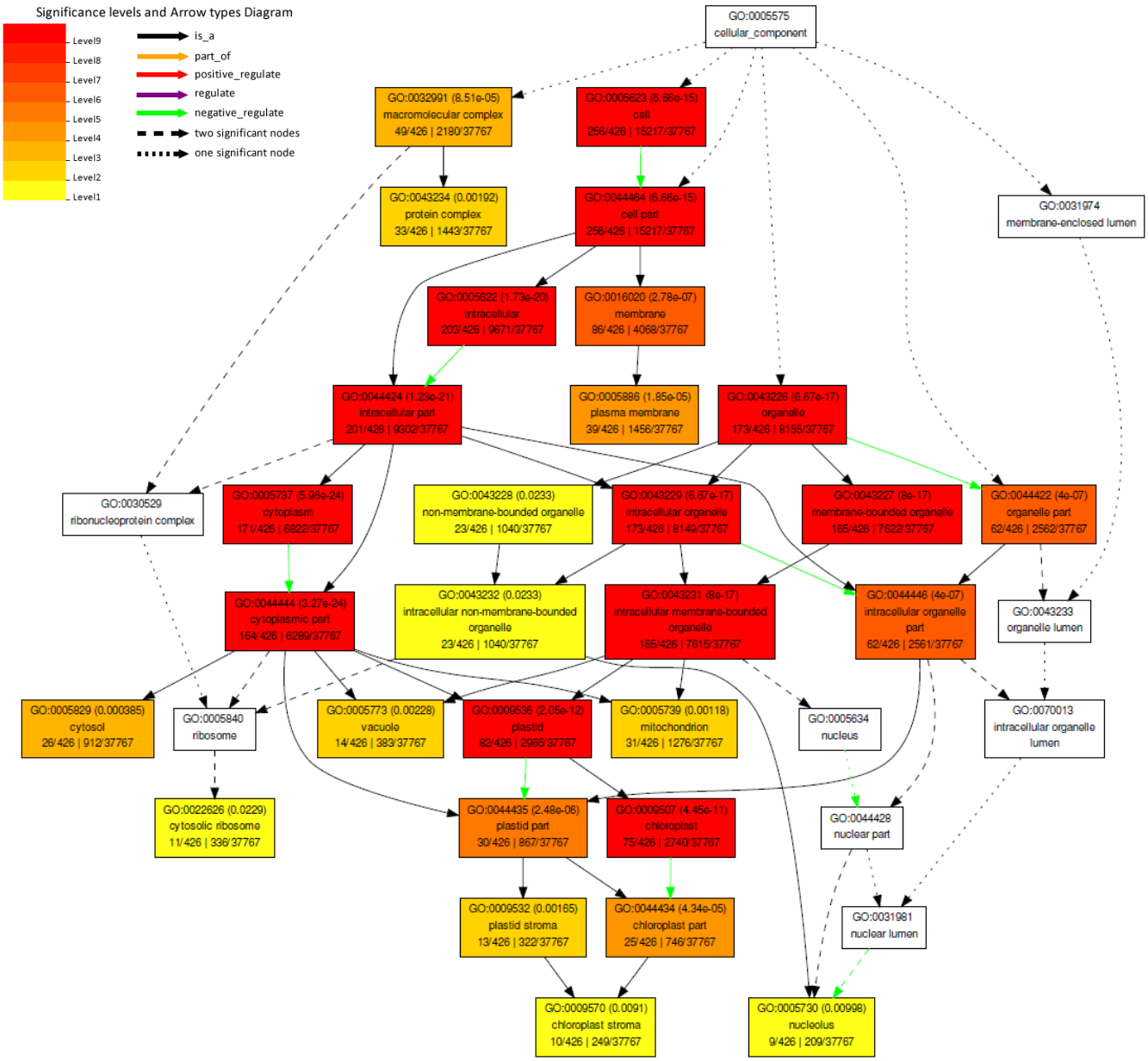



Figure 3.16 – The output of AgriGO for cellular components of the significantly different genes (P ≤0.05, no FDR) between spring elevated and spring ambient. The higher the significance level the more represented a GO category is.

The cellular component analysis for control elevated and control ambient can be seen in Figure 3.17. You can immediately see a lot more activity with almost double the amount of GO categories being differentially expressed. This again highlights the levels of change occurring in the control plants compared to the spring plants.

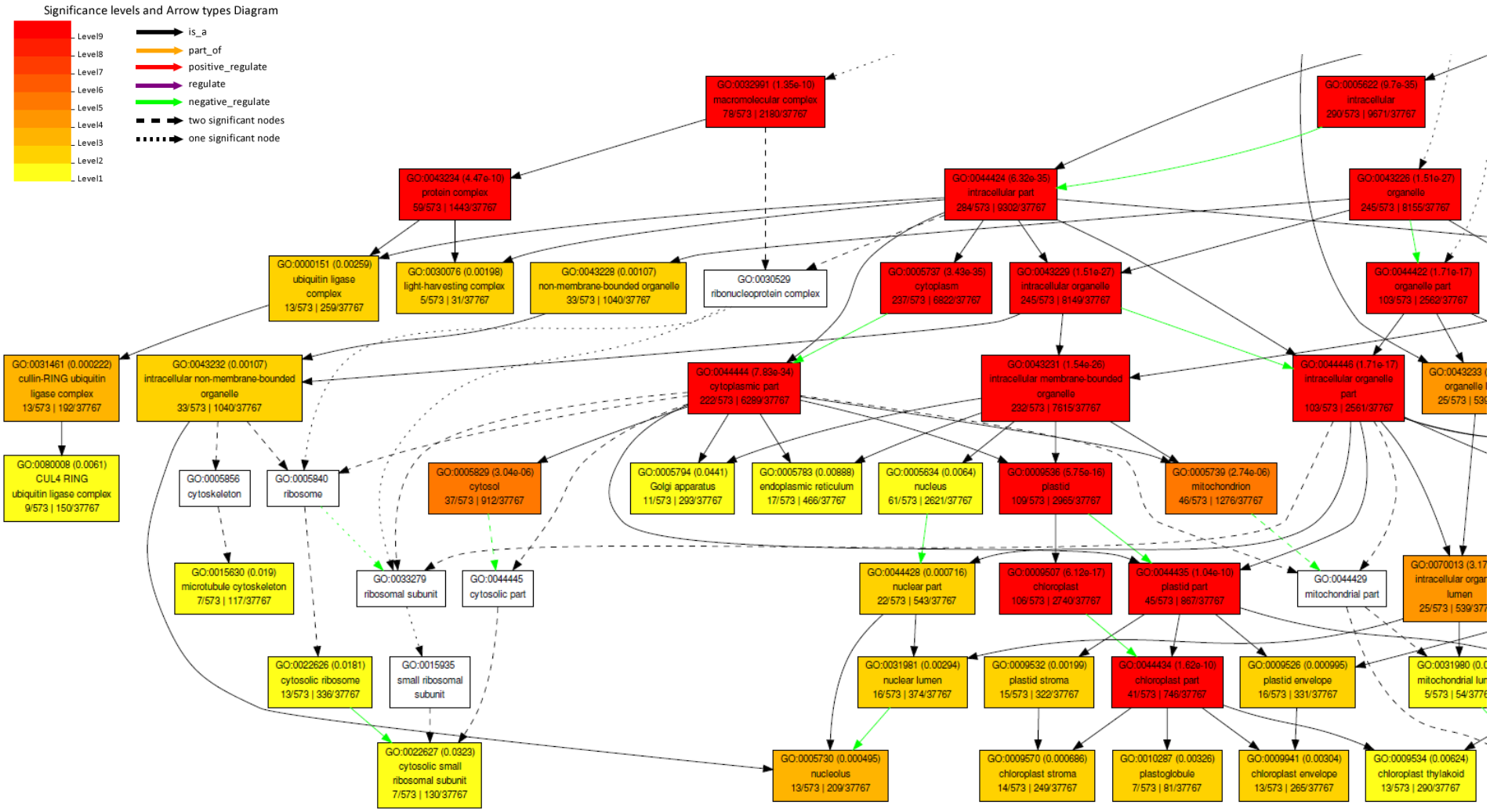



Figure 3.17 – A section of the output of AgriGO for cellular components of the significantly different genes (P ≤0.05, no FDR) between control elevated and control ambient. The higher the significance level the more represented a GO category is.

From analysis of the number of differentially expressed genes, it was clear there were a large number of genes that were constitutive differences between Spring and Control plants, regardless of the $CO_2$ stimuli. These differences are important as these are the adaptive differences that have occurred over generations from long term exposure to elevated $CO_2$. Three analyses were carried out to look at these potential adaptive traits, genes significantly different in spring ambient compared to control ambient (Figure 3.18), genes significantly different in spring elevated compared to control elevated (Figure 3.19), and finally genes significantly different in both of the previous comparisons (Figure 3.20). The cellular components pathway was again analysed. The statistical cut off for these analyses was set at 0.05% FDR as the number of genes differentially expressed was much higher than when looking at differences between $CO_2$ treatments. Again a Fisher statistical test was used with a significance level of 0.05, with a minimum number of 5 mapping entries to give a significance value to the representation of each GO category.

The biological processes was also looked at in the genes that are differentially expressed between both control ambient and spring ambient, and control elevated and spring elevated to look for further genes that could be adaptive to elevated $CO_2$ (Figure 3.21).

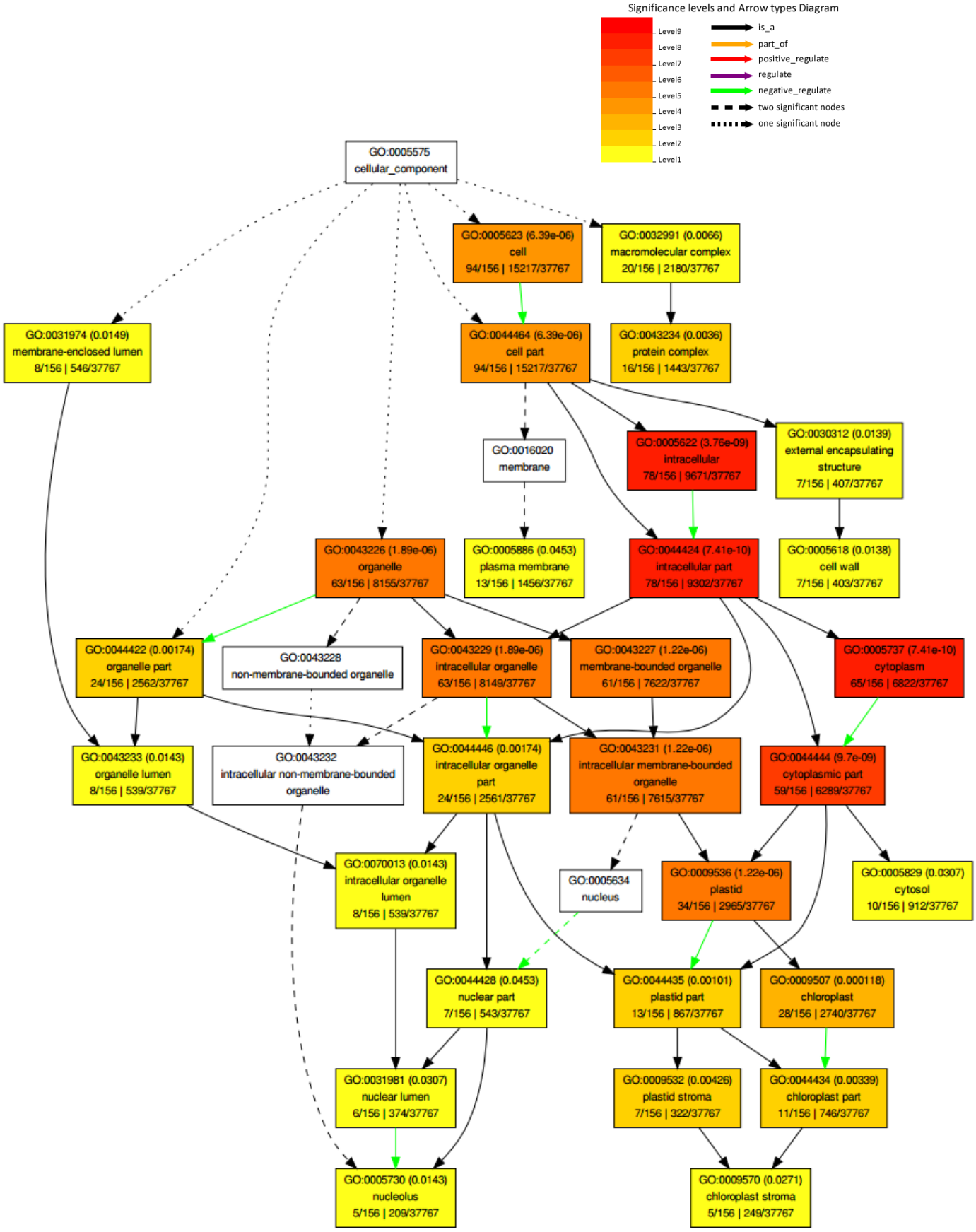



Figure 3.18 – The output of AgriGO for cellular components of the significantly different genes (FDR ≤ 0.05) between control ambient and spring ambient. The higher the significance level the more represented a GO category is.

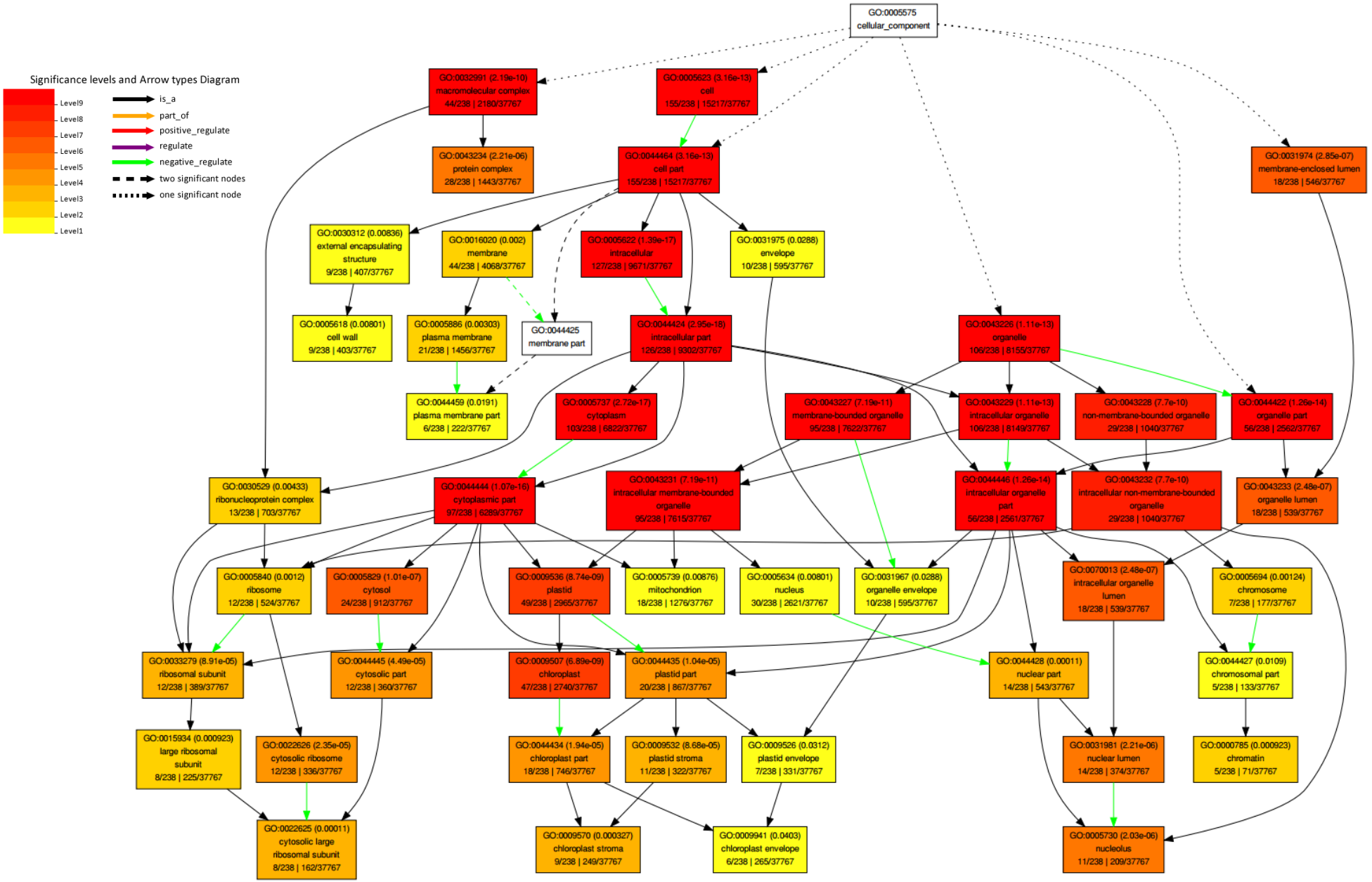



Figure 3.19 – The output of AgriGO for cellular components of the significantly different genes (FDR ≤ 0.05) between control elevated and spring elevated. The higher the significance level the more represented a GO category is.

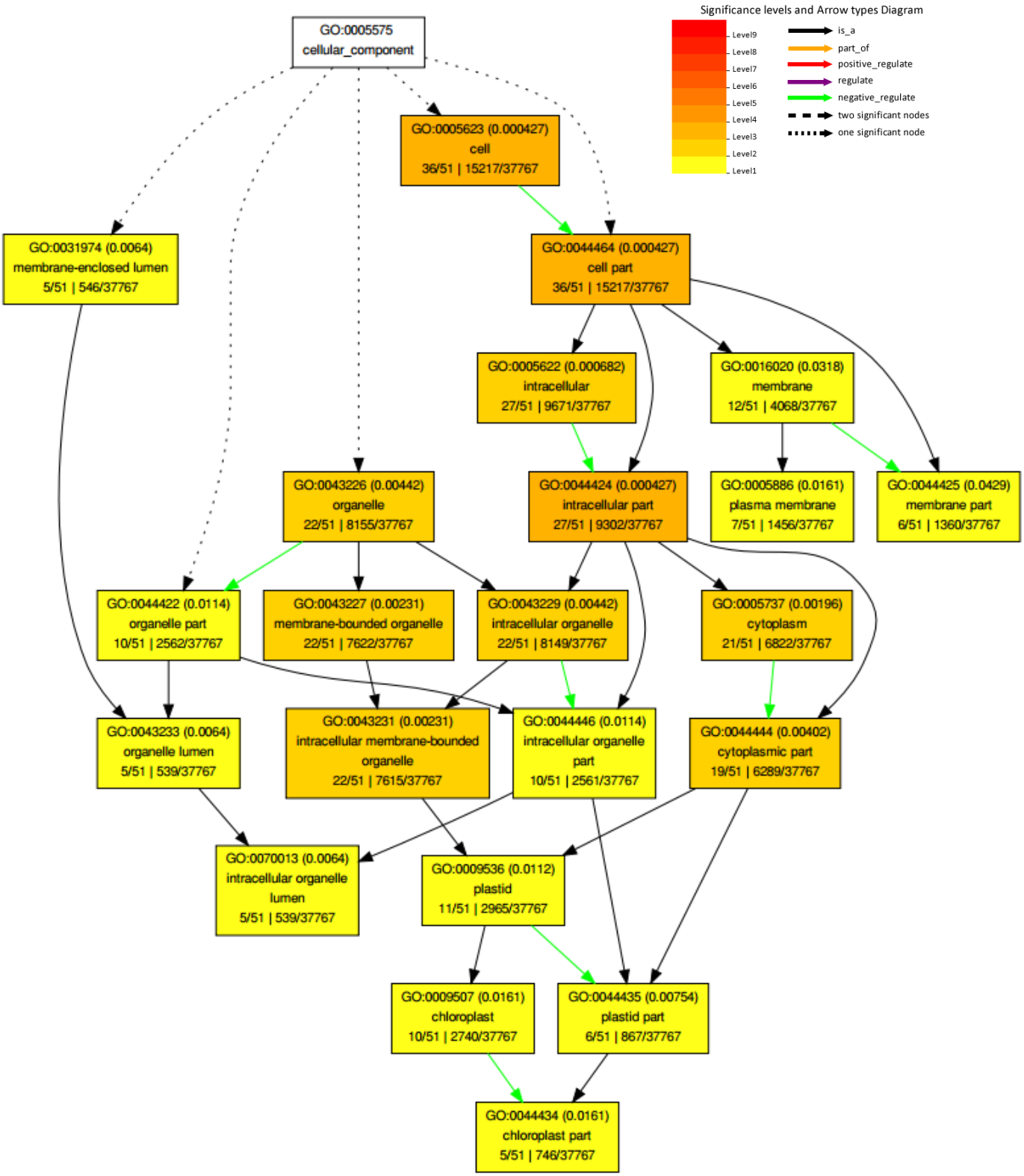



Figure 3.20 – The output of AgriGO for cellular components of the significantly different genes (FDR ≤ 0.05) in both control elevated and spring elevated, and control ambient and spring ambient. The higher the significance level the more represented a GO category is.

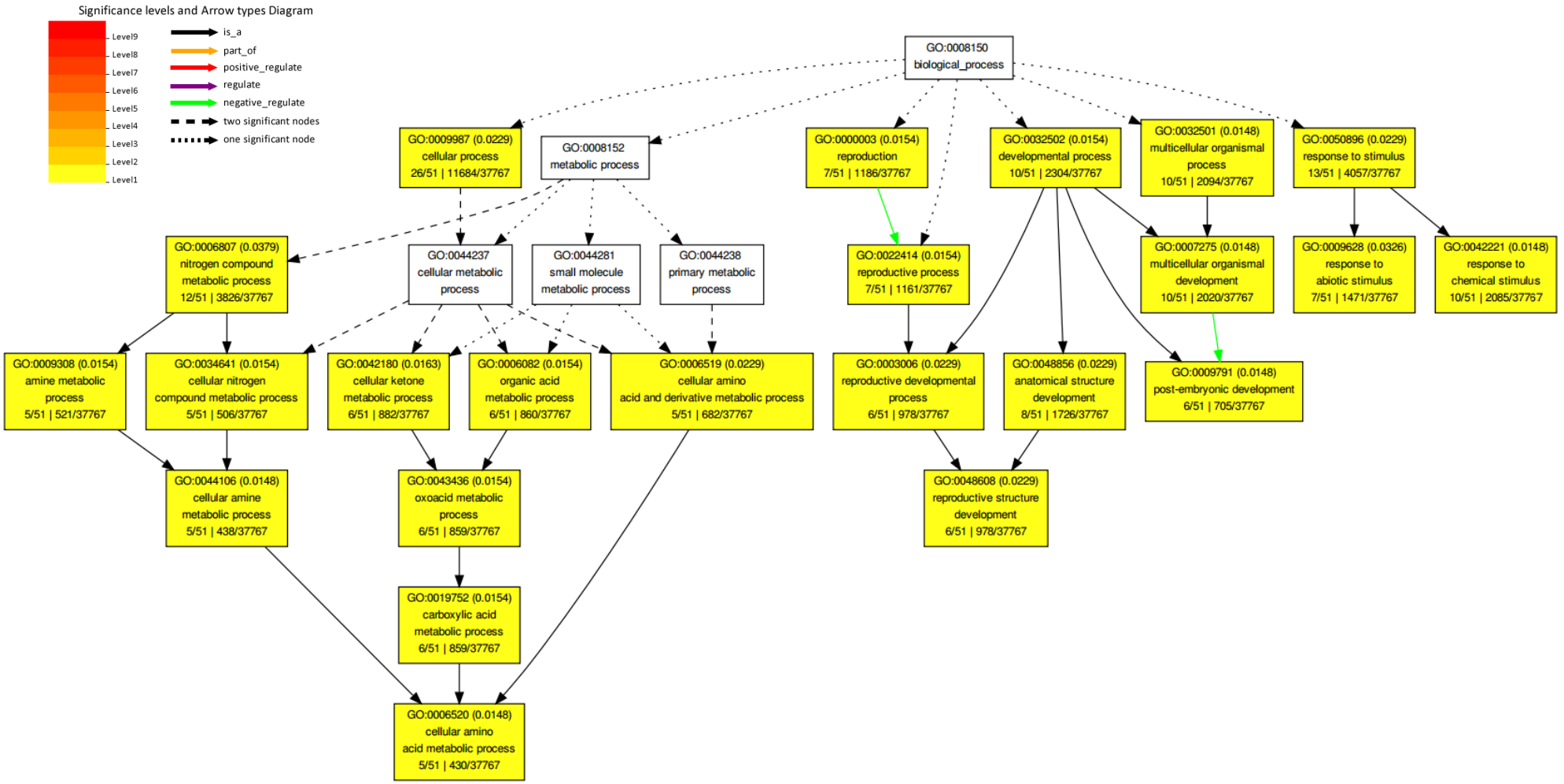



Figure 3.21 – The output of AgriGO for biological processes of the significantly different genes (FDR ≤ 0.05) in both control elevated and spring elevated, and control ambient and spring ambient. The higher the significance level the more represented a GO category is.

#### 3.3.1.2 MapMan

The AgriGO analysis highlighted the chloroplast as containing numerous differentially expressed genes in both control and spring plants in response to elevated $CO_2$. The chloroplast is where photosynthesis occurs and so MapMan was used to have a more in depth look at the photosynthesis pathway to see how the pathway was affected. The respiration pathway was also looked at as the mitochondria appeared to also be over represented, and due to the recent developments in the respiration pathways response to elevated $CO_2$. All genes that matched to an AT number were used in the MapMan analysis. Figure 3.22 shows the output of the pathway analysis of the significant genes, of sub-group analysis of photosynthesis related genes. The list of the genes used and their significance can be seen in Appendix Table 7.2.

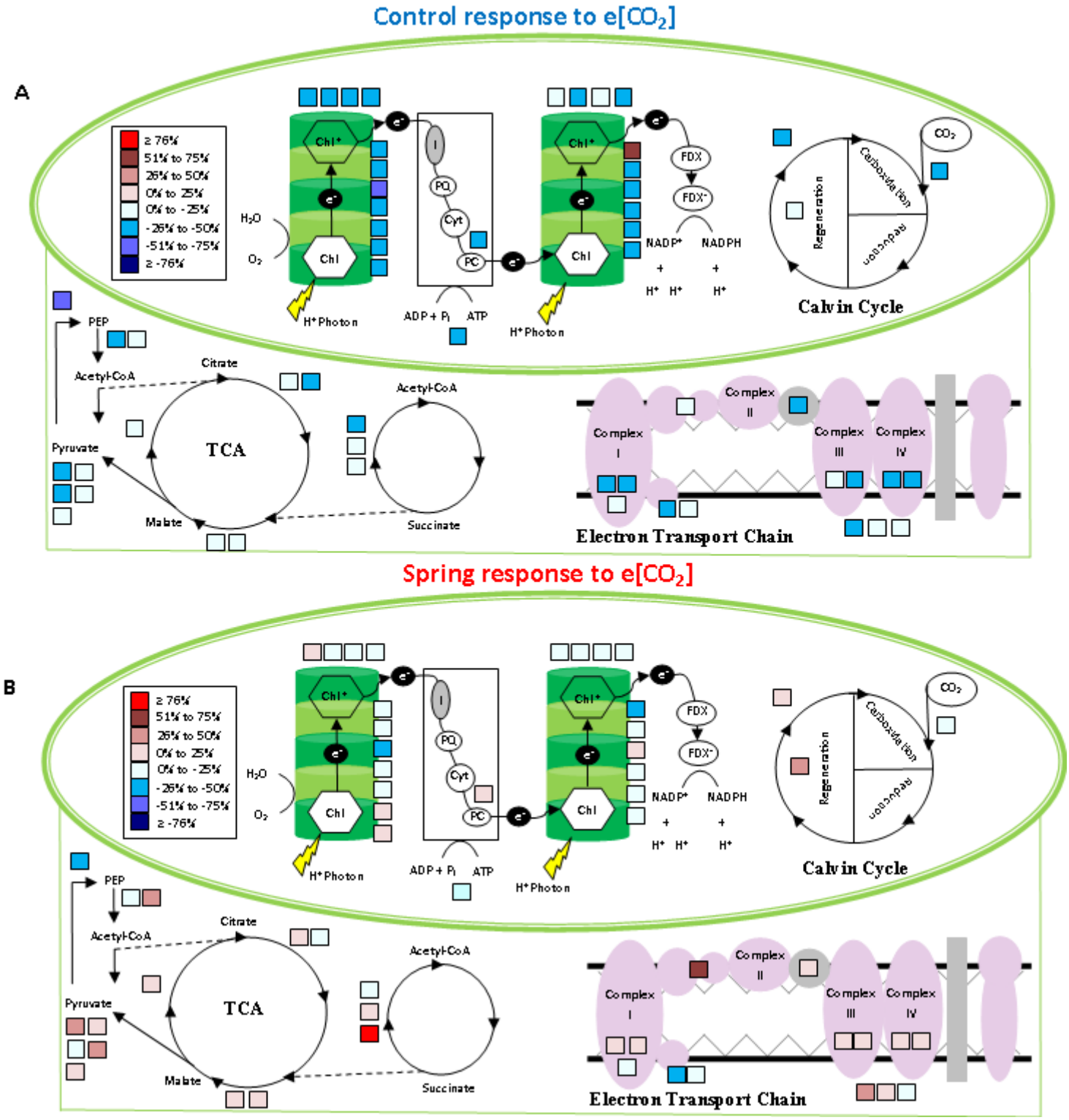



Figure 3.22 - The differentially expressed genes from the output of the photosynthesis pathway from MapMan. (**A**) Control response to e[$CO_2$] refers to control elevated/control ambient, and (**B**) Spring response to e[$CO_2$] refers to the spring elevated/spring ambient.

Using percentage change is a good way to give an overall look at a pathway, but genes with large changes need to be looked at in more detail to see both overall expression and changes between spring and control plants. A selection of key genes from the photosynthesis pathway can be seen in Figure 3.23.

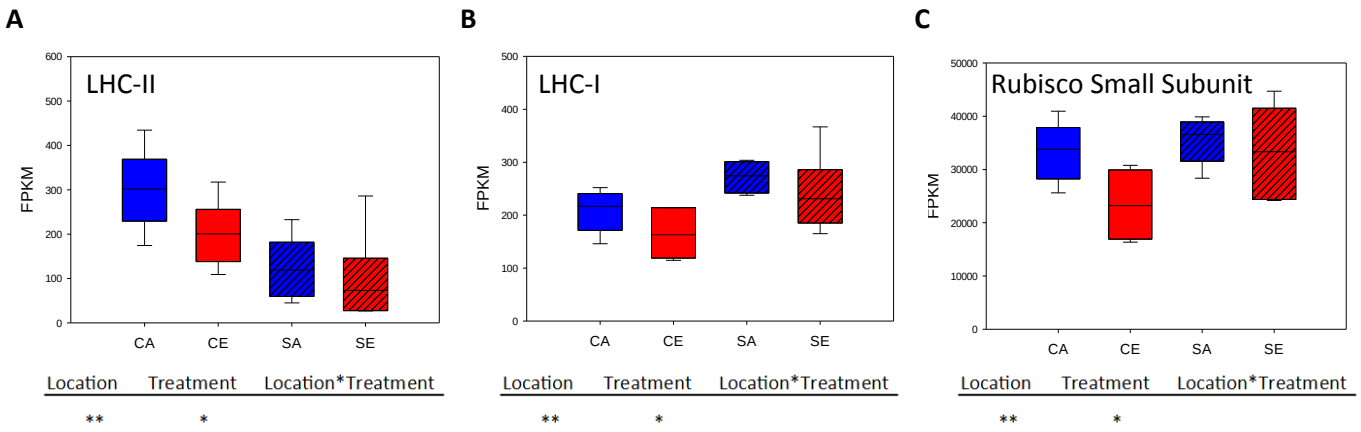



Figure 3.23 –FPKM for three genes involved in the photosynthesis pathway, (**A**) LHC-11, (**B**) LHC-1 and (**C**) Rubisco small subunit. The statistics are from a two-way ANOVA (see table 2.3), * P≤0.1; ** P≤0.05.

### 3.3.3 Real-Time PCR

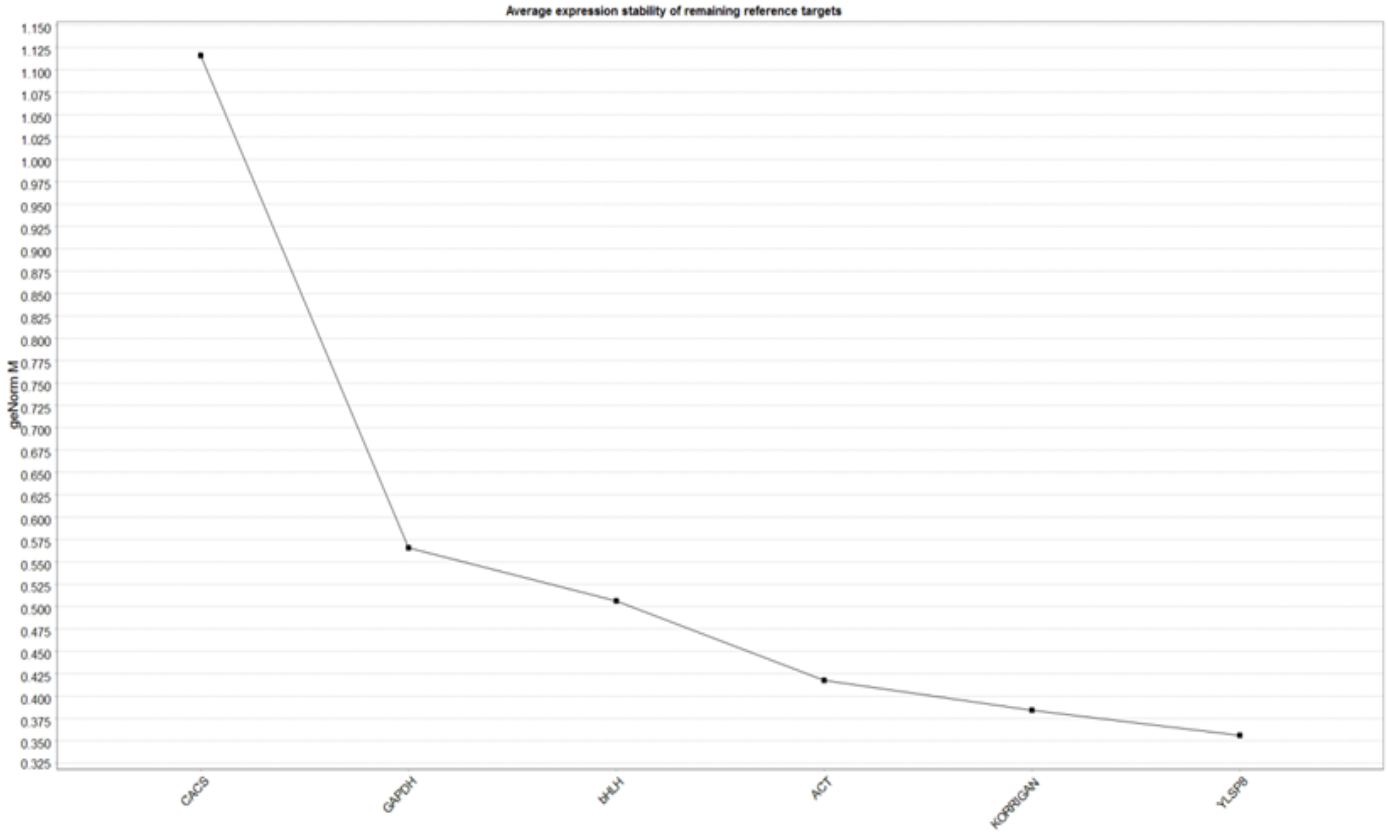
A selection of known conserved housekeeping genes were selected from the literature to try and utilise the GeNORM software to find a set of housekeeping genes. Any genes which could be identified from the BLAST search in our data set were selected. The list of housekeeping genes selected can be seen in Table 3.3.

| *Gene* | *AT Number* | *Study Where Gene Was Used* |
|---|---|---|
| CACS (Clathrin Adaptor Complex Subunit) | At5g46630 | (Demidenko, Logacheva *et al.* 2011) |
| GAPDH (glyceraldehyde-3-phosphate-dehydrogenase) | At1g16300 | (Barber, Harmer *et al.* 2005) |
| ACT (β-actin) | At3g46520 | (Du, Zhang *et al.* 2013) |
| bHLH (Basic helix-loop-helix family protein) | At4g38070 | (Pohjanvirta, Niittynen *et al.* 2006) |
| KORRIGAN | At5g49720 | (Czechowski, Stitt *et al.* 2005) |
| YLSP8 (Yellow-leaf-specific protein 8) | At5g08290 | (Štajner, Cregeen *et al.* 2013) |

Table 3.3 –List of the housekeeping genes used attempt to find a housekeeping gene that is conserved in *P. lanceolata*.

Primers were designed for all housekeeping genes from the components which matched to the housekeeping genes using Primer3 (Untergasser, Cutcutache *et al.* 2012). All of these genes were run with all 24 samples that were sequenced. The C(t) values were then imported into GeNORM to see which genes were the most conserved. The results can be seen in Figure 3.24.

**A**




**B**

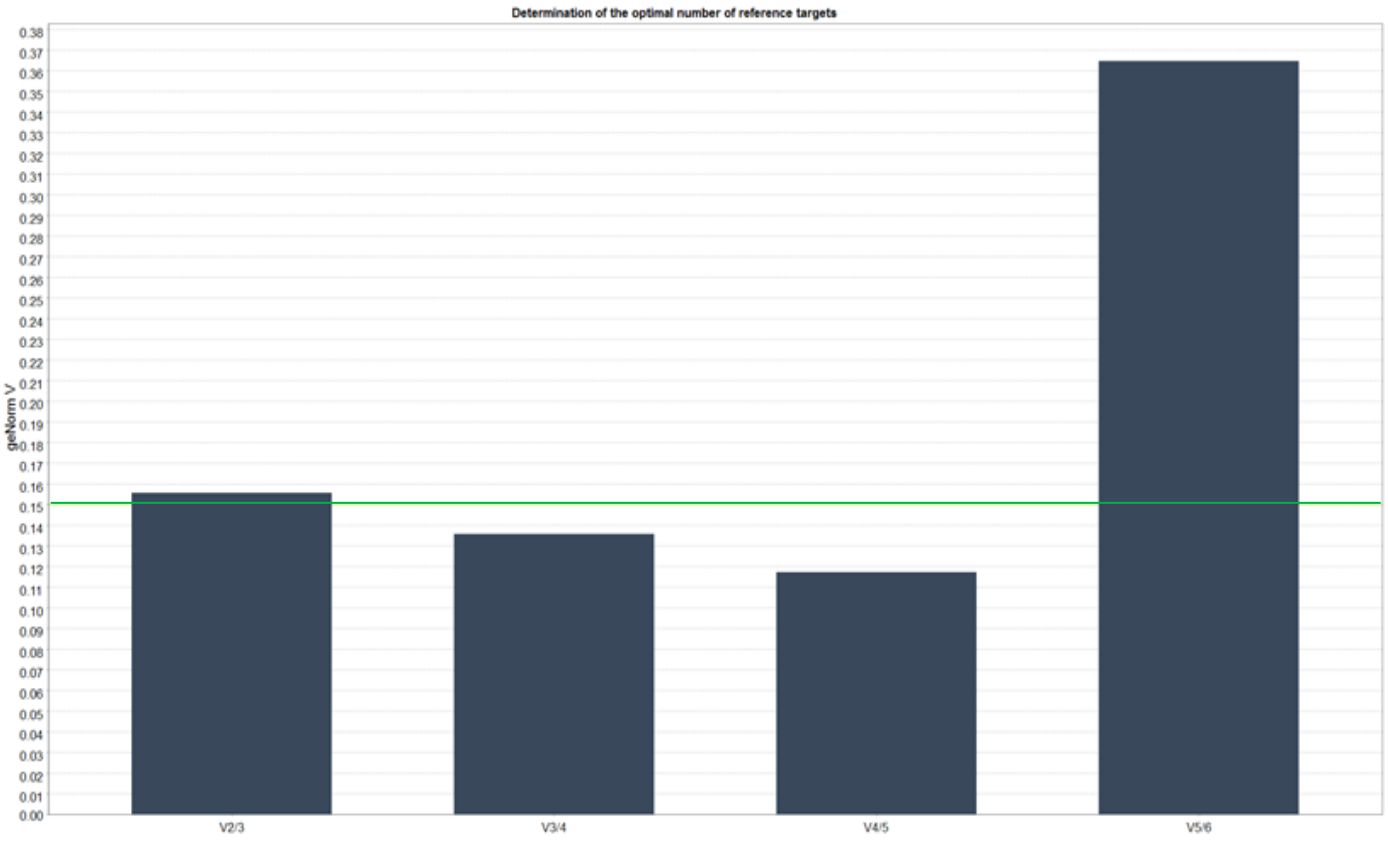



Figure 3.24 – (**A**) A graph of the M values for each housekeeping gene. The lower the value the more conserved the gene. (**B**) A bar chart of the V values using multiple genes in order of their M value (e.g. 2/3 would mean the first 2 most conserved genes used).The bar has to be lower than the green line for that number of genes to be classed as conserved.

The results showed that the geometric mean of 3 housekeeping genes would be needed, YLSP9, KORRIGAN and ACT for correct normalisation. This caused problems as the RT-PCR has to be run on 96 well plates, which would mean if three housekeeping genes were used there would be no room for a gene of interest on the plates due to the number of samples we have. This would then induce another bias as housekeeping genes would have to be on split on different plates. A method which has also been adopted is to find the most conserved component in the data set and use that as the housekeeping gene (Zemp, Minder *et al.* 2014), regardless of function. This method was used so we could use a single housekeeping gene for the RT-PCR. Only components with a total expression >100 were considered as small changes in lowly expressed components could create issues. The variance was then calculated for every component that met the criteria using Formula 3.4. The component with the lowest variance was selected to be the housekeeping gene. This component was contig36689.

$$\sigma = \sqrt{\sum \frac{(x - x)^2}{N - 1}}$$

$\sigma$= Measure of Variance

$x$= Value of samples

$N$= Number of samples

$x$= Mean of sample values

Formula 3.4 – The formula used to calculate variance.

Three genes of interest were selected for RT-PCR to confirm the RNA-Seq results, BCA1, ER and RuBisCO Subunit. The RNA-Seq expression pattern and the RT-PCR expression pattern of each gene can be seen in Figure 3.25.

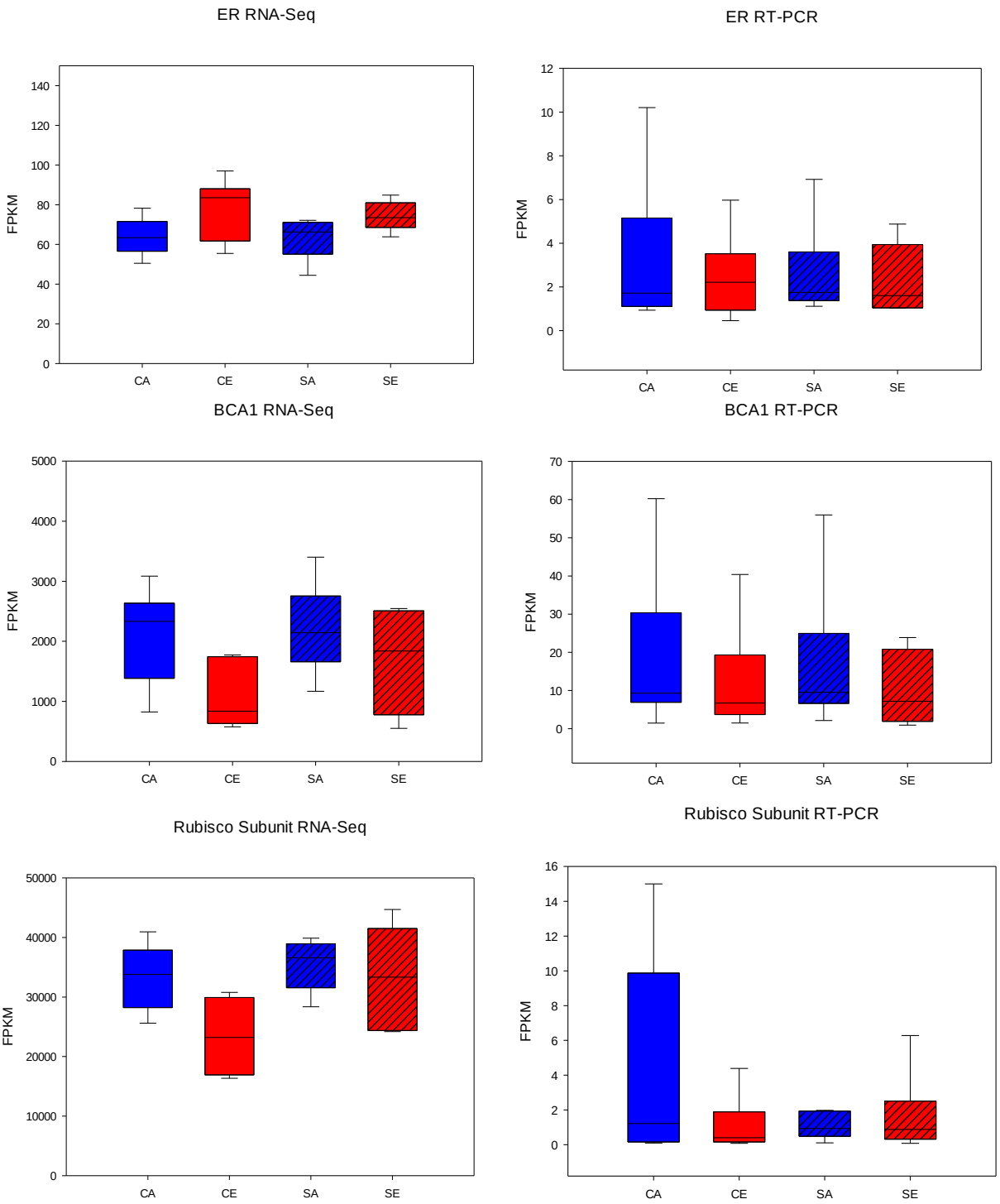



Figure 3.25 - Box plots of three genes chosen for RT-PCR analysis. The boxes on the left are the RNA-Seq expression patterns and the boxes on the right are the RT-PCR expression patterns CA= Control Ambient, CE = Control Elevated, SA = Spring Ambient, SE = Spring Elevated. The central line in each boxplot shows the median. Whiskers indicated the $5^{th}$/$95^{th}$ percentiles.

There did seem to be large variance in the RT-PCR for the RuBisCO subunit, but on analysis of the individual sequences for this gene there seemed to be highly heterozygous (Figure 3.26) meaning there could be potential problems with amplification with the RT-PCR primers in some samples.

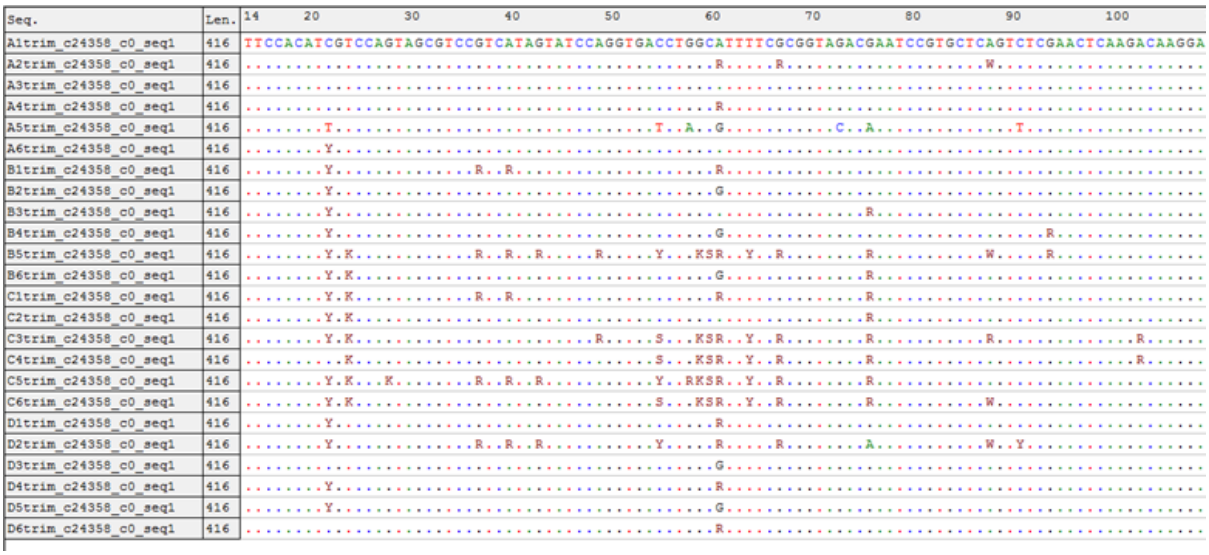



Figure 3.26 – The sequences of all of the individual samples (A1-D6) for the component which matches RuBisCO small subunit. Where there is a dot there is no variation at that base between the top sequences and the rest, where there is a letter there is variation in that base.

Multiple RuBisCO small subunits are encoded in the *Arabidopsis* genome and are highly conserved; this result in trinity not being able to differentiate between different copies of the gene and designing primers from this sequence coupled with the variation between sequences could mean multiple or different RuBisCO subunits are being expressed in different samples leading to the variation observed.

## 3.4. Discussion

Using RNA-Seq, the entire gene expression profile of *P. lanceolata*, originating from both a natural spring and control site has been obtained. Due to this, for the first time, the identification of genes potentially involved in both adaptation and acclimation to elevated $CO_2$ can be analysed side by side, on a whole genome scale. Morphological changes in both spring and control plants in response to elevated $CO_2$ were observed in Chapter 2, and via linking these phenotypes with the RNA-Seq data, more comprehensive conclusions can be made.

The PCA plot revealed clear gene expression differences in the control plants when exposed to elevated $CO_2$, whereas the spring plants appeared relatively unaffected by exposure to elevated $CO_2$. This suggests that the spring plants may be pre-adapted to elevated $CO_2$, and so do not respond to the environmental stimulus to the same degree as the control plants. The diverged expression patterns noted in the control plants is expected, as gene expression across several key pathways has been shown to differ in plants exposed to elevated $CO_2$, compared to ambient $CO_2$ (Ainsworth, Rogers *et al.* 2006). The response of plants to a decreasing $CO_2$ concentration has been shown to be limited in the literature, with experiments using sub-ambient $CO_2$ finding limited changes in the same pathways in response to elevated $CO_2$ (Johnson, Polley *et al.* 1993, Malone, Mayeux *et al.* 1993).This could explain why the spring plants are having such a limited response to ambient $CO_2$ compared to the control plants, as the stress treatment for the spring plants is the sub-ambient $CO_2$.

Differentially expressed genes between locations (spring and control) under the same $CO_2$ condition could to be due to long term adaptations (i.e. fixed divergence in gene expression) which have evolved over generations. The number of significantly differentially expressed genes between experimental groups showed there were over 10 times more significantly differentially expressed genes between locations (1,739 at FDR ≤0.05), than between treatments (117 at FDR ≤0.05). This suggests that the major genetic differences are between the plants location of origin. A list was made of genes that were significantly differentially expressed between spring plants grown in elevated $CO_2$, and control plants grown in ambient $CO_2$, as these genes could be the constitutional adaptations in the spring plants to elevated $CO_2$. To analyse these genes, the list was inputted into AgriGO (Du, Zhou *et al.* 2010) to see which GO categories were being significantly over represented using the Fishers-exact test (Figure 3.16 - 3.21). Within biological processes, several of the significantly over represented categories stood out, of which one comprised of GO categories related to nitrogen metabolic processes (P = 0.0154). Nitrogen is the second highest compound of the gas that is expelled from the spring (Bettarini, Grifoni *et al.* 1999) and so it is expected that the plants inside the spring would adapt to take advantage of these

conditions. The others were response to abiotic stimuli (P = 0.0326) and response to chemical stimuli (P = 0.0148). Again several chemical compounds have been found to present in the spring, including hydrogen sulphide ($H_2S$) (Bettarini, Grifoni *et al.* 1999), emitted along with $CO_2$ which could be driving these changes in gene expression. These changes confer with responses you would expect to see, considering the environment of the natural spring. The cellular component analysis showed the chloroplast was significantly over represented (P = 0.0161). The chloroplasts main role is to conduct photosynthesis via the photosynthetic pigment chlorophyll, which captures the energy from sunlight, and stores it in energy storage molecules (ATP and NADPH), whilst freeing oxygen from water (Klein and Mullet 1986). The ATP and NADPH are then used to make organic molecules from $CO_2$ via the Calvin cycle (Raines 2003). Elevated $CO_2$ is known to have a large effect on photosynthesis (Long, Ainsworth *et al.* 2004), and so it is reassuring that this category is over represented.

The genes which have been identified as differentially expressed between treatments, are genes where acclimation could be occurring (i.e. expression changes depending on the $CO_2$ treatment), and represent the plastic part of the transcriptome. The results of the chamber experiment (Chapter 2), showed some similar phenotypic differences in response to elevated $CO_2$ from both populations, of which some conserved acclamatory genes could be responsible. Only two genes showed a significant difference in plants from both locations in response to elevated $CO_2$, and only one matched to an *A. thaliana* sequence. This gene was an early auxin induced protein (AT3G04730), and auxin is known to antagonise effects of ABA on stomatal opening and closure (Tanaka, Sano et al. 2006). The expression pattern of this gene did not indicate a conserved acclamation, and in fact showed the opposite expression pattern in both populations (in control plants the gene was expressed in ambient $CO_2$, but showed no evidence of expression in elevated $CO_2$; in spring plants the gene showed no evidence of expression in ambient conditions, and expressed in elevated $CO_2$). The stomatal patterning pathway is a phenotype of which the control and spring plants responded differently (see Chapter 2). It could therefore be that plants acclimating to a change in $CO_2$ concentration by a change in stomatal aperture to control their $CO_2$ intake. This gene, along with such a low number of genes significantly differentially expressed in both spring and control in response to elevated $CO_2$, suggests the plants are responding in opposite ways. This is not unexpected, as control and spring plants are experiencing the opposite $CO_2$ stress, so you would expect them to acclimate with different expression profiles.

As the number of genes with significantly different expression in response to elevated $CO_2$ was so limited using the stringency of 5% FDR (117), and due to the requirement of a match to a AT

number, we did not meet the minimum number of genes required for AgriGO analysis. The stringency of the gene list to be entered into AgriGO was lowered to 0.05% significance with no FDR, so we could still use AgriGO to give a general idea of where any changes were occurring in response to elevated $CO_2$. Genes significantly different in control plants to elevated $CO_2$, and those significantly different in spring plants, were analysed separately. It was immediately noticeable that the control plants had a lot more significantly differentially expressed genes than the spring plants, with around double the number of GO categories being significantly over represented in the control plants, compared the spring plants. These results tie in with the PCA plot, indicating there is a lot more change in the control plants than the spring plants in response to elevated $CO_2$. It was again notable in both spring and control plants that GO categories related to photosynthesis were significantly over represented. The chloroplast was highly significant in both control and spring ($P = 6.12e^{-17}$ and ($P = 4.45e^{-11}$, respectively), and the category 'light harvesting complex' was significantly over represented in the control plants ($P = 0.00198$). Using MapMan we were able to look at the photosynthesis pathway in more detail to see any differences. Again it was immediately notable that there is change across the whole pathway in the control plants in response to elevated $CO_2$ (35.14% decrease), whereas the spring plants only showed minor differences in parts of the pathway in response to elevated $CO_2$ (7.23% decrease). All genes used in the photosynthetic pathway showed a significant treatment effect (FDR ≤0.05), further indicating elevated $CO_2$ is a major driver in these gene expression differences. Several of these genes also showed a significant location and interaction effect (FDR ≤0.05), suggesting the response to elevated $CO_2$ the two populations exhibit differs, which is what was observed phenotypically (Chapter 2). The most common previously documented response to elevated $CO_2$ is a down regulation in the photosynthesis pathway (Faria, Wilkins *et al.* 1996), particularly in ribulose-1,5-*bis*phosphate carboxylase (RuBisCO) (Vu, Allen *et al.* 1997). This is the response we see in the control plants, with a significant down regulation across all photosynthetic machinery and RuBisCO. The limited change in gene expression of the spring plants is unusual, but has been previously identified in specific conditions. In species under drought, the down regulation of photosynthesis is diminished (Huxman, Hamerlynck *et al.* 1998), suggesting that water status ultimately controls photosynthesis. The spring plants grown in the chamber would not have been water stressed compared to the control plants, but their site of origin has a much higher temperature than the control site due to the gases causes a greenhouse effect in the spring. Therefore the spring plants could have innate adaptations to water status leading to this expression profile. It has also been shown that in mature forest trees after three years of exposure to elevated $CO_2$ six (out of six analysed) species no longer down regulated leaf

photosynthesis (Zotz, Pepin *et al.* 2005). This shows that in the long term the previously documented adaptation to elevated $CO_2$ might not be valid as the 'normal' response, and this is why we see no down regulation of gene expression in photosynthesis related genes in the spring plants.

The group 'mitochondria,' was also over represented in the GO analysis in response to elevated $CO_2$ in both spring and control (0.00118 and $2.74e^{-06}$, respectively) plants. Mitochondria are the site of respiration, making this category of interest. The respiration pathway in response to elevated $CO_2$ has undergone extensive investigation, and in terms of gene expression recent development has shown the variety of gene expression changes that can occur depending on leaf development (Markelz, Vosseller *et al.* 2014). The respiration pathway was investigated using MapMan, and the control plants showed an overall down regulation of gene expression across the pathway (decrease of 11.87%), whereas the spring plants showed an overall up regulation of gene expression (increased by 7.36%). When looking at the significance of the genes in the respiration pathway only one showed a significant treatment effect (≤0.05), whereas the majority showed a significant location effect (≤0.05). This suggests it a local adaptation to the spring site which is causing this shift in expression. The previously documented response is an increase in respiration gene expression in response to elevated $CO_2$ (Leakey, Xu *et al.* 2009), which interestingly we do not see in the control plants. An increase would be expected, as all other responses in the control plants follow the previous literature. What can be seen though is an apparent multigenerational adaptive phenotype where the continual exposure to elevated $CO_2$ in the spring plants has produced an increase in gene expression in the respiration pathway, compared to the control plants. This could suggest the spring plants are metabolising more and so you would expect larger plants in the spring. The results in Chapter 2 showed smaller leaves and no change in above ground biomass in the spring plants in response to elevated $CO_2$, which doesn't tie in, and so remains to be elucidated.

The stomatal phenotype was one of the most interesting results from the chamber experiment, and so the stomatal patterning pathway was investigated in more detail. Eight genes were identified as having significantly different expression in the stomatal patterning pathway, showing changes across the pathway. Several genes that act as negative regulators early in the pathway (ER, BCA1, BCA4) showed a decrease in expression in the spring plants in response to elevated $CO_2$, coinciding with the increased stomatal index phenotype, but interestingly they were even more down regulated in the control plants in response to elevated $CO_2$, which was not expected based on the phenotype they produced. BCA1 and BCA4 showed the same expression pattern in

response to elevated $CO_2$ in both populations, which is expected as the two genes act together (Hu, Boisson-Dernier *et al.* 2010), and the decreased stomatal density phenotype is only exhibited in a double mutant in the two genes, and not in single mutants of either BCA1 or BCA4. The stomatal density and index in the control plants showed no significant difference in response to elevated $CO_2$, and so the expression pattern does not tie in. As the stomatal patterning pathway is complicated, the combined expression of several other genes could be responsible for the increase in the spring plants in response to elevated $CO_2$ (Torii 2012).

Two genes which act during the asymmetric division stage of stomatal patterning (SCRM2 and CDBK1;1 ; refer to Figure 3.15), showed expression patterns which could explain the increased stomatal index phenotype in the spring plants in response to elevated $CO_2$. They are both positive regulators, and both show an increase in expression under elevated $CO_2$ in the spring plants, and no response in the control plants in response to elevated $CO_2$ (both with a significant ($P < 0.05$) location effect). This indicates that the extra proliferation of stomata in the spring plants may be occurring during this stage in the stomatal patterning pathway. Plants with reduced CDKB1;1 activity have a decreased stomatal index because of an early block of meristemoid division and inhibition of satellite meristemoid formation (Boudolf, Barrôco *et al.* 2004), the increased expression in the spring plants would therefore indicate the opposite, an increase in stomatal index. Expression of CDKB;1 also is correlated with epidermal cell size (Boudolf, Barrôco *et al.* 2004), with increased expression leading to smaller epidermal cells, which again is what was found in the chamber experiment (Chapter 2) in the spring plants, highlighting the importance of this gene for the spring plants stomatal patterning pathway. The gain-of-function mutant of SCRM2 is complete over proliferation of stomata, so much so they cover the whole surface, highlighting the influence of SCRM2 in stomatal initiation (Hofmann 2008). Conversely, successive loss of *SCRM* and *SCRM2* recapitulated the phenotypes of *fama*, *mute*, and *spch*, indicating that a gene dosage of *SCRM* and *SCRM2* determines successive initiation, proliferation, and terminal differentiation of stomatal cell lineages (Kanaoka, Pillitteri *et al.* 2008). We could not identify a transcript that matched to SCRM, but would expect expression to be similar if SCRM and SCRM2 act in tangent as they do in *Arabidopsis*. The importance of SCRM2 and CDKB1;1 in the stomatal patterning pathway and the expression we see from the chamber experiment suggest the changes at this part of the pathway are key for why we see the stomatal phenotype we do in spring originating *P. lanceolata*.

Genes associated with cell wall synthesis were extracted from the whole gene list, and expression patterns were investigated, with a substantial change being observed between control and spring

plants in response to elevated $CO_2$. Control plants expressed a 10.46% decrease in response to elevated $CO_2$ across all cell wall regulating genes, whereas the spring plants expressed a 15.88% increase. This could be associated with the stomatal phenotype, as the spring plants appear to show more asymmetric divisions with greater numbers of stomata and smaller cells, and need to create more cell walls, as smaller epidermal cells have a total larger cell wall area than larger epidermal cells (Cutler, Rains *et al.* 1977).

The RT-PCR was used to confirm the expression in a selection of genes to give confidence in the RNA-Seq data. The results showed the patterns were quite varied in the real-time PCR, but on the whole the pattern of expression was similar to that of the RNA-Seq (Figure 3.25). Due to the intricacy of RNA expression it is difficult to produce exact results, and the RNA used can degrade over time, and so it would not be expected to have exact matches. Taking this into account the real-time does not disagree with the RNA-Seq data and so would indicate the RNA-Seq data can be trusted.

Due to the complicated nature of using NGS, it is important to fully understand what our data represents, and to expect complications. Although it can be suggested from the data that the genes identified are due to adaptations and acclimations, it cannot be ruled out at this point that epigenetic parental effects are responsible, as the seed used was collected directly from the spring and control sites. It does however; provide a substantial set of genes that are potentially important in adaptation and acclimation to elevated $CO_2$, which could be utilised for enhancing capacities to elevated $CO_2$ in other species. It important to ensure the use of the best methods when analysing our data, but it is not as simple as choosing the most popular method, as it has been shown that for instance, different assemblers perform better depending on the organism used (Baker 2012). For the human chromosome 14, one widely used assembler omitted only 1.7% of the reference assembly sequence. Although it had omitted the least for this metric with the human chromosome, the program omitted the most for *Rhodobacter sphaeroides*, omitting 7.5% of the genome (Baker 2012). It has to be considered that these assemblers were genome assemblers, which do differ to transcriptome assemblers, but variation has also been observed in transcriptome assemblers (Zhao, Wang *et al.* 2011). This highlights the variability between, and we have to take into account we may be missing data once assembled. Trinity has been identified as one of the consistently top assemblers (Grabherr, Haas *et al.* 2011), and so we can be confident that at this point in time, the coverage obtained is at its optimum. There are currently a number of on-going projects that are analysing all available assemblers and which hold competitions such as 'the Assemblathon', dnGASP and GAGE (Salzberg, Phillippy *et al.* 2012), not to identify the best

assembler at a particular point in time, but to find ways to assess and improve assemblers in general. Keeping on top of the information these competitions provide will give a good insight and understanding to move forward with any data analysis. The field of NGS itself is still also continually advancing, producing cheaper and more efficient methods of sequencing. This year Life Technologies released a machine that is able to sequence the entire human genomes in one day for only $1,000. Not only that, but Illumina have also released a machine of its own with the same capabilities. So there is now the potential to sequence more samples to further analyse these changes for a fraction of the cost.

The RNA-Seq results have provided us with an understanding of the stomatal pathway, and a strong potential explanation to the phenotypes identified. It has also enlightened the important biochemical pathways that could be associated with an increase to elevated $CO_2$. This is important as the changes in biochemical pathways could be the reasoning behind the stomatal phenotype, and also responsible for the other phenotypic changes. The gene expression alone does not necessarily mean the physiological mechanisms will be affected, but is a strong indication that there may be an effect, and a platform to base further investigations into these pathways.

# Understanding the Gas Exchange Response of Spring and Control Plants to Ambient and Elevated $CO_2$

## 4.1 Introduction

The previous chapter highlighted clear areas of interest in the genome for further analysis, and the photosynthetic pathway, with altered gene expression for light harvesting and RuBisCO, being one such place. RuBisCO acts as part of Calvin cycle (Raines 2003), which is the light independent reaction of photosynthesis. The role of this enzyme is to catalyse a reaction between $CO_2$ and RuBP to create 2 molecules of 3-phosphoglyceric acid (3-PGA). This process is referred to as carbon fixation, as the $CO_2$ is fixed from an inorganic form into organic molecules. Three molecules of $CO_2$ and RuBP are fixed in one cycle, forming six molecules of 3-PGA. ATP and NADPH are then used to convert the 3-PGA into 6 molecules of glyceraldehyde 3-phosphate (G3P). This is known as the reduction part of the reaction as it involves the gaining of electrons by 3-PGA. The G3P molecules then leave the Calvin cycle to be used in the formation of compounds for the plant such as glucose. Six molecules of G3P are created per cycle, and five are transported out of the Calvin cycle and one remains and is used to regenerate RuBP in preparation for more $CO_2$ to be fixed. The Calvin cycle is illustrated in Figure 4.1.

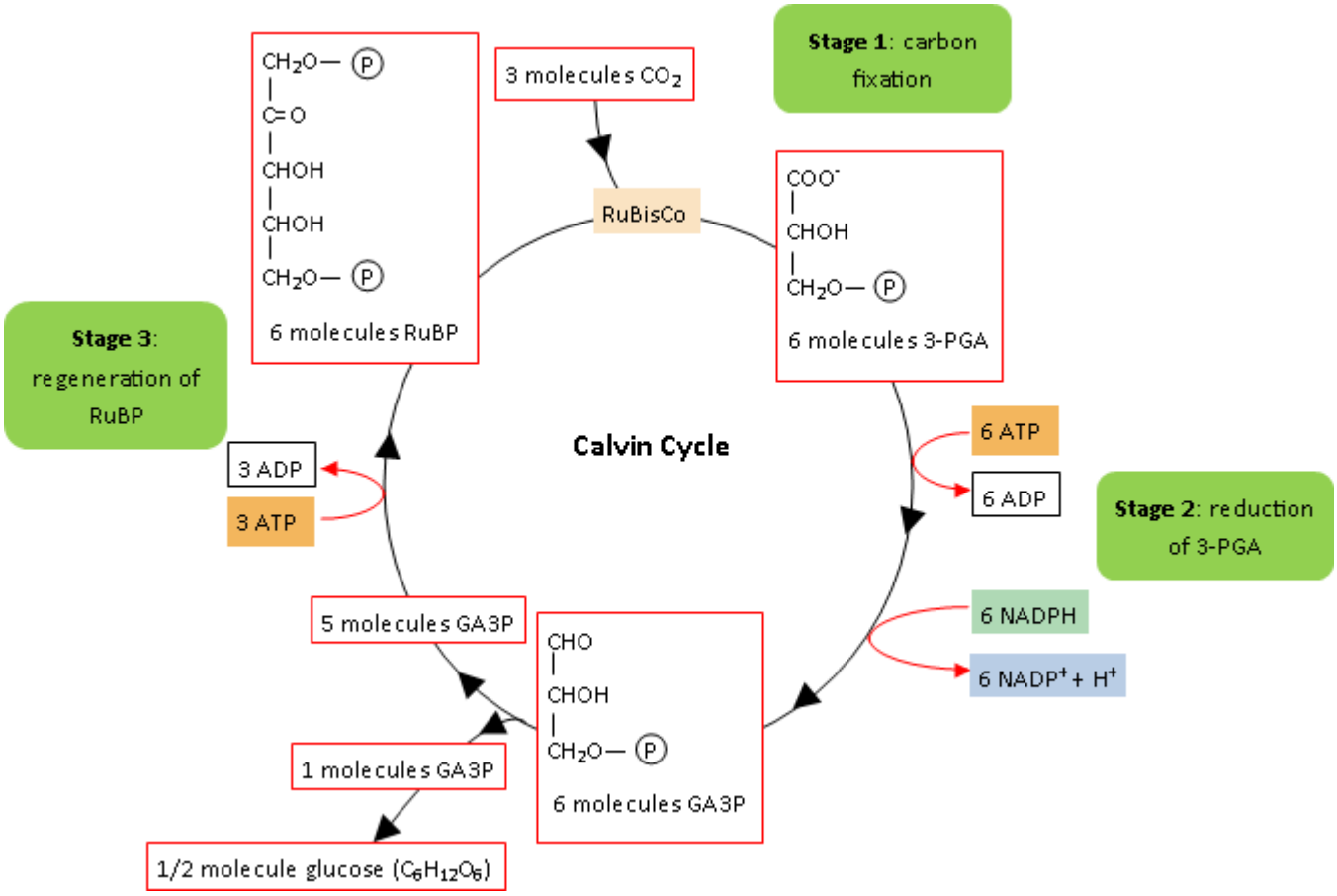



Figure 4.1 – The three stages of the Calvin cycle. Stage 1 - the enzyme Rubisco incorporates carbon dioxide into an organic molecule, 3-PGA. Stage 2 - the organic molecule is reduced using electrons supplied by NADPH and ATP. Stage 3 - RuBP, the molecule that starts the cycle, is regenerated so that the cycle can continue. Only one carbon dioxide molecule is incorporated at a time, so the cycle must be completed three times to produce a single three-carbon GA3P molecule, and six times to produce a six-carbon glucose molecule.

Although it is known as the light independent reaction, the Calvin cycle is still dependant on the light dependant reaction to create the ATP and NADPH to reduce the 3-PGA. The light dependent reaction uses light energy to split water and extract electrons into photosystem II (PS II) (Renger and Renger 2008). These electrons are then transported to photosystem I (PS I) via cytochrome b6f to be re-energised (Eberhard, Finazzi *et al.* 2008). This is known as the electron transport chain. The high energy electrons are then used to reduce $NADP^+$ into NADPH in PS I. As the electron is transported through cytochrome b6f, one of two reactions can happen. One is non-cyclic photophosphorylation, where cytochrome b6f uses energy from the electrons from PS II to pump hydrogen ions from the lumen of the stroma to allow ATP synthase to attach a third phosphate group to ADP to form ATP. The second is cyclic photophosphorylation, where cytochrome b6f uses the energy of the electrons from both PS I and PS II to create more ATP and stop the production of NADPH to maintain a balance of the two molecules. An overview of the light dependent reaction can be seen in Figure 4.2.

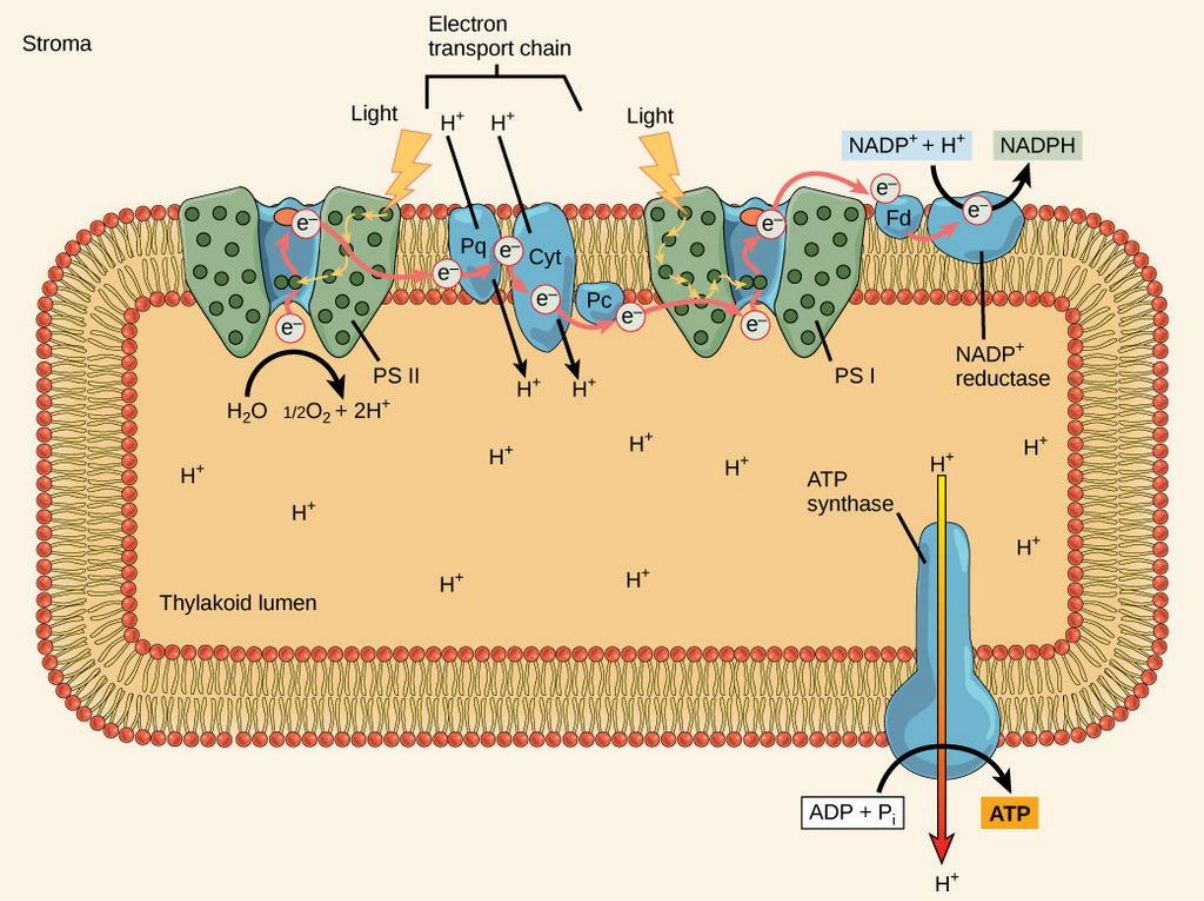



Figure 4.2 - In the photosystem II (PSII) reaction centre, energy from sunlight is used to extract electrons from water. The electrons travel through the chloroplast electron transport chain to PS I, which reduces $NADP^+$ to NADPH. The electron transport chain moves protons across the thylakoid membrane into the lumen. At the same time, splitting of water adds protons to the lumen while reduction of NADPH removes protons from the stroma. The net result is a low pH in the thylakoid lumen and a high pH in the stroma. ATP synthase uses this electrochemical gradient to make ATP (Berg 2007).

RuBisCO exists in two forms, one a dimer and the other a hexadecamer. The form which is present in all higher plants is the hexadecamer form (Kellogg and Juliano 1997). This form consists of eight large subunits (rbcL) and eight small subunits (rbcS). The rbcL's have a mass between 50-55 kDa whilst the rbcS's have a mass between 12-18 kDa, giving a total molecular mass of around 560 kDa for the complete protein (Andersson and Backlund 2008). The rbcL's are transcribed and translated within the chloroplasts, whereas the rbcS's are translated on cytoplasmic ribosomes as a precursor that is targeted to the plastids (Bowman, Patel *et al.* 2013). The active sites of RuBisCO are located on the large subunits, between the dimer pairs, thus the catalytic properties of RuBisCO are dependent on the large subunits. rbcS is not essential for catalysis as several configurations of rbcL's have been shown to retain carboxylase activity and have unperturbed specificity , such as the *Synechococcus* large subunit octamer (Andrews 1988). On the other hand, hybrids containing foreign small subunits and enzymes with mutations in the small subunit often display altered holoenzyme stability and/or specificity (Spreitzer 2003). Furthermore, interspecific hybrid enzymes have indicated that rbcS's are required for maximal catalysis. Plants contain families of differentially expressed rbcS, raising the possibility that these subunits may regulate the structure or function of RuBisCO.

Under elevated $CO_2$ the efficiency of RuBisCO, and therefore photosynthesis, is increased due the increased ratio of $CO_2$ to $O_2$ in the atmosphere. $CO_2$ and $O_2$ compete to bind to RuBisCO, and $O_2$ is often oxygenated as part of the reaction (Bowes 1991). This inhibits carboxylation of $CO_2$, and reduces photosynthetic productivity. It has been shown that RuBisCO has a higher affinity to bind to $CO_2$ when concentrations of $O_2$ and $CO_2$ are equal, especially in rbcS's, where twice as much $CO_2$ binds over $O_2$, compared to the rbcL's (Van Lun, Hub *et al.* 2014). Unfortunately this higher affinity is not enough to negate the effects of oxygenation. Under elevated $CO_2$ the likely hood of oxygenation is reduced, increasing the efficiency of carboxylation, therefore increasing photosynthetic productivity (Drake, Gonzalez-Meler *et al.* 1997).

The spring plants from this study have been grown in elevated $CO_2$ for many generations, and evidence from the RNA-Seq data suggests the spring and control plants have pronounced differences, including genes underpinning the photosynthetic processes described here. It is therefore important to investigate how elevated $CO_2$ has an effect on photosynthesis over multiple generations. Studies which looked at this long term effect, such as in cotton, found that plants grown at elevated $CO_2$ had lower photosynthetic capacity than ambient grown plants when measured at numerous $CO_2$ concentrations (Delucia, Sasek *et al.* 1985). Rey and Jarvis (1998) looked at how elevated $CO_2$ would affect photosynthesis in *Betula pendula* (Birch) over four

growing seasons. They found elevated $CO_2$ concentration stimulated photosynthesis by 33% on average over the fourth growing season. However, comparison of maximum photosynthetic rates at the same $CO_2$concentration (350 or 700 µmol $mol^{-1}$) revealed that the photosynthetic capacity of trees grown in an elevated $CO_2$ concentration was reduced (Rey and Jarvis 1998). Analysis of the A/Ci curves showed decreases in the initial slope (carboxylation capacity) and the plateau of the curve (RuBisCO regeneration capacity) in response to elevated $CO_2$. A result that has now been observed on numerous previous occasions (Moore, Cheng *et al.* 1999, Rogers and Humphries 2000), and also summarised in the meta-analysis of Long *et al.* (2004). This indicates that acclimation to elevated $CO_2$ concentration involves decreases in carboxylation efficiency and RuBisCO regeneration capacity. They also noted that despite a 21% reduction in stomatal conductance in response to the elevated $CO_2$ treatment, stomatal limitation of photosynthesis was significantly less in the elevated, than in the ambient, $CO_2$ treatment. Thus, after four growing seasons exposed to an elevated $CO_2$ concentration in the field, the trees maintained increased overall photosynthesis, although their photosynthetic capacity was reduced compared with trees grown in ambient $CO_2$. These results suggest the increased efficiency of RuBisCO has allowed for the down regulation of the key enzymes of the photosynthetic carbon reduction cycle. This in turn allows for the resource efficiency of the plant to be increased, by less investment in these enzymes and increasing light use efficiency, whilst increasing net photosynthesis (Drake, Gonzalez-Meler *et al.* 1997). The mechanism which controls the regulation of photosynthesis has been linked with the source-sink balance of carbohydrates (Koch 1996). Sink regulation of photosynthesis is highly dependent on the physiology of the rest of the plant. This physiological state regulates photosynthesis through signal transduction pathways that co-ordinate the plant carbon : nitrogen balance, which match photosynthetic capacity to growth and storage capacity and underpin and can override the direct short-term controls of photosynthesis by light and $CO_2$ (Paul and Foyer 2001). Increased amount of assimilates from elevated $CO_2$ causes down regulation of the photosynthesis pathway due to an imbalance of carbon, between source and sink, which agrees with the experimental data (Arp 1991).

Photosynthetic measurements have been previously measured using *P. lanceolata* samples from the spring in Bossoletto (Bettarini, Vaccari *et al.* 1999). They used 20 plants growing in the spring and 20 growing in a nearby control site for gas exchange measurements. One mature healthy leaf was selected from each plant and leaf photosynthesis measured using the portable photosynthesis system, CIRAS-1 (PPS, Hitchin, UK). Assimilation (A), intercellular $CO_2$ concentration ($C_i$) and stomatal conductance ($g_s$) was measured. At that time (1995), the equipment was not as advanced as is currently available, and so these measurements did not

create an A/Ci curve as would be the case now. The measurements were conducted at a constant vapour pressure deficit (1 kPa), saturating irradiance (1600 µmol $m^{-2}s^{-1}$) and at external $CO_2$ concentrations ($C_a$) of 355 and 700 µmol $mol^{-1}$. Bettarini *et al.*'s (1999) results can be seen in Table 4.1.

| | ***$C_a$*** | ***Spring*** | ***Control*** | ***P value*** |
|---|---|---|---|---|
| A (µmol $m^{-2}s^{-1}$) | 355 | 10.8 ± 0.8 | 16.1 ± 1.1 | <0.001 |
| | 700 | 18.4 ± 1.6 | 26.8 ± 1.9 | <0.001 |
| $C_i$ (µmol $mol^{-1}$) | 355 | 257 ± 6.6 | 241 ± 2.7 | <0.001 |
| | 700 | 489 ± 11.9 | 468 ± 6.9 | <0.001 |
| $C_i/C_a$ | 355 | 0.72 | 0.68 | <0.001 |
| | 700 | 0.70 | 0.67 | <0.001 |

Table 4.1 – Gas exchange parameters (A = Assimilation, $C_i$ = internal $CO_2$ concentration, $C_a$ = external $CO_2$ concentration) measured on leaves of *P. lanceolata* from the $CO_2$ spring in Bossoleto by Bettarini *et al*. (1999).

The results showed that the plants originating from the spring had a significantly lower photosynthetic rate than those from the control site (-33% when $C_a$ = 355 µmol $mol^{-1}$ and -31% when $C_a$ = 700 µmol $mol^{-1}$). Van Gardingen *et al.* (1997) showed the photosynthetic rate and capacity of *P. australis* was again down regulated within the Bossoleto spring compared to a nearby control site. These results agree with the findings that the increased efficiency of RuBisCO under elevated $CO_2$ leads to down regulated key enzymes within the pathway to increase resource efficiency. Net photosynthesis was not measured in these studies so it is unclear whether this would be increased compared to the control site as might be expected, but is likely to be the case. Several other species measured within naturally elevated $CO_2$ springs have been shown to show no difference when compared to control sites. Within the Bossoleto spring *Q. pubescens* showed no difference in photosynthetic rate or capacity in two separate studies (Johnson, Michelozzi *et al.* 1997, Van Gardingen, Grace *et al.* 1997). At two different sites in Italy, Baccaiano and Solfatara, *A. canina* and *P. major* seed was collected from naturally elevated $CO_2$ springs and nearby control sites, then grown in both controlled environments, in both elevated and ambient $CO_2$ (Fordham and Barnes 1999). Photosynthetic analysis of all experimental groups

showed no significant differences in photosynthetic characteristics. Studies utilising the Japanese springs, again showed no evidence of down regulation of the photosynthetic machinery in seven different species inside a naturally elevated $CO_2$ spring, compared to an ambient $CO_2$ control site (Onoda, Hirose *et al.* 2007, Onoda, Hirose *et al.* 2009). These studies showed that in response to elevated $CO_2$, there was a significant increase in photosynthetic capacity of both plants originating from the spring and the control site, but no adaptation between the sites was apparent. These inconsistencies mean that a down regulation of photosynthetic machinery may not necessarily be the only response to multi-generational elevated $CO_2$, and suggest other variables within the springs may be determining this phenotype.

A number of variables can interfere with RuBisCO activity under elevated $CO_2$, both negatively and positively. One variable which has been noted is Nitrogen (N) content, which when varied has resulted in a wide variety of A/Ci relationships (Pettersson and McDonald 1994). The role of N in the carbon : nitrogen balance indicates it would have a major role in photosynthesis regulation, and it has been suggested that genes linked with concentrations of N produce signals to regulate photosynthesis when the balance is affected (Paul and Foyer 2001). Wong (1979) showed in maize and cotton that under low N (0.06mM $NO_3^-$) plants showed less photosynthetic activity under elevated $CO_2$, as expected, but under high N (24mM $NO_3^-$) the photosynthesis seemed unchanged or even increased (Wong 1979). Other studies contradict this, showing low/high levels of N have no feedback on photosynthetic rate (Sage, Sharkey *et al.* 1990). It is proposed that the effect of N on photosynthetic rate is dependent on the plants ability to store nitrogen, and its ability to acclimate to any nitrogen differences that may occur due to increases in growth, leading to the variety of A/Ci curves under these conditions. Onoda *et al.* (2009) furthered their work using the Japanese springs; where there was no apparent photosynthetic adaptation in seven species, but found that varying N concentration has a major influenced on the photosynthetic rate. As N availability increased, photosynthetic rate increased in elevated $CO_2$ conditions. On the other hand, when N availability decreased photosynthetic rate was decreased in elevated $CO_2$ conditions (Osada, Onoda *et al.* 2010).This could explain the differences in photosynthetic response between the Japanese and the Italian springs, but is unlikely to explain differences within the Bossoleto spring, as availability of N should be similar.

The same principles of N concentration can also be expanded to innate nutrient content, where cereal crops have been shown to have no acclimation to elevated $CO_2$. It is suggested this response occurs due to the balanced supply of nutrients and carbohydrates contained in the crops. In a long-term (full growing season) study with winter wheat at two levels of N application and

two levels of $CO_2$, there was no evidence for acclimation of photosynthesis to $CO_2$ (Mitchell, Mitchell *et al.* 1993). This was explained as a consequence of sink strength in a cereal crop being large both during the vegetative phase (when it can respond to an increased assimilate supply by increased tilling), as well as during the reproductive phase when the ear is a large sink. This may explain why, in most studies of cereals where gas exchange has been measured, no negative feedback responses have been found (Pettersson, Lee *et al.* 1993). Another variable that can influence the curve through reduction in the efficiency of RuBisCO is salinity. High concentrations of NaCl have been shown to reduce the effectiveness of RuBisCO via regeneration capacity and activity leading to lower photosynthetic rates (Seemann and Sharkey 1986).

The isotopes of carbon itself can have an effect on RuBisCO, and can also be used as an indicator of photosynthetic activity. C12 and C13 are stable isotopes of carbon which both occur naturally in the environment and do not change over time. However, because of a small difference in their molecular weights due to an extra neutron in the C13 isotope, these two isotopes are processed differently in the photosynthetic pathway and so can be used to measure photosynthetic activity. RuBisCO preferentially binds to C12 as the lighter isotope is easier and quicker to use in biological processes. This is only true for C3 plants, as in C4 plants the more efficient PEP-carboxylase has no preference over C12 and C13. This affinity produces different internal to external concentrations of C13 at different photosynthetic rates, which can indicate a higher or lower capacity of the chloroplasts to fix atmospheric $CO_2$ depending on the ratio (Farquhar, Ehleringer *et al.* 1989). This ratio is often referred to as the carbon discrimination (Δ) ratio. Bettarini *et al.* (1999) showed that the carbon discrimination was lower in species from the spring, and suggested that this indicated that carboxylation was restricted (relative to diffusion through the stomata) leading to a lower photosynthetic capacity. Miglietta *et al.* (1998) also showed this was present, but concluded that this affects groups of species differently. Grassland species had significantly reduced carbon discrimination within the spring compared to a control site, but ruderal species showed no significant difference. Tree species showed varied results, and interestingly, *Q. pubescens* (the species which showed no down regulation of photosynthetic capacity within the spring), did show a significantly lower carbon discrimination within the spring compared to a control site. If this was the variable that was causing the phenotype of *Q. pubescens*, you would not expect this result. Carbon discrimination under elevated $CO_2$ has been shown to be alleviated by high levels of N availability (Bettarini, Calderoni *et al.* 1995). Using *Erica arborea* in a spring near Naples (Italy), it was shown that carbon discrimination increased as $CO_2$ concentration increased, but this response did not occur when N availability was increased. This again suggests the N content

affected the source-sink balance is stopping the down regulation of aspects in relation to photosynthesis.

Elevated temperature has been shown to have a significant negative effect on RuBisCO efficiency (Delucia, Sasek *et al.* 1985, Crafts-Brandner and Salvucci 2000). Crafts-Brandner and Salvucci (2000) showed a decrease in RuBisCO activation occurred when leaf temperatures exceeded 35°C, whereas the activities of isolated activase and RuBisCO were highest at 42°C and >50°C, respectively, indicating RuBisCO activase is the limiting factor at high temperatures. Sage (2001) also showed that low temperature can have a negative effect on photosynthesis, showing that below 20 °C, C3 photosynthesis at ambient and elevated $CO_2$ is often limited by the capacity to regenerate phosphate for photophosphorylation. Long (2006) used simulations from the equations which A/Ci curves are based on to predict what would happen under elevated $CO_2$ under different temperatures in C3 plants. The simulations show the increase, with elevation of $CO_2$ from 350 to 650 µmol $mol^{-1}$, in light saturated rates of $CO_2$ uptake (Asat) and maximum quantum yields rise with temperature. An increase in $CO_2$ from 350 to 650 µmol $mol^{-1}$ can increase Asat by 20% at 10°C and by 105% at 35°C, and can raise the temperature optimum of Asat by 5°C (Long 1991). These results confirm experimental data shown in other studies (Zhang and Dang 2013) who found maximal carboxylation rate (Vcmax), PAR-saturated electron transport rate (Jmax) and triose phosphate utilization (TPU) varied with $CO_2$, and the Vcmax and Jmax were significantly higher at 37$^{o}$C than at 26$^{o}$C under elevated $CO_2$ in white birch seedlings.

Any of these variables could be responsible for no down regulation of photosynthesis being seen on multiple occasions, but the phenotype could also represent a different adaptation to elevated $CO_2$, where a down regulation no longer occurs. From these studies it is hard to determine, but using samples from a controlled environment study could provide answers to explain this phenotype, and show the true adaptation to elevated $CO_2$.

## 4.1.1 A/Ci Curve Analysis

A/Ci curves work by taking various aspects of the biochemistry of photosynthetic carbon assimilation and integrating them into a form compatible with studies of gas exchange in leaves (Farquhar, Caemmerer *et al.* 1980). There are three variables that limit the carboxylation rates across an A/Ci curve (Manter and Kerrigan 2004). The first is the amount, activity, and kinetics of RuBisCO (Vcmax), and the second is the rate of RuBisCO regeneration supported by electron transport (Jmax). The third is triose phosphate availability. Triose phosphate is used to regenerate the enzyme $CO_2$ is fixed too RuBP, hence limiting photosynthesis when availability is limited. The production of triose phosphate is cyclic with the photosynthesis cycle, and is produced as a product of photosynthesis, so this variable occurs only occasionally.

### 4.1.1.1 A/Ci Curve Calculations

A generalised A/Ci curve at ambient $CO_2$ was suggested by Long *et al.* (1993). This showed the limitations of $CO_2$ assimilation in standard conditions, and where they occur (Figure 4.3). This shows the rate which would occur if there was no stomatal or other gas phase diffusive limitation interpolated from the A/Ci curve, and so is a good basis to compare responses of the photosynthetic machinery to a change in $CO_2$ concentration.

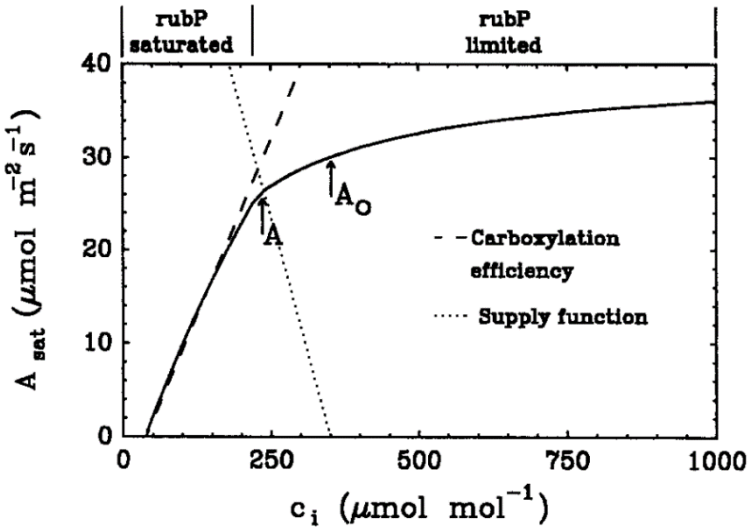



Figure 4.3 – A generalised A/Ci curve at ambient $CO_2$. Where A indicates the photosynthetic rate with a Ca (atmospheric $CO_2$ concentration) of 354 µmol $mol^{-1}$ and $A_o$ the rate with a Ci (intracellular $CO_2$ concentration) of 354 µmol $mol^{-1}$ (Long, Baker *et al.* 1993).

Two equations from Farqhar *et al.* (1980), and later modified by Harley and Sharkey (1991), can be used to estimate Vcmax and Jmax to fit photosynthetic data to an A/Ci curve. The first equation (Formula 4.1) calculates Wc (the RuBisCO limited-limited rate of carboxylation), which is determined by the Vcmax, and so the maximum Wc equals the Vcmax (Harley and Sharkey 1991). The second (Formula 4.2) calculates J by taking into account all parameters that effect the electron transport chain (Farquhar, Caemmerer *et al.* 1980). J is then related to Jmax according to Formula 4.3 (Bernacchi, Pimentel *et al.* 2003, Lenz, Host *et al.* 2010).

$$W_c = \frac{V_{max} C}{C + K_c (1+O/K_0)}$$

Formula 4.1 –Equation used to calculate Vcmax. Vmax is the maximum rate of carboxylation; $K_c$ and $K_0$ are the michaelis constants for $CO_2$ (C) and $O_2$ respectively.

$$J = \left( \frac{F'_m - F_s}{F'_m} \right) \times 0.5 \times PPFD \times \alpha_{leaf}$$

Formula 4.2 – Equation used to calculate J. $F_s$ is chlorophyll fluorescence measurements of steady-state fluorescence; $F'_m$ is maximal fluorescence; PPFD is photosynthetically active radiation; $\alpha_{leaf}$ is leaf absorbance.

$$\Theta PSII J^2 - (Q_2 + J_{max})J + Q_2 J_{max} = 0$$

Formula 4.3 – Equation relating J to Jmax. ΘPSII is the quantum yield of photosystem II; $Q_2$ is the irradiance absorbed by photosystem II (µmol $m^{-2}$ $s^{-1}$); J is the rate of electron transport (µmol $m^{-2}$ $s^{-1}$); Jmax is the maximum rate of electron transport (µmol $m^{-2}$ $s^{-1}$).

These equations were modelled to specifically fit C3 species photosynthetic response curves by Sharkey *et al.* (2007). A/Ci curves have been carried out on a wide variety of species to give an insight into the variations in the capacities of the biochemical reactions that drive photosynthesis. In 1993, Wullschleger analysed A/Ci curves from 109 C3 species at elevated $CO_2$ to see how they differed (Wullschleger 1993). He found that average maximum Vcmax was 64µmol $m^{-2}$ $s^{-1}$ across all species. Interestingly though, the range of Vcmax was from 194µmol $m^{-2}$ $s^{-1}$ (in agricultural species *Beta vulgaris*) to just 6µmol $m^{-2}$ $s^{-1}$ (in coniferous species *Picea abies*). This range was also

seen in the Jmax, with the highest at 372µmol $m^{-2}$ $s^{-1}$ (in desert annual *Malvastrum rotundifolium*) and the lowest at 17µmol $m^{-2}$ $s^{-1}$ (again in *P. abies*). This shows the range of photosynthetic capability between species, and suggests the ability to utilise elevated $CO_2$ concentrations is dependent on differences in the biochemical pathway of each species.

### 4.1.2 AQ Curve Analysis

Plant response to photon flux density (PPFD) has been subjected to much investigation over the last century (Björkman 1981). AQ curves measure plant assimilation rate under a range of PAR intensities to see how plants can acclimate to high light intensities. This result, along with A/Ci curves give a more holistic view of the photosynthetic pathway, as it helps identify the underlying biochemical limitations underpinning observed rates of photosynthesis. Under non-limiting $CO_2$ concentrations increasing light intensity increases photosynthetic capacity, but this effect becomes diminished as $CO_2$ concentrations become the limiting factor (Wilson and Cooper 1969). The saturation point (Asat) is used to identify the point at which light becomes saturated and the plant can no longer use the excess light to increase assimilation.

In elevated $CO_2$, Asat control is primarily determined by the rate of RuBP regeneration, following the steady-state model of photosynthesis (Farquhar, Caemmerer *et al.* 1980). Due to this the control in not down to the control of RuBisCO, and so Asat has the potential to be unaffected even when a down-regulation of RuBisCO occurs (Long and Drake 1991, Sage 1994). It has been shown that maximum Asat increases under elevated $CO_2$ (Kubiske and Pregitzer 1996, Bernacchi, Morgan *et al.* 2005). This increase has been shown to be correlated with N content, and Asat increases are associated with high N conditions (Medlyn, Badeck *et al.* 1999). This relationship appears similar to relationship of N and A/Ci curves, which is expected, as Asat is dependent on the rate of $CO_2$ limited photosynthesis. In high N concentrations it appears both photosynthesis and Asat can increase whilst RuBisCO is down regulated, but the source-sink balance affects this response under low N concentrations. AQ curves have been previously carried out within the Bossoleto spring investigating the green-algal lichen *Parmelia caperata* (Balaguer, Manrique *et al.* 1999). It was shown that Asat under both ambient and $CO_2$ saturating conditions was similar in plants originating from the spring compared to the control site. There was no evidence of photosynthetic differences between the sites, but a decrease in RuBisCO was noted, in keeping with dependence of RuBP regeneration on changes in Asat.

### 4.1.3 Stomatal Response

Stomatal response is measured by measuring the exchange of water vapour through the stomata (Pearcy, Schulze *et al.* 2000). As previously stated, the most documented response to elevated $CO_2$ is a decline in stomatal conductance following partial stomata closure (Morison 1998, Ainsworth and Rogers 2007). Ainsworth and Rogers (2007) carried out a meta-analysis on stomatal conductance in response to elevated $CO_2$ (Figure 4.4). They showed the magnitude of decline in stomatal conductance varied, but there was always a significant decrease.

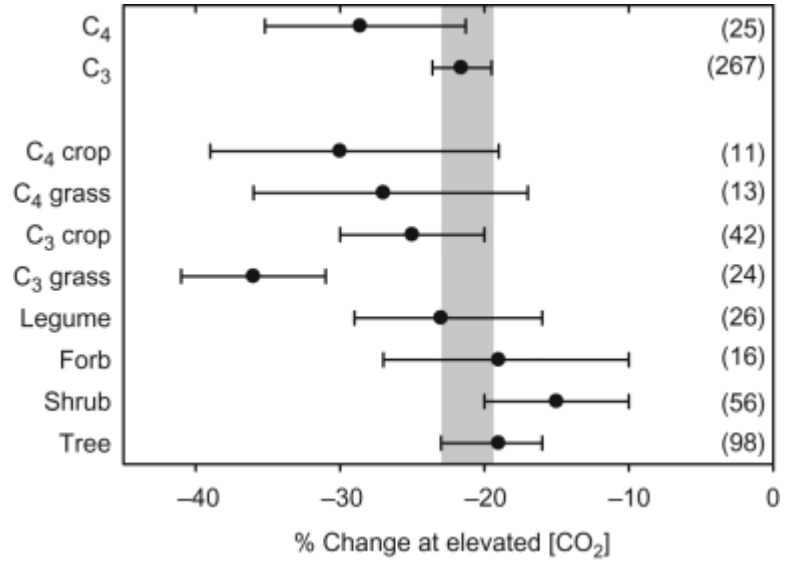



Figure 4.4 - Meta-analysis of the response of stomatal conductance to elevated $CO_2$ concentrations in free-air $CO_2$ enrichment experiments. The ambient and elevated $CO_2$ concentration for all studies averaged 366 and 567 $\mu$mol mol$^{-1}$, respectively. The grey bar represents the overall mean and 95% confidence interval (CI) of all measurements. The symbol represents the mean response (± 95% CI) of $C_3$ and $C_4$ species, and different functional groups. The degrees of freedom for each measurement are shown in parenthesis (Ainsworth and Rogers, 2007).

However, some species show no change in stomatal conductance in response to elevated $CO_2$. Saxe *et al.* (1989) found no significant decrease in trees, particular woody coniferous trees, showing there are exceptions to the rule. Several studies using long term elevated $CO_2$ FACE experiments showed that stomatal conductance does not acclimate independently of the other factors (Nijs, Ferris *et al.* 1997, Leakey, Bernacchi *et al.* 2006). Only under water stressed conditions did stomatal conductance acclimate (Medlyn, Barton *et al.* 2001), suggesting elevated $CO_2$ may only cause stomatal closure in the short term. Changes in the long term are therefore

likely to include both functional and morphological responses, including decreased stomatal numbers – a response commonly, but not consistently, seen in elevated $CO_2$. Onoda *et al.* (2010) showed that iWUE was significantly higher in plants originating from naturally elevated springs in Japan compared to a control site at both ambient and elevated $CO_2$. This indicates that an adaptation has occurred within the spring to allow for this increase in iWUE which could be due to effects on either, or both, of stomatal morphology and function.

It has been shown that smaller stomata react faster to environmental factors (Hetherington and Woodward 2003, Raven 2014), and so size may play a role in stomatal response. Raven (2014) showed that calculations of the influence of stomatal size on the potential rate of osmolarity increase, assuming size-independent ion influx rate per unit area of guard cell plasmalemma set at the value found in large (60 µm long) stomata, show that 10 µm long stomata could have at least a 6-fold higher rate of osmolarity increase. This could lead to a corresponding decrease in the time taken in going from the closed to the fully open state from about 1 hour to about 10 minutes. Drake *et al.* (2013) also showed a negative correlation between stomatal size and maximum rate of stomatal opening (Figure 4.5).

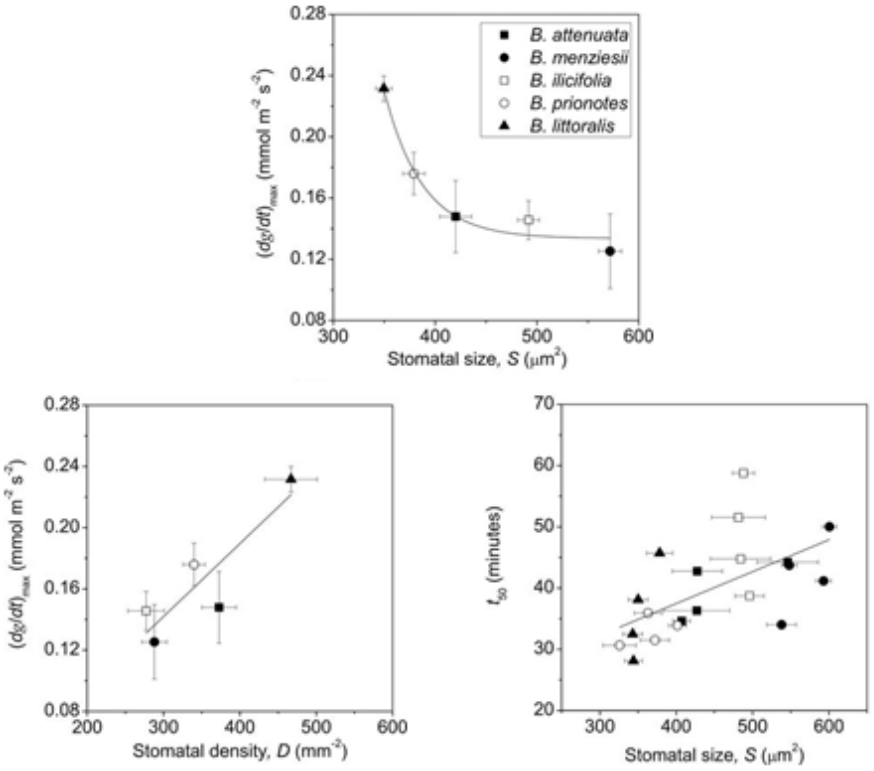



Figure 4.5 - The maximum rate of stomatal opening $(dg/dt)_{max}$ was negatively correlated with maximum stomatal size, S and positively correlated with stomtal density, D. The time to reach 50% of the range between $g_{min}$ (dawn) and $g_{op}$ ($t_{50}$) was positively correlated with stomatal size (Drake, Froend *et al.* 2013).

Although smaller stomata have this potential for faster response rates, it has been suggested that the size alone does not fully determine response rate (Franks and Farquhar 2007, Raven 2014). The concentration of the enzymes involved in opening and closing such as carbonic anhydrases have an affect (Raven 2014), as well as the shape of the stomata and the role of the subsidiary cells (Franks and Farquhar 2007). These factors have to be considered when investigating changes in stomatal conductance in response to elevated $CO_2$.

### 4.1.4 Aims

- Using seeds collected from the Bossoleto spring and a nearby control site, grow *P. lanceolata* in both ambient and elevated $CO_2$ in controlled environments to investigate both acclimation and adaptation to elevated $CO_2$ to confirm the phenotype discovered in Chapter 2.

- To investigate the physiological mechanisms underlying the gene expression changes identified, by carrying out A/Ci, AQ and stomatal response measurements on both spring and control plants, grown in both ambient and elevated $CO_2$ concentrations.

# 4.2 Materials and Methods

## 4.2.1 Chamber Experiment

After completion of the first experiment, plants were cut back and moved to the greenhouse to regrow for seed. Plants were separated into families using netting to avoid cross pollination between families. When possible seed was collected from as many plants as possible and stored in cold room at the University of Southampton (0˚C).

Seed was chosen from plants which had been grown under ambient $CO_2$ conditions in the previous chamber experiment (Chapter 2) to avoid any biases from the previous experiment. Where possible seed was also chosen from plants and families which had undergone the transcriptome sequencing to increase the probability any gene expression patterns would relate to phenotype.

Ten families were chosen in total, five families from the control site and five families from the spring site, of which four of the families were ones that were sequenced. Seeds were germinated in the chambers under ambient $CO_2$ conditions (384.15±0.45 µmol $mol^{-1}$). Temperature (Table 4.2), flow (Table 4.2) and light intensity (Figure 4.6) measurements were taken.

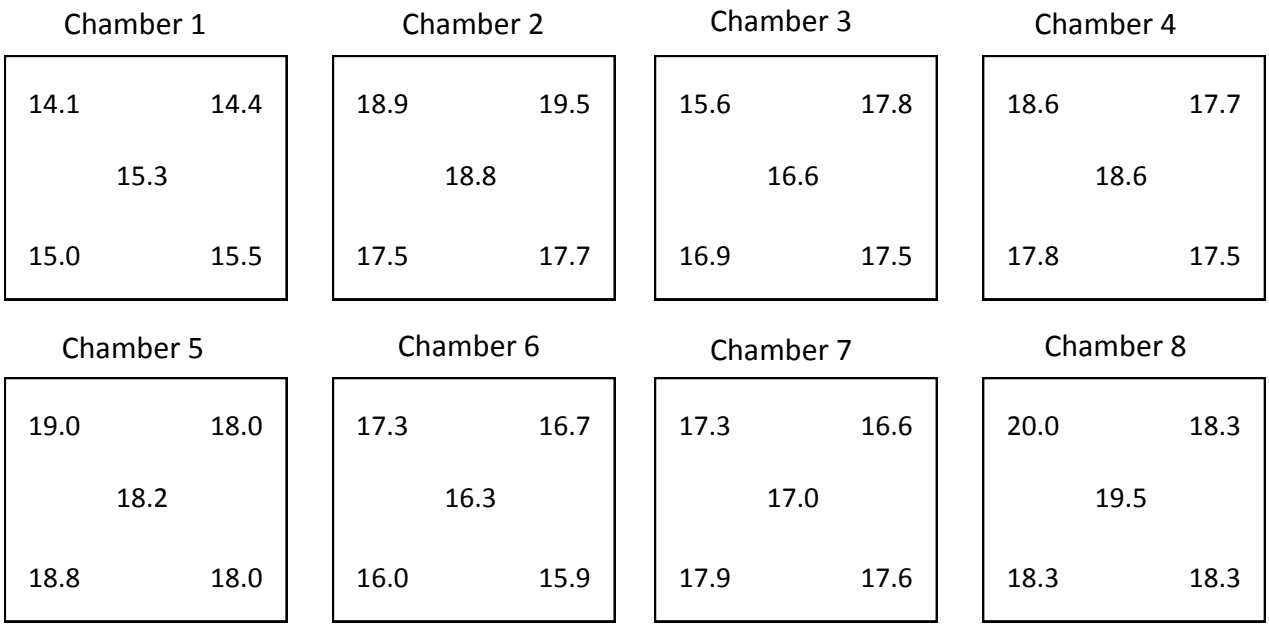



Figure 4.6 – The light intensity measurements taken in each chamber. The sensor was elevated to 21cm (to mimic the height of the plants). Units = x10µmol $m^{-2}$ $s^{-1}$.

| | Temperature (˚C) | Flow (m/sec) |
|---|---|---|
| **Chamber 1** | 22.93 ± 0.0003 | 2.53 |
| **Chamber 2** | 20.88 ± 0.0002 | 2.55 |
| **Chamber 3** | 22.37 ± 0.0003 | 2.53 |
| **Chamber 4** | 20.86 ± 0.0002 | 2.47 |
| **Chamber 5** | 22.78 ± 0.0003 | 2.45 |
| **Chamber 6** | 20.65 ± 0.0002 | 2.55 |
| **Chamber 7** | 21.51 ± 0.0002 | 2.48 |
| **Chamber 8** | 20.45 ± 0.0002 | 2.45 |

Table 4.2 – Table of the temperature measurements taken over the course of the experiment (one reading per hour), and the flow in each chamber measured at the vent at the start of the experiment.

Due to slight variation between chambers which could not be addressed due to the nature of the equipment a rigorous randomisation plan was followed throughout the experiment. This involved switching trays between chambers every 2 days, and then every 8 days switch the concentration of the chambers (not during germination as all chambers were ambient for this period) and then switch between the remaining chambers. This means every 16 days the trays have spent 2 days in every chamber. Every 6 days the pots were also randomised with 30 random movements within each tray to randomise the pots around the tray due to differences in light intensity and flow around individual chambers.

The seeds were planted in trays, each tray containing the seed from one family. The trays were filled with John Innes No. 2 compost and topped with 1cm of fine sand. After 2 weeks the seedlings were transferred to individual pots. These pots were filled with a 1:1:1 ratio of John Innes No. 2 compost, seed and molecular compost, and vermiculite. 56 seedlings from each family were potted. Once transferred to pots the chambers were switched to elevated and ambient conditions. The $CO_2$ condition's for the ambient chambers were 384.15±0.45 μmol $mol^{-1}$, and for the elevated chamber 803.92±3.46 μmol $mol^{-1}$. The plots of all the $CO_2$ data recorded for each chamber (post germination) can be seen in Figure 4.7.

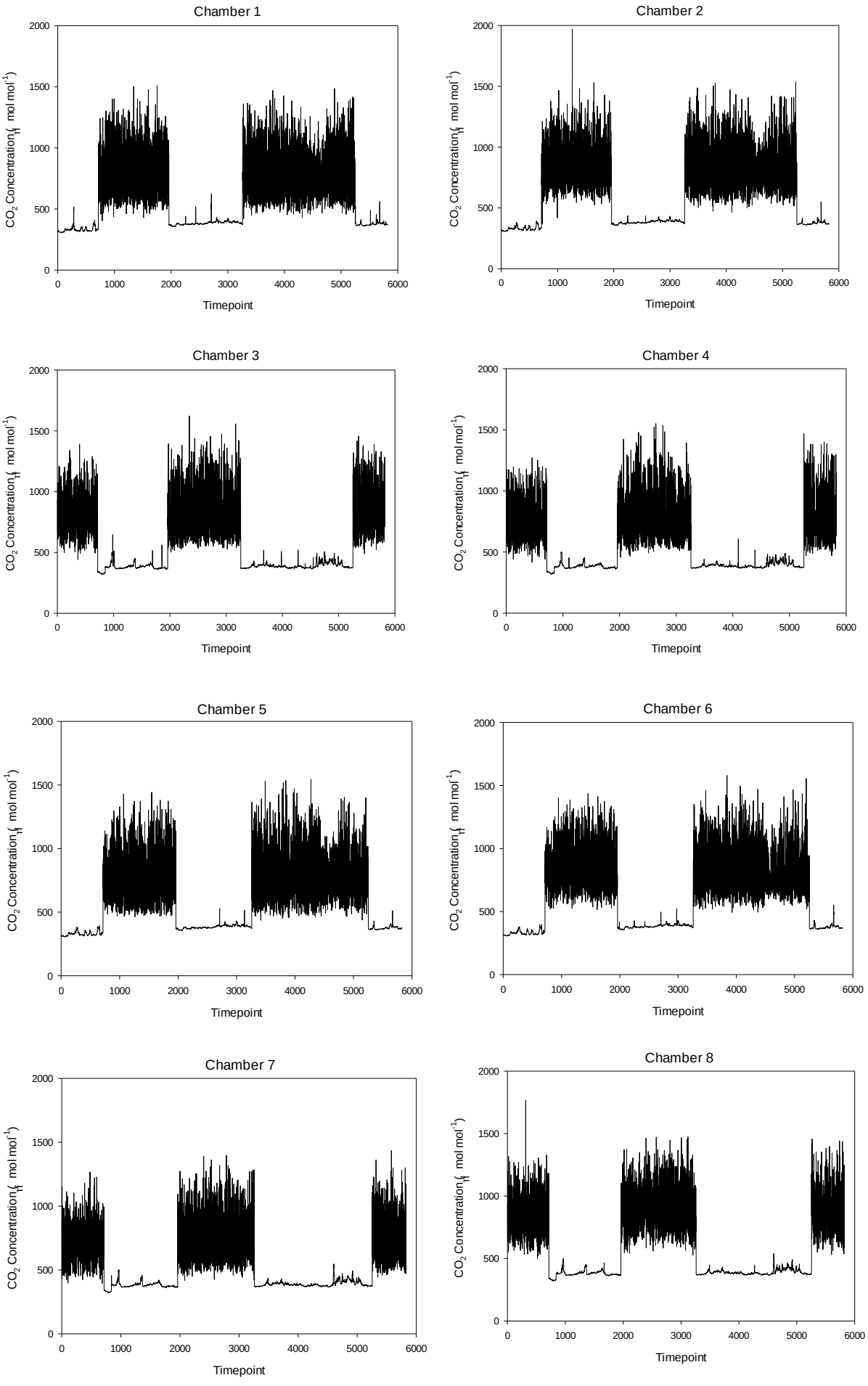



Figure 4.7 – The $CO_2$ concentration (µmol $mol^{-1}$) in each chamber at each time point (one reading every 10 minutes was taken and represents a timepoint).

At 66-67 days after establishment cell imprints were taken from each plant for analysis. The cell imprints were taken from the abaxial surface of the leaves. They were taken by coating the middle of the abaxial leaf surface with nail varnish. Once the nail polish was dried the coated area of the leaf was removed using sellotape. This left an imprint on the sellotape, which was then placed on a glass slide and labelled.

The slides were analysed using a Zeiss upright light microscope. Five 400 µm x 400 µm images of each slide were taken using a x10 magnification lens, making sure each image had a calibration bar. These images were then analysed using a piece of software called ImageJ. Five random 100 µm x 100 µm boxes were then selected from each image and five random epidermal and guard cells were selected from each box and measured for area. The number of epidermal cells and stomata were also counted in each box. The software allows you to trace around the cells or from point to point, and uses the calibration bar to give distances and areas. Averages were taken to give one value for each image. Using these measurements the following calculations were made;

- Stomatal density (See Section 1.6.1)
- Stomatal index (See Section 1.6.2)
- Epidermal cell size
- Epidermal cell number

The plants were then harvested and additional measurements were taken;

- Total Leaf area
- Specific leaf area
- Above ground biomass (Dry)
- Above ground biomass (Fresh)

The statistics were carried using the same model as the previous chamber experiment (See section 2.2.4).

### 4.2.2 AQ and A/Ci Curve Analysis

At 60-61 days after establishment 16 plants were chosen for A/Ci curve analysis. 4 plants from each treatment group were selected. Plants from the same family were selected for both control conditions and both spring conditions to avoid any family bias where possible. To carry out the A/Ci curve a LiCor-6400XT portable photosynthesis system was used. This analysis was done in collaboration with Dr Tracy Lawson and Jack Matthews of the University of Essex. The LiCor-

6400XT was first used to complete a light curve for each plant to determine the light saturation for the A/Ci curve. Conditions were set inside the chamber of the LiCor-6400XT according to the growing conditions. Temperature was set to 22˚C and light intensity to 15 (x10µmol/$m^2$/sec). The LiCor-6400XT contains an automatic programme which produces the light curve so this was run for each plant. The chamber clamps onto a single leaf covering 2$cm^2$, so a young mature leaf was chosen on each plant, and the chamber was clamped around the top of the leaf. After the first two light curves it appeared that the curve plateaued at around 1000 PAR (Photosynthetically Active Radiation) and so this was used as the light saturation level for the rest of the AQ curves as to increase the light too much can damage and stress the plant. AQ curves were conducted at the $CO_2$ concentration the plant was grown under in the chambers, and also under the opposite $CO_2$ concentration to see how the plants would react to the change in $CO_2$ concentration.

As the light saturation had been identified at 1000 PAR from the AQ curves, the A/Ci curves were carried out on each plant using saturating light of 1000 µmol $m^{-2}$ $s^{-1}$. The same temperature was kept in the chamber, but the light level was adjusted to the light saturation measured from the light curve. The A/Ci measurements were taken manually; when lower than 500 µmol $mol^{-1}$ measurements were taken in 50 µmol $mol^{-1}$ increments, then when above 500 µmol $mol^{-1}$ in 100 µmol $mol^{-1}$ increments. Measurements were stopped as soon as the measurements levelled out over several measurements. From the A/Ci curve the Vcmax and Jmax were calculated using the slope and the plateau of the curve. To calculate the slope the first six measurements were used, and to calculate the plateau the highest value was used. For the AQ curve Asat was calculation from the plateau, again taken from the highest point. A two-way ANOVA was then carried out to see if these values were significantly different.

### 4.2.3 Stomatal Response

At 60-61 days after establishment 16 plants were chosen for analysis of stomatal response. Four plants from each treatment group were selected. Plants from the same family were selected for both control conditions and both spring conditions to avoid any family bias where possible (the same plants were used as for the AQ and A/Ci curve analysis where possible). To measure stomatal response a PP systems CIRUS-1 photosynthesis system was used. A young mature leaf was chosen and the chamber clamped on to the leaf covering $2cm^2$. The $CO_2$ concentration in the chamber was set to 400 µmol $mol^{-1}$. The light level was set at low light (146 PAR) and left for 15 minutes. After 15 minutes the intensity was switched to high light (1036 PAR). This was held for 30 minutes to measure the stomatal response to light. After 30 minutes The $CO_2$ in the chamber was increased from 400 µmol $mol^{-1}$ to 800 µmol $mol^{-1}$ and held for another 30 minutes to measure the stomatal response to elevated $CO_2$.

# 4.3 Results

## 4.3.1 Phenotypic Data

The data from the morphological traits of chamber experiment were analysed (Figure 4.8), showing similar results to those found in Chapter 2. An increase in stomatal index in the spring plants in response to elevated $CO_2$ was observed, one of the key traits the experiment was looking to confirm. Statistics were carried out on the morphological traits using the same model as described in section 2.2.4 (Table 4.3). The change in stomatal index was again significant, along with a significant treatment effect in above ground biomass, leaf area, leaf number and guard cell length.

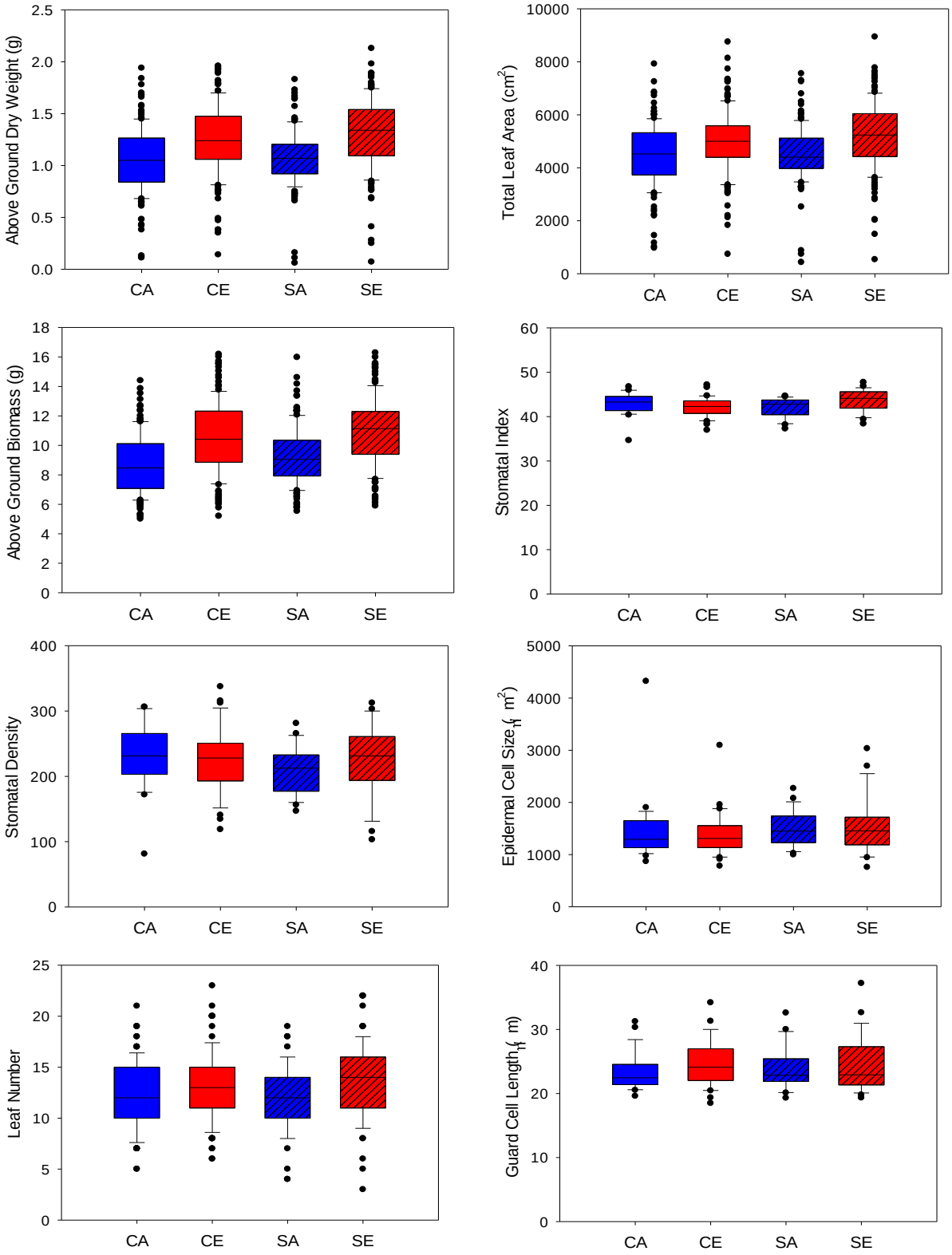



Figure 4.8 – The impact of elevated $CO_2$ on the morphological traits of *P. lanceolata*, originating from either a naturally high $CO_2$ spring (S) or a nearby ambient $CO_2$, control (C), site exposed to either ambient (A) or elevated $CO_2$ (E) at a target concentration of 700 umol $mol^{-1}$ . CA= Control Ambient, CE = Control Elevated, SA = Spring Ambient, SE = Spring Elevated. The central line in each boxplot shows the median. Whiskers indicated the $5^{th}$/$95^{th}$ percentiles. Each dot indicates the observation of an outlier.

| Source | d.f. | Above Ground Dry Weight | | Above Ground Biomass | | Total Leaf Area | | Leaf Number | |
|---|---|---|---|---|---|---|---|---|---|
| | | *T* | *P* | *T* | *P* | *T* | *P* | *T* | *P* |
| **Location** | 1 | 0.726 | 0.394 | 3.455 | 0.063 | 2.610 | 0.107 | 1.252 | 0.263 |
| **Treatment** | 1 | 54.886 | **0.001**** | 74.301 | **0.001**** | 35.62 | **0.001**** | 11.035 | **0.001**** |
| **Location*Treatment** | 1 | 0.096 | 0.757 | 0.025 | 0.874 | 0.670 | 0.414 | 1.216 | 0.270 |
| **Treatment*Family(Location)** | 16 | 21.093 | **0.007**** | 20.514 | **0.009**** | 14.016 | 0.081 | 27.149 | **0.040*** |
| **Location*Chamber (Treatment)** | 11 | 2.994 | 0.810 | 3.774 | 0.707 | 1.857 | 0.603 | 19.520 | **0.012*** |



| Source | d.f. | Epidermal Cell Size | | Guard Cell Length | | Stomatal Index | | Stomatal Density | |
|---|---|---|---|---|---|---|---|---|---|
| | | *T* | *P* | *T* | *P* | *T* | *P* | *T* | *P* |
| **Location** | 1 | 1.403 | 0.236 | 0.469 | 0.493 | 0.161 | 0.688 | 1.266 | 0.260 |
| **Treatment** | 1 | 1.599 | 0.206 | 5.596 | **0.018*** | 1.232 | 0.267 | 0.360 | 0.543 |
| **Location*Treatment** | 1 | 0.621 | 0.431 | 1.870 | 0.171 | 7.874 | **0.005**** | 2.235 | 0.135 |
| **Treatment*Family(Location)** | 16 | 16.295 | 0.433 | 33.403 | **0.007**** | 41.241 | **0.001**** | 17.633 | 0.346 |
| **Location*Chamber (Treatment)** | 11 | 16.485 | 0.124 | 97.423 | **0.001**** | 8.196 | 0.696 | 13.582 | 0.257 |

Table 4.3 – Statistical analysis of morphological data from *P. lanceolata*. A generalized Linear Model was used (see section 2.2.4). P = significance value, T = wald chi-square value. Significance level: * P≤0.05; ** P≤0.01.

## 4.3.2 A/Ci and AQ Curve Analysis

### 4.3.2.1 AQ Curve Analysis

The results of the AQ curve analysis, measured at the $CO_2$ concentrations the plants were grown in (standard), can be seen in Figure 4.9. The results of the AQ curve analysis exposed to the opposite $CO_2$ concentration than they were grown in (flipped) can be seen in Figure 4.10. Averages were taken for each family to give a single curve for each experimental group. Standard deviation and standard error were calculated and added as error bars. The light saturation (Asat) was calculated for each curve (Table 4.4) to see if the curves differed significantly via a two way ANOVA. For the standard AQ curve, the results showed a significant increase in assimilation in elevated $CO_2$, but no significant location effect. The flipped experiment produced no significant differences.

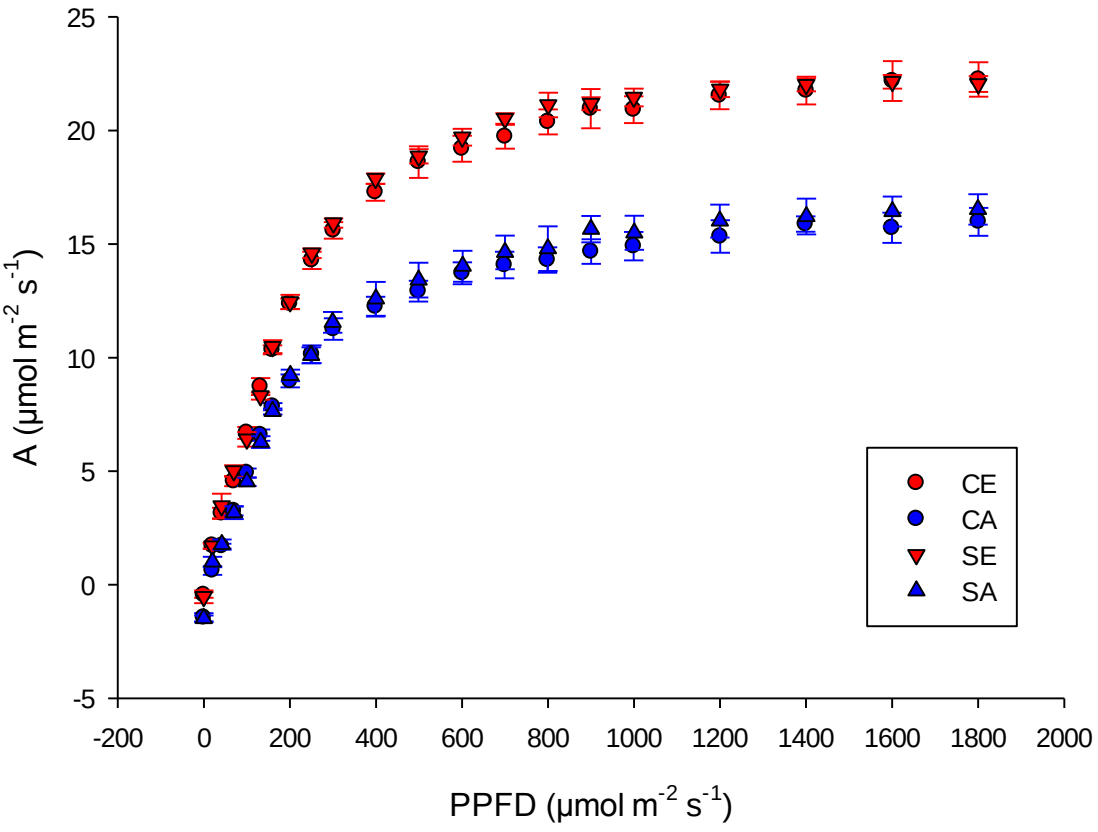



Figure 4.9 – AQ curve analysis carried out at the $CO_2$ concentration the plants were grown in, for each experimental group. S refers to Spring; C refers to Control. E refers to Elevated $CO_2$ concentration; A refers to Ambient $CO_2$ concentration. Error bars are the standard error calculated.

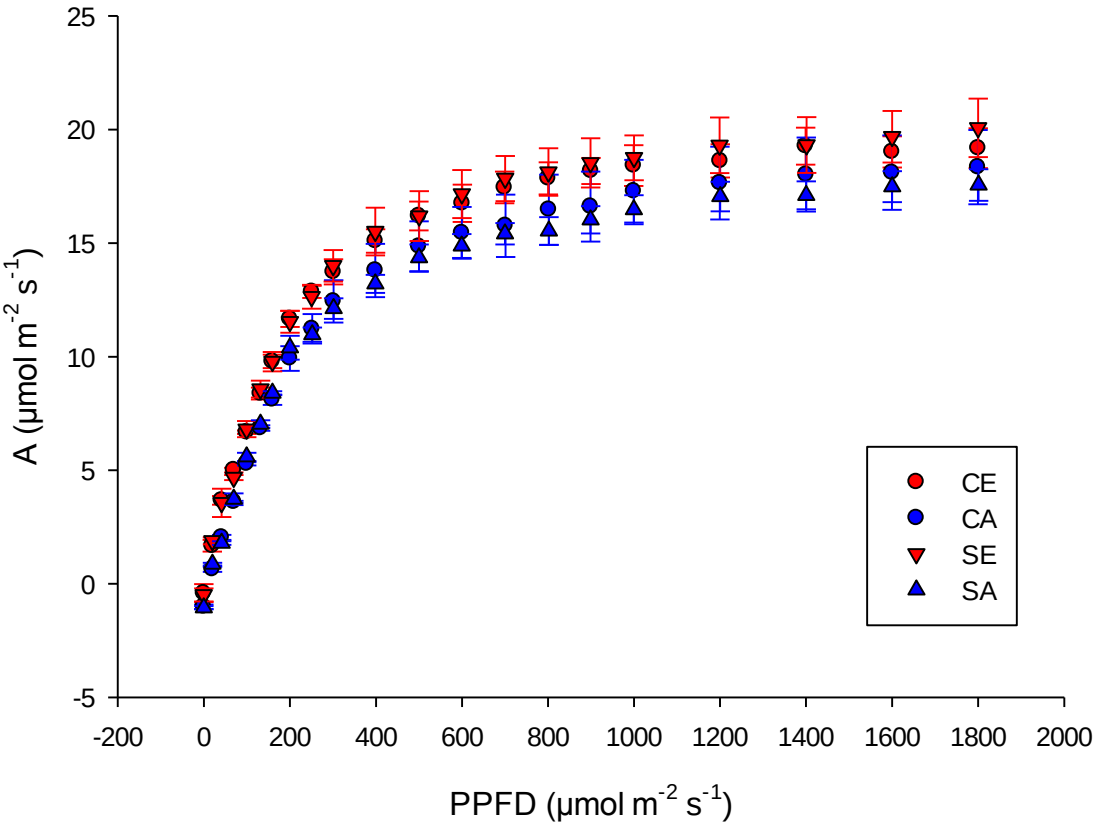



Figure 4.10 – AQ curve analysis carried out at the opposite $CO_2$ concentration the plants were grown under, for each experimental group. S refers to Spring; C refers to Control. E refers to Elevated $CO_2$ concentration; A refers to Ambient $CO_2$ concentration. Error bars are the standard error calculated.

**A**

| | SE | SA | CE | CA |
|---|---|---|---|---|
| *Asat* | 22.247 | 15.983 | 22.049 | 16.526 |

| | Location | Treatment | Location*Tretament |
|---|---|---|---|
| *Asat* | 0.745 | **0.001**** | 0.487 |

**B**

| | SE | SA | CE | CA |
|---|---|---|---|---|
| *Asat* | 17.561 | 20.078 | 18.350 | 19.181 |

| | Location | Treatment | Location*Tretament |
|---|---|---|---|
| *Asat* | 0.958 | 0.102 | 0.410 |

Table 4.4 – The average for each experimental group, and the results of the two-way ANOVA carried out on Asat for (**A**) the experiment carried out at the concentrations the plants were grown in, and (**B**) the flipped $CO_2$ experiment. Significance level: * P≤0.05; ** P≤0.01

### 4.3.2.2 A/Ci Curve Analysis

The A/Ci curve carried out can be seen in Figure 4.11. Again an average was taken for each family to give a single cure for each experimental group, and the standard deviation error calculated and added as error bars. The Vcmax and Jmax were calculated for each sample, and then a two-way ANOVA was carried out to see if the differences were significant (Table 4.5). A significant location effect can be seen, but no apparent $CO_2$ treatment effect.

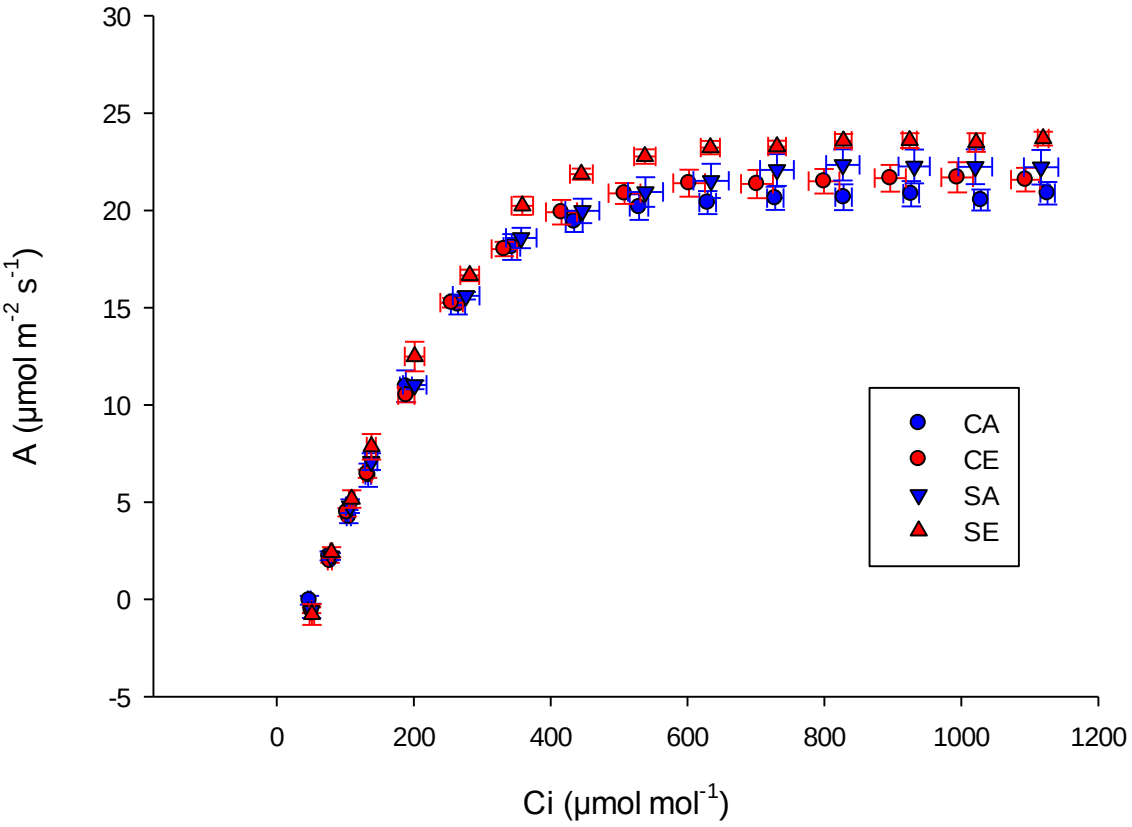



Figure 4.11 – A/Ci curve analysis carried out under saturating light (1000 µmol $m^{-2}$ $s^{-1}$) on each experimental group. S refers to Spring; C refers to Control. E refers to Elevated $CO_2$ concentration; A refers to Ambient $CO_2$ concentration. Error bars are the standard error calculated.

| | SE | SA | CE | CA |
|---|---|---|---|---|
| *Vcmax* | 71.01822 | 65.29397 | 61.7557 | 54.1642 |
| *Jmax* | 139.2941 | 129.0646 | 126.6056 | 117.6742 |

| | Location | Treatment | Location*Tretament |
|---|---|---|---|
| *Vcmax* | **0.031*** | 0.136 | 0.826 |
| *Jmax* | **0.03*** | 0.074 | 0.897 |

Table 4.5 – The average Vcmax and Jmax for each experimental group, and the results of the two-way ANOVA carried out on the Vcmax and Jmax values. Significance level: * P≤0.05; ** P≤0.01.

### 3.3.3 Stomatal Response

Several measurements were taken to analyse the stomatal response of *P. lanceolata*. The stomatal conductance was measured to analyse the amount of water transfer between the leaves and the environment (Figure 4.12). Averages were taken for each family to give a single curve for each experimental group. Standard deviation and standard error were calculated and added as error bars. The light was increased from 146 PPFD to 1036 PPFD at 14 minutes, and then $CO_2$ was increased from 400ppm to 800ppm at 45 minutes.

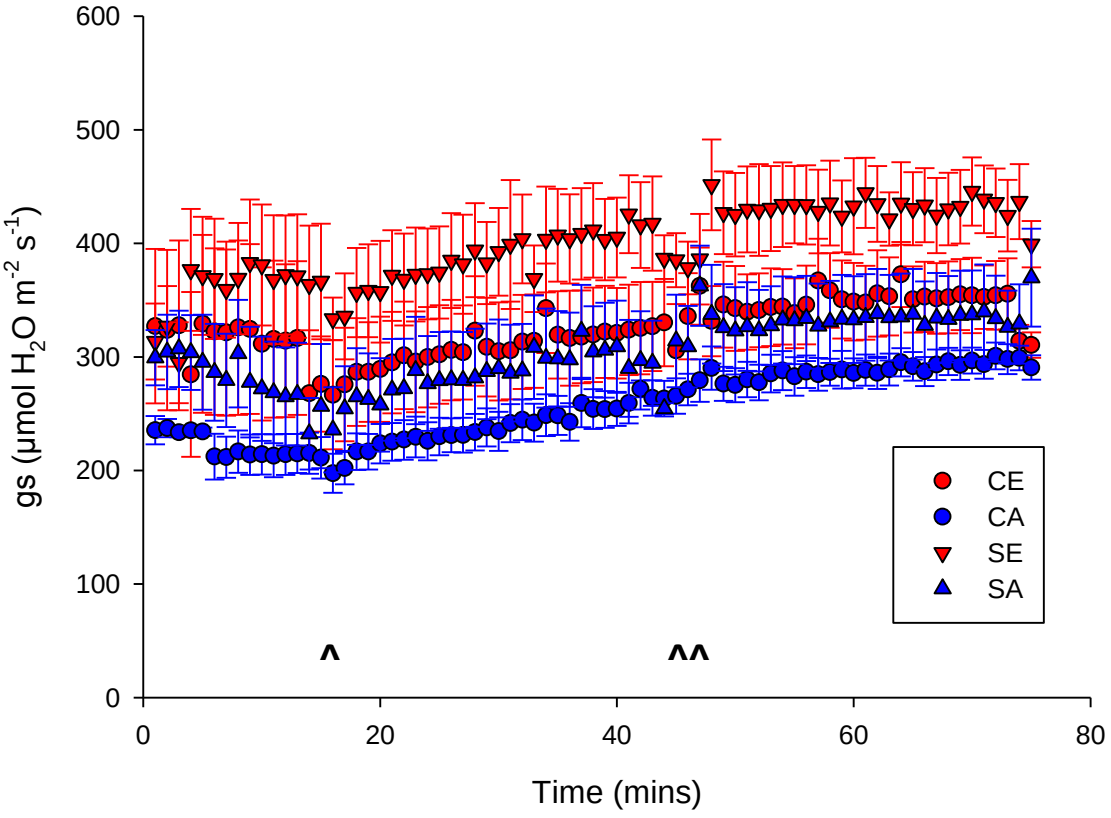



Figure 4.12 – The stomatal conductance (gs) results for each experimental group. Measurements were taken every minute for 75 minutes. Measurements were carried out at 146 PPFD, and then raised to 1036 PPFD at 14 minutes (^). Also carried out at 400 ppm, and the raised to 800 ppm at 45 minutes (^^). S refers to Spring; C refers to Control. The number is family used. E refers to Elevated $CO_2$ concentration; A refers to Ambient $CO_2$ concentration.

The assimilation (A; Figure 4.13) and intracellular $CO_2$ concentration (Ci; Figure 4.14) were both measured. The ratio between the assimilation rate and stomatal conductance was then calculated to give the water use efficiency (iWUE; Figure 4.15). Again, averages were taken for each family to give a single curve for each experimental group. Standard deviation and standard error were calculated and added as error bars. The light was increased from 146 PPFD to 1036 PPFD at 14 minutes, and then $CO_2$ was increased from 400ppm to 800ppm at 45 minutes.

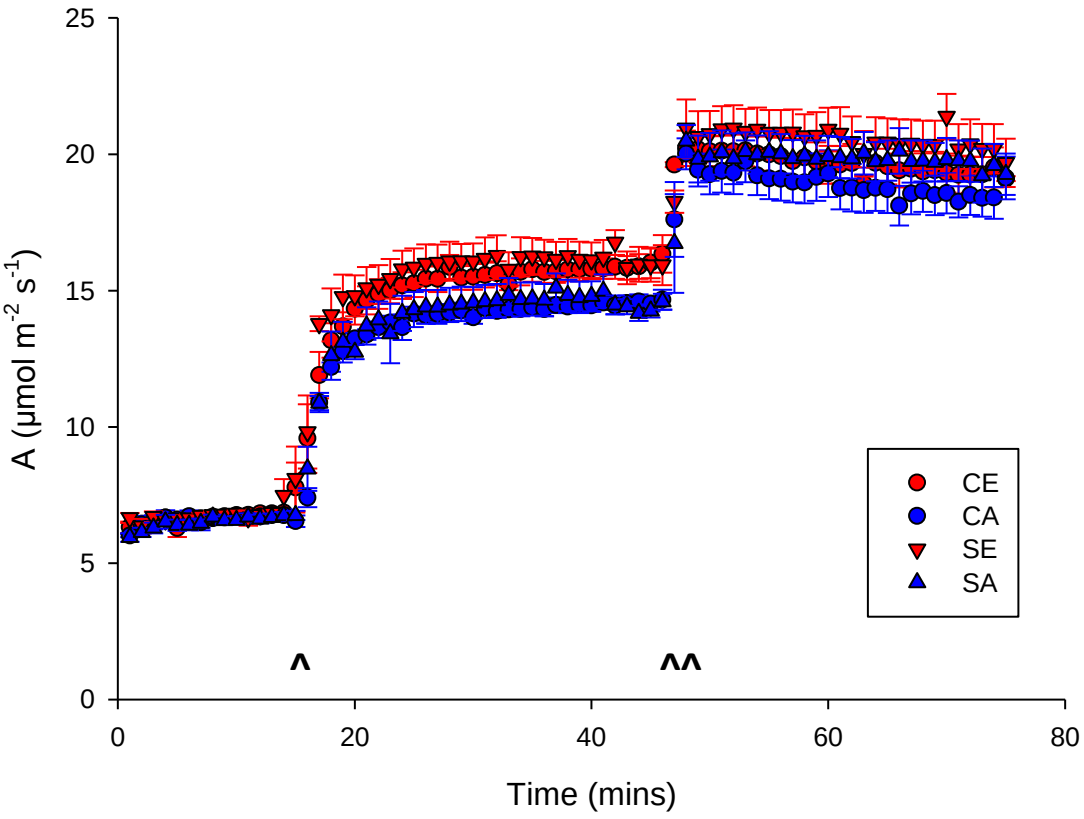



Figure 4.13 – The assimilation (A) results for each experimental group. Measurements were taken every minute for 75 minutes. Measurements were carried out at 146 PPFD, and then raised to 1036 PPFD at 14 minutes (^). Also carried out at 400 ppm, and the raised to 800 ppm at 45 minutes (^^). S refers to Spring; C refers to Control. The number is family used. E refers to Elevated $CO_2$ concentration; A refers to Ambient $CO_2$ concentration.

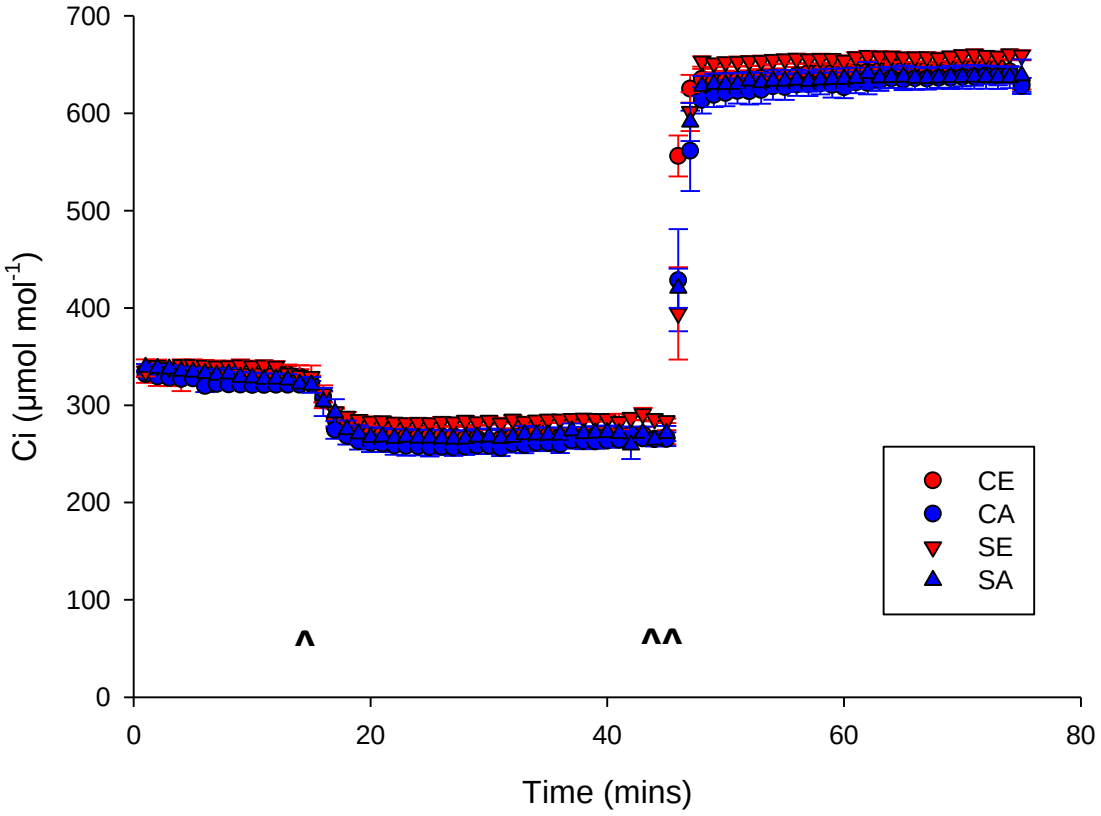



Figure 4.14 – The intra cellular $CO_2$ concentration (Ci) results for each experimental group. Measurements were taken every minute for 75 minutes. Measurements were carried out at 146 PPFD, and then raised to 1036 PPFD at 14 minutes (^). Also carried out at 400 ppm, and the raised to 800 ppm at 45 minutes (^^). S refers to Spring; C refers to Control. The number is family used. E refers to Elevated $CO_2$ concentration; A refers to Ambient $CO_2$ concentration.

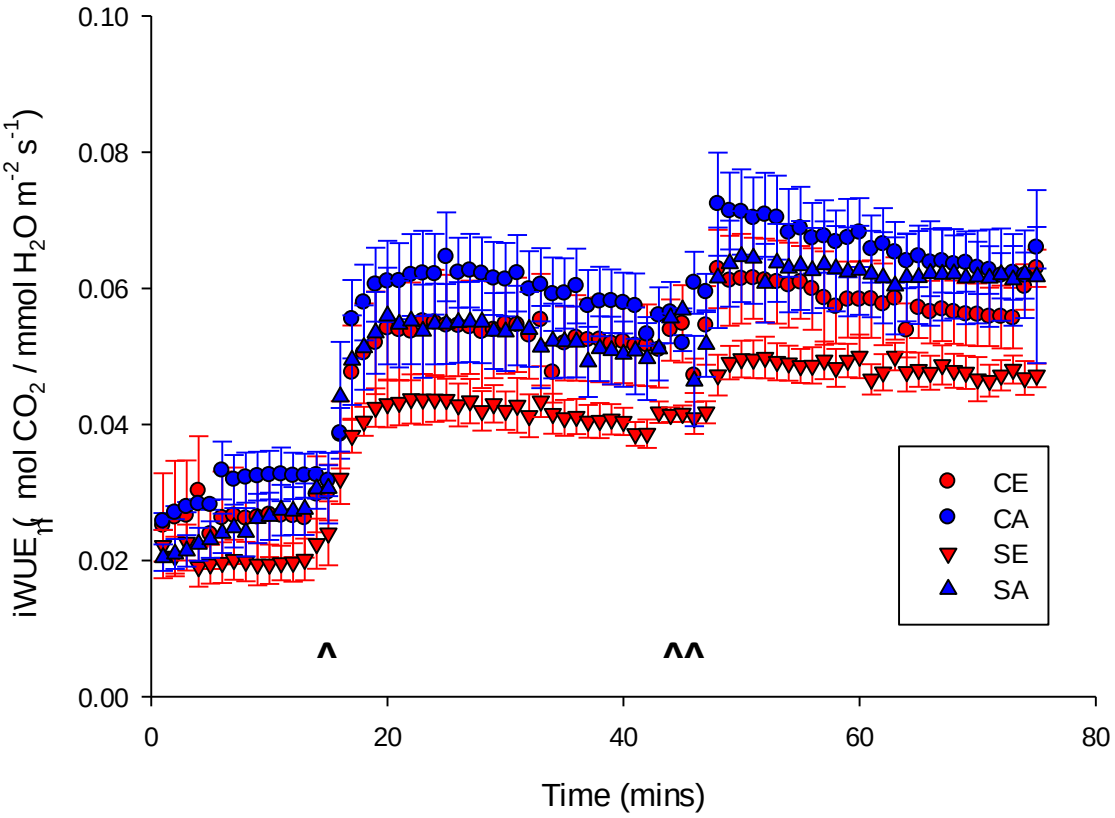



Figure 4.15 – The water use efficiency (iWUE) results for each experimental group. Measurements were taken every minute for 75 minutes. Measurements were carried out at 146 PPFD, and then raised to 1036 PPFD at 14 minutes (^). Also carried out at 400 ppm, and the raised to 800 ppm at 45 minutes (^^). S refers to Spring; C refers to Control. The number is family used. E refers to Elevated $CO_2$ concentration; A refers to Ambient $CO_2$ concentration.

## 4.4 Discussion

This extension of the findings of the Chapter 2 allowed for both confirmation of the phenotypes identified, and analysis of the physiological mechanisms underpinning the observed phenotype of the *P. lanceolata*. The phenotypic measurement showed a significant increase in stomatal index in the spring plants (SE/SA = 3.93% increase ($P < 0.05$, Tukey's HSD)) in response to elevated $CO_2$, confirming the stomatal phenotype found in Chapter 2. This was vital as any investigations into physiological mechanisms would be non-comparable to the first experiment if the phenotype was not consistent and present. A significant increase in above ground biomass was noted (SE/SA = 17.84% increase, and CE/CA = 18.45% increase ($P < 0.01$)) in this experiment for both spring and control plants in response to elevated $CO_2$, but no significant difference between spring and control plants was apparent. No significant differences in above ground biomass were noted in Chapter 2. The increase in above ground biomass in control plants in response to elevated $CO_2$ is expected (Klus, Kalisz *et al.* 2001), but the increase in the spring plants does not follow what was previously identified by Woodward (1999). Total leaf area significantly increased in response to elevated $CO_2$ in both spring and control plants (SE/SA = 15.49% increase and CE/CA = 11.98% increase ($P < 0.01$)). The number of leaves also significantly increased in the both populations in response to elevated $CO_2$ ($P < 0.01$). This would suggest the larger total leaf area in response to elevated $CO_2$, is due to more leaves. A decrease in leaf size in spring plants, compared to control plants in response to elevated $CO_2$ was seen in Chapter 2, but was not apparent here. The literature shows leaf area can be variable in response to elevated $CO_2$ (Ferris and Taylor 1994, Körner, Asshoff *et al.* 2005), but the only response previously shown for *P. lanceolata* originating from the Bossoleto spring is decreased leaf size under elevated $CO_2$, compared to ambient $CO_2$ (Woodward 1999). There was a difference in harvest time between the two experiments, with the plants in Chapter 2 being almost twice the age (in Chapter 2 plants harvested at 123 days; in this experiment plants were harvested at 66 days). As plants age, the root : shoot ratio has been shown to be affected, and they relocate more biomass to the roots (Shipley and Meziane 2002). Woodward (1999) showed the root size increased in spring plants in response to elevated $CO_2$ compared to control plants. Here as young plant are being analysed, they may still be placing more resources into the leaves, and not yet the roots, which is why no decrease in leaf size is seen, and may also explain the increase in above ground biomass in the spring plants.

A/Ci analysis gives a good indication of both RuBisCO efficiency and maximum capacity (Farquhar, Caemmerer *et al.* 1980). A significant location effect ($P < 0.05$) was found between the A/Ci curve for both Vcmax and Jmax, suggesting the spring plants have both a higher efficiency and maximum photosynthetic capacity. Spring plants had consistently higher Vcmax (SA/CA = 17.05%

increase and SE/CE = 15% increase) and Jmax (SA/CA = 7.05% increase and SE/CE = 7.93% increase) values across both treatments, which could be indicative of an adaptation to elevated $CO_2$ in the spring. Interestingly neither the spring or control plants showed the previously documented, most common response to elevated $CO_2$ (Long, Ainsworth *et al.* 2004), both increasing Vcmax and Jmax under elevated $CO_2$. Further to this, the results do not support the gene expression analysis of RuBisCO in Chapter 3, where a down regulation was found for CE/CA, but no effect for SE/SA. Although such an increase has previously been noted (Rey and Jarvis 1998), this is the first time that this increase has occurred with plants from an elevated $CO_2$ population. These results do not agree with the results previously found in the spring, where Bettarini *et al.* (1999) showed assimilation significantly was decreased in the spring plants at both ambient (-33%) and elevated (-31%) $CO_2$ concentrations and require further investigation beyond the scope of this study. One possible explanation is that the youngest fully grown leaf used here differed to that previously used, since Bettarini *et al.* (1999) used mature leaves during their experiment. It has been shown that leaf age can negatively affect photosynthetic rate (Constable and Rawson 1980), and so could explain why an increase is noted here and a decrease is found in mature leaves. This could have an effect on the source/sink balance of the leaves which could be affecting the efficiency of the A/Ci curve (Paul and Foyer 2001, Kaschuk, Hungria *et al.* 2010), but this cannot explain the inconsistencies between our gene expression, and photosynthetic data observed within this study.

The AQ curve analysis produced only a significant treatment effect ($P < 0.01$), and both populations followed a very similar response to elevated $CO_2$. Asat increased by 33.42% in control plants, and 39.19% in spring plants in response to elevated $CO_2$. An increase is Asat under elevated $CO_2$ is well documented, with a meta-study showing the average increase was 31% in response to elevated $CO_2$ (Long, Ainsworth *et al.* 2004). The fact that there was no difference between spring and control is probably due to the fact that light intensities are very similar at the two sights, due to the closeness of the sites and similar canopy cover. Neither sight would need to adapt to changing light conditions, and so harvest light at similar rates.

Analysis of the stomatal response data produced some interesting results, including the spring plants showing lower water use efficiency (iWUE) than the control plants, again, a result clearly in contrast to those reported previously and summarised in several meta-analyses (Eamus 1991, Long, Ainsworth *et al.* 2004). However, here, the decrease in iWUE is coupled with an increased stomatal conductance in the spring plants, which could be working in tandem with the increased RuBisCO efficiency and maximum capacity, as the increase in stomatal conductance correlates with an increase in assimilation in the spring plants. This increase is in the spring plants is

independent of $CO_2$ concentration, and occurs in both ambient and elevated $CO_2$ conditions. An increase in stomatal conductance was previously shown in spring plants, again by Bettarini *et al.* (1999). The response to elevated light and $CO_2$ was very similar in both control and spring plants, and no discernable differences can be noted, so the only differences in the stomatal response data is between the populations in respect to consistently higher (for stomatal conductance and assimilation), or lower (for iWUE) values at high and low light or $CO_2$.

These results paint an interesting picture, in an attempt to link underlying gene expression mechanisms to those of physiological processes. It is apparent that inside the spring, the efficiency of the physiological mechanisms associated with photosynthesis has increased. The increased Vcmax and Jmax of the A/Ci curve correlates with an increased stomatal conductance in the spring plants, which suggests in the spring plants, flux of both carbon dioxide and water vapour increases through stomatal pores. This may reflect the fact that both carbon dioxide and water are in plentiful supply in the spring, and are utilised at an entirely different steady state-an unlimited state. This correlates with the increased number of stomata in the spring plants, as more stomata allows for more portals for $CO_2$ entry and therefore increased stomatal conductance. The consequence of this is lower instantaneous iWUE for spring plants. From the previous chapter it was noted the gene expression of the photosynthesis pathway, differed in response to elevated $CO_2$. The control plants showed decreased gene expression across the pathway in response to elevated $CO_2$, as might be predicted from well-worn literature on RuBisCO and elevated $CO_2$, but here, for gas exchange there was no treatment effect on Vcmax or Jmax. This would suggest the control plants are able to use the excess $CO_2$ to photosynthesise more efficiently, and so decrease their gene expression, but this could not be detected in the gas exchange measurements reported here. On the other hand spring plants did not decrease their gene expression for RuBisCO in response to elevated $CO_2$, and in some cases increase gene expression across the photosynthesis pathway was observed. Clearly it is an over simplification to attempt to link gene expression data in this way to physiological processes since we have no information at the protein level, of RuBisCO investment that could explain the results observed here. What is undisputable, is that photosynthetic acclimation to elevated $CO_2$ as understood from the numerous reports in the literature, may not be occurring in plants grown within the spring environment, and that significant differences at the level of gene expression exist between spring and control plants that warrant further investigation.

# DNA Polymorphism and Evidence for Selection

## 5.1 Introduction

Differential gene expression can be related to phenotypic evolution (Shapiro, Marks *et al.* 2004), but coding sequence (exons) can also correspond to any phenotypic differences due to allelic differences (Gardiner, Guethlein *et al.* 2001). Variation can also occur in the non-coding region of DNA (introns) which can affect the coding regions and expression of associated genes (Laurie and Stam 1994). It is therefore important to sequence the genomic DNA of genes of interest as to not miss any indicators of selection, as the RNA-Seq only provides the sequence of coding DNA. Also as we are interested in differentially expressed genes there can be varied levels of expression, and often lowly expressed genes have missing sections of sequence. This does not mean the RNA-Seq data is not useful, as it provides the coding sequence, which is likely to also be in linkage disequilibrium with upstream regions and can pick up any evidence of increased divergence in genes where a DNA sequence is not available. Analysing both coding sequence and upstream regions is important as it can be used to indicate if any selective sweeps have occurred which could suggest selection to the spring environment, and if one set of plants is derived from the other. Selective sweeps leave what is known as DNA footprints, where characteristics such as less variation are observed (Oleksyk, Smith *et al.* 2010). Population genetics has long been used to identify these characteristics, and a range of metrics have been developed over the years to provide measurements of aspects related to genetic variation. These are inevitably based on the number of polymorphic sites between sequences. There are two main measures of polymorphism, the first is known as nucleotide diversity, or Pi ($\pi$) (Formula 5.1), which is used as a measure of the average diversity between sequences. This was first introduced by Nei and Li (1979) and takes the average number of nucleotide differences per site between any two DNA sequences chosen randomly from the sample population. This takes into account the extent of differentiation between sequences, as well as the relative frequencies of the sequence. This is also sometimes referred to as the pairwise method, as pairwise samples are randomly chosen before the average is taken. The second is based on the number of segregating sites or, theta ($\theta$; Formula 5.2). $\theta$ is similar to $\pi$, but measures the total diversity by assuming what is known as the infinite sites model (Kimura 1969). This model calculates the number of segregating sites (DNA sites that are polymorphic) and assumes that the number of nucleotide sites is large enough that each new mutation occurs at a site that has not mutated before so mutations can never overlay or reverse each other.

$$\pi = \frac{n}{n-1} \sum_{ij} x_i x_j \pi_{ij}$$

Formula 5.1 – The equation used to define π where $x_i$ and $x_j$ are the respective frequencies of the ith and jth sequences, $\pi_{ij}$ is the number of nucleotide differences per nucleotide site between the ith and jth sequences, and n is the number of sequences in the sample(Nei and Li 1979).

$$\ominus = \frac{S}{\sum_{i=1}^{n-1} \frac{1}{i}}$$

Formula 5.2 – The equation used to define θ where s is the number of segregating sites, n is the sample size and i is the index of summation (Kimura 1969).

These two measures are often used in conjunction to give a comprehensive measure of polymorphism, as although they are expected to be equal according to the neutral theory of molecular evolution (see below), rare alleles can influence θ counts as all segregating sites are counted equally, an effect which does not affect π due to the way it incorporates heterozygosity. For both metrics, which vary between 0 and 1, a value of 0 would indicate no polymorphism, with the value increasing as polymorphism increases. Lower values of polymorphism in a population compared to a control would be indicative of a selective sweep occurring in that population, as variation is lost when selecting towards an environmental factor (Clark, Schweikert *et al.* 2007).

The neutral theory of evolution was proposed by Kimura (1984), to explain evolution at a molecular level. It suggests that most evolutionary changes, along with most of the variation within and between species, are not caused by natural selection, but by random drift of alleles that are neutral. A neutral mutation is one that does not affect an organism's ability to survive and reproduce. The neutral theory allows for the possibility that most mutations are deleterious, but because these are rapidly purged by natural selection, they do not make significant contributions to genetic variation within and between species. Mutations that are not deleterious

are assumed to be mostly neutral rather than beneficial (Kimura 1984). In addition to assuming the primacy of neutral mutations, the theory also assumes that the fate of neutral mutations is determined by the sampling processes described by specific models of random genetic drift. The strength of genetic drift depends on the population size. If a population is at a constant size with a constant mutation rate, the population will eventually reach equilibrium of gene frequencies.

A measure that takes into account the level of polymorphism and the neutral theory of evolution is known as Tajima's D (Tajima 1989). Tajima's D is a measurement which distinguishes between a DNA sequence evolving randomly (neutral theory of evolution), and one evolving under selection, a non-random process. These selections include directional or balancing of selection, demographic expansion or contraction, genetic hitchhiking, or introgression. Tajima's statistic computes a standardized measure of θ in the sampled DNA and π. If these two numbers differ by as much as one could reasonably expect by chance, then the null hypothesis of neutrality cannot be rejected. Otherwise, the null hypothesis of neutrality is rejected and so selection is suggested. The way Tajima calculates this difference has led to its popularity in being able to give a value across species (Formula 5.3 and Formula 5.4). By considering the variance of π and θ of the two samples it expresses the difference between the two estimates relative to their standard error.

$$D = \frac{\pi - \theta}{\sqrt{V}}$$

Equation 5.3 – The formula to define Tajima's D where π is the average number of pairwise mutations, θ is the number of segregating sites and V is calculated using Equation 5.4 (Tajima 1989).

$$V = Var[\pi - \theta]$$

Equation 5.4 - The formula to calculate V where π is the average number of pairwise mutations, θ is the number of segregating sites and Var is the variance.

If the difference between π and θ was normally distributed, then we would expect D to lie between –2 and 2 about 95% of the time. This can be used in relation to the amount of selection;

values that deviate further from 0, and are closer too or over 2 and -2 would indicate more selection is present. Which way the value trends can also be used to gain insight into what is happening in the populations. A negative Tajima's D infers a population expansion after a recent bottleneck, or a selective sweep. A bottle neck refers to a subsampling of the original population; a sub-population originating from a parent population, but a selective sweep refers to selection towards environmental stimuli in specific alleles. Both can occur simultaneously, as the sub-population can undergo selection not present in the parent population. In contrast, a positive Tajima's D is likely to be due to a sudden population contraction where balancing of selection is occurring. Using + or - 2 is an indication to whether a gene has a significant Tajima's D value, but it is not a robust statistical test, and so can only be used as a guide. As there is no true statistical test, the values which segregate the most from the data set (the highest and lowest 5%) are often reported as significant (Thomas, Godfrey *et al*. 2012), as the test is unlikely to identify false positives in interesting regions of a chromosome, if only the greatest outliers are reported.

Another important population statistic to consider is the fixation index (FST), which gives a measure of population differentiation based on genetic structure. It was originally developed from Wright's F-Statistic (Wright 1965), which describes the expected level of heterozygosity in a population, and is now one of the most commonly used statistics in population genetics. In its basic form FST is an estimator of the differences in π between and within populations (Formula 5.5). In this form there are biases in variables, such as sample size, and so the equation has been adapted to be more complex and comprehensive (Weir and Cockerham 1984), but the basic principles are still present.

$$FST = \frac{\pi Between - \pi Within}{\pi Between}$$

Formula 5.5 – The formula used to define FST where π is the average number of pairwise differences between different sub-populations (Between) and within the same sub-population (Within) (Hudson, Slatkin *et al.* 1992)

The output of FST is a value between 0 and 1, where 0 implies the two populations are interbreeding freely, and a value of 1 implies that all of the genetic variation is explained by population structure and the populations do not share genetic variation. Values of FST below 0 indicate spatially uniform selection is occurring that favours the same phenotype in different environments (De Kort, Vandepitte *et al.* 2013). There is no statistical test to evaluate the

differences between FST, but the top 5% of FST values are often considered as significance (Gardner, Williamson *et al.* 2007, Narum and Hess 2011).

Dxy (Formula 5.6) is a measure which gives a similar estimation to FST, and looks at the pairwise differences between two randomly chosen sequences within two populations (Takahata and Nei 1985). Dxy quantifies the extent of divergence between populations (Wakeley 1996), but a lack of knowledge about the variation in Dxy and its components under particular genetic and demographic models has both discouraged their use and made their interpretations more difficult (Simmons, Kreitman *et al.* 1989). FST therefore remains the main indicator for aspects of divergence between populations.

$$Dxy = \frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} k_{ij}$$

Formula 5.6 – The equation used to define Dxy, where i and j denote sequences from populations x and y respectively (Takahata and Nei 1985).

All these statistics are often used because together they give a comprehensive overview of the nucleotide differentiation that is evident between sequences and populations (Chapman, Hiscock *et al.* 2013).

### 5.1.1 DNA Sequencing

These metrics can be applied to the RNA-Seq data obtained, but these are in most cases only partial sequences, and most variation can occur in the non-coding sequence and thus, they are limited (Villablanca, Roderick *et al.* 1998). DNA sequencing can be used to obtain the full sequence of the genes of interest to minimalize missing any variation in the introns and missing exon data of the DNA. Sanger sequencing (Sanger and Coulson 1975) was one of the original methods of DNA sequencing, but has now been superseded by NGS. Sanger sequencing is still widely used in cases where only small numbers of samples need to be analysed and long contiguous reads of DNA are required. Sanger sequencing can produce DNA reads >800 bases, so a forward and reverse reaction can be used to produce >1500 base sequences. Sanger sequencing is based on the selective incorporation of chain-terminating di-deoxynucleotides

(ddNTPs), by DNA polymerase during in vitro DNA replication. Each base terminator contains a different coloured automatic dye. The fragments are passed through a capillary to separate them, the dyes then indicate which base is present, and the amount of migration indicates the position of the base, together forming the DNA sequence.

### 5.1.2 Aims

-To investigate the genetic diversity of the coding regions of the genome for the genes of interest via the RNA-Seq data set.

- Use DNA sequencing to obtain full genomic sequences (where possible) of genes of interest to further investigate the genetic diversity of these genes in terms of the RNA-Seq data set.

## 5.2 Methods

DNA was extracted from 24 samples (12 control and 12 spring) from the chamber experiment conducted in Chapter 3, 55 days after establishment, using a CTAB-based protocol. Samples were frozen in liquid nitrogen, ground into a powder, and placed into individual eppendorfs. 900 µl of pre-warmed (65 ºC) CTAB ( 2% hexadecyltrimethylammonium bromide (weight (w)/v), 2% polyvinylpyrrolidone (w/v), 100 mM Tris-HCl (hydrochloric acid), 25 mM EDTA Ethylenediaminetetraacetic acid and 2 mole (M) NaCl (sodium chloride)) and 2 µl β-nercaptoethanol was added to each Eppendorf tube, and tubes were incubated at 65 ºC for 45 minutes. After incubation, 900 µl CHISAM (Chloroform/Isoamyl alcohol (ratio of 24:1 respectively)) was added into each tube and tubes were centrifuged for 10 minutes at 12,000 x gravity (g). The upper layer (aqueous phase) was transferred to a new tube and 50 µl of 3M sodium acetate and 333 µl of cold Isopropanol was added. After incubation at -20 ºC for 60 minutes, the samples were centrifuged for 15 minutes at 12,000 x g at 4 ºC to form a pellet. The liquid phase was discarded and 50 µl of TE buffer and 1 µl of RNAase was added to each Eppendorf and left over night. 500 µl of cold 70% ethanol was then added and centrifuged for 2 minutes at 12,000 x g at 4 ºC. The liquid phase was discarded then 200 µl of TE solution, 100 µl of 3M sodium acetate and 1ml of cold absolute Ethanol was added and centrifuged for 10 minutes at 12,000 x g at 4 ºC. The liquid phase was again discarded and again 500 µl of cold 70% ethanol was then added and centrifuged for 2 minutes at 12,000 x g at 4 ºC. The liquid phase was discarded and the pellet was left to dry for 20 minutes. Once dry the pellet was re-suspended in 50 µl of DEPC treated water. The DNA concentration was measured using the Nanodrop spectrophotometer (ND100, NanoDrop Technologes, Delaware, USA).

Primers were designed for 12 sequences of interest (Table 5.1) chosen for their involvement in either the stomatal patterning or photosynthesis pathway using Primer3 (Untergasser, Cutcutache *et al.* 2012). The primers were designed to encapsulate the parts of the sequence with the most homology to the genes of interest. The primers were between 18 and 23 bases, have a primer Tm between 58 ºC and 62 ºC and a GC content between 40% and 60%. They were also designed to give a product of around 800 bases, as this size is optimum for Sanger sequencing. Polymerase Chain Reaction (PCR) was then used to amplify each sample with each gene of interest using the program:

- (Denature) 95 ºC for 10 minutes
- 35 cycles of;
    - (Denature) 95 ºC for 30 seconds
    - (Anneal) 50 ºC for 30 seconds
    - (Elongation) 72 ºC for 1 minute
- (Elongation) 72 ºC for 10 minutes
- 4 ºC for 10 minutes

| Gene | AT Number | Primers | |
|---|---|---|---|
| ERECTA | AT2G26330 | *Forward* | AGGGGATTGTTCTGCACTCA |
| | | *Reverse* | GGAGCTTTGTCATGTTCCCC |
| YODA | AT1G63700 | *Forward* | GGGTTTCCTGGATCTGGTCA |
| | | *Reverse* | TTGTCCCAACTGCTTCACAC |
| MYB88 | AT2G02820 | *Forward* | ATTGGGCTGGGAATTGATGC |
| | | *Reverse* | TGCCAATCAGCTTCATGTGG |
| CDKB1;1 | AT3G54180 | *Forward* | CAATACGAGCCATGTCCTGC |
| | | *Reverse* | GAGAAAGTCGGCGAAGGAAC |
| SCREAM2 | AT1G12860 | *Forward* | TGATTTGGGTTGCCATGGTG |
| | | *Reverse* | GAGCTACTCGGGATGGACTC |
| Epidermal Patterning Factor 1(EPF1) | AT2G20875 | *Forward* | CTGTTTATGCATGCAAGGCG |
| | | *Reverse* | TGGCAAAATCTGAGGTCGAC |
| Beta Carbonic Anhydrase 1 (BCA1) | AT3G01500 | *Forward* | GGAAATGGGGAGGAGGATCA |
| | | *Reverse* | ACCCTGCTTTGTATGGTGAAC |
| RuBisCO Small Subunit (RUBS) | AT1G67090 | *Forward* | TAGTAGTCGTCTGGCTTGGC |
| | | *Reverse* | GGCACTCCTCTCCATTCTGT |
| RuBisCO Interacting (RUBINTER) | AT2G39730 | *Forward* | AGGCATTCGGGTCAGTACAA |
| | | *Reverse* | ATCAGGCAAAGGTACCGTGA |
| Photosystem I.LHC-I (LHC_I16) | AT1G61520 | *Forward* | CTGCCAAATGATCCAACAGGT |
| | | *Reverse* | TGCAGTCAGGAAATAGGCAAC |
| Photosystem II.LHC-II (LHC_II299) | AT1G29910 | *Forward* | CCACTCTAGATGGGCAATGC |
| | | *Reverse* | GCAAGCCCAGATAGCCAAAA |
| Rib5P Isomerase (RIB5P) | AT3G04790 | *Forward* | TGGCCGCCACATTATCATTC |
| | | *Reverse* | AACAAACCATGCTCCACCAC |

Table 5.1 – The list of genes and their respective AT numbers chosen for DNA sequencing.

Sanger sequencing was chosen as the sequencing method, which required preparation post PCR to remove contamination, such as random DNA oligos to prevent the sequencing reaction proceeding in both directions, due to the presence of both forward and reverse primers. This was accomplished by adding 0.5 µl of EXO1 exonuclease, 1 µl of Alkaline Phosphate and 1.5 µl of DEPC treated water to 15 µl of each PCR product. This was then run in a PCR machine using the following programme:

- 37 ºC for 45 minutes
- 80 ºC for 15 minutes
- 4 ºC for 5 minutes

The samples were then sent in 96 well plates to SourceBioscience (Nottingham, UK) where a sequencing PCR was run before the Sanger sequencing was carried out. Forward and reverse Sanger sequencing was carried out on each sample for each gene of interest.

### 5.2.1 Population Genetics

The sequences were imported into BioEdit (Hall 1999) to trim the ends of the sequences where the quality was poor. Where clear distinguishing peaks with no bridging to the next peak could not be seen the bases were removed. When no sequence was useable, the sample was removed from the analysis. In the case of the gene Photosystem II.LHC-II (AT1G29910), none of the samples returned good sequencing data and the gene was removed from the analysis. Before analysis the sequences were run through fastPHASE (Scheet and Stephens 2006) to phase any heterozygous bases. The software works by statistically calculating the probability of the most likely base pair for each heterozygous base from the other sequences. The sequences were then aligned and trimmed, so all sequences were the same length. They were then imported into ProSeq (Filatov 2009), a sequence alignment program, which also calculates population genetic equations from the sequences. Using ProSeq $\pi$, $\theta$, Tajima's D and FST were calculated. The general settings used were to analyse the whole sequence, to remove all indels and to ignore ambiguous sites/missing data. This ignore does not mean that all bases where missing data is present will actually be removed, rather if the base is only missing from one or two of the sequences then the base will be assumed from the other sequences aiding in filling in sequencing gaps. If three or more sequences being analysed are missing the base, then the base is ignored. For $\pi$, $\theta$ and Tajima's D calculations were made within the control and spring groups, as well as between the two. For FST, due to the nature of the calculation, only between control and spring could be calculated.

Calculations were made for all of the genes of interest from the sequenced DNA, along will all contigs from the RNA-Seq data (Chapter 3). For the RNA-Seq data, each contig had to be 100 bases or more in length to be included in the analysis in order to avoid bias in the statistics due to short sequences. The sequences were again called for heterozygous bases using fastPHASE before analysis, and aligned in ProSeq.

The genes of interest were searched against the Arabidopsis genome using BLAST, to identify introns within the sequences. The sequences were also converted into amino acid sequences using ExPAsy (Artimo, Jonnalagedda *et al.* 2012), and a protein blast was carried out to identify the correct reading frame for each sequence. The sequences were then compared to identify any protein differences between sequences.

## 5.3 Results

The results for each test on the genes of interest can be seen in Table 5.2 and Table 5.3. The calculated values were initially compared between spring and control, for each gene of interest (Figure 5.1). This could not be calculated for FST or Dxy due the comparison being between spring and control itself.

| | | Pi (π) | | | | Theta (θ) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Groups | | Compiled | | Groups | | Compiled | |
| Gene | Origin | Number of bases | Pi | Number of Bases | Pi | Number of bases | Theta | Number of Bases | Theta |
| BCA1 | Control | 301 | 0.0391 | 301 | 0.04786 | 301 | 0.0491 | 301 | 0.06268 |
| | Spring | 695 | 0.03497 | | | 695 | 0.04098 | | |
| CDKB1;1 | Control | 712 | 0.02566 | 475 | 0.01694 | 712 | 0.0419 | 475 | 0.03667 |
| | Spring | 477 | 0.00737 | | | 477 | 0.01187 | | |
| EPF1 | Control | 837 | 0.04062 | 834 | 0.03918 | 837 | 0.03529 | 834 | 0.03692 |
| | Spring | 835 | 0.03795 | | | 835 | 0.03465 | | |
| ERECTA | Control | 185 | 0.0237 | 185 | 0.0339 | 185 | 0.0187 | 185 | 0.02456 |
| | Spring | 185 | 0.02932 | | | 185 | 0.02769 | | |
| LHC_I16 | Control | 326 | 0.01575 | 326 | 0.01536 | 326 | 0.01387 | 326 | 0.01447 |
| | Spring | 484 | 0.01656 | | | 484 | 0.01868 | | |
| MYB88 | Control | 322 | 0.01543 | 322 | 0.01542 | 322 | 0.01543 | 322 | 0.02197 |
| | Spring | 322 | 0.01242 | | | 322 | 0.02051 | | |
| RIB5P | Control | 445 | 0.01485 | 403 | 0.01285 | 445 | 0.0149 | 403 | 0.01417 |
| | Spring | 403 | 0.01007 | | | 403 | 0.01346 | | |
| RUBINTER | Control | 521 | 0.033 | 377 | 0.02965 | 521 | 0.05669 | 377 | 0.06652 |
| | Spring | 405 | 0.02377 | | | 405 | 0.03125 | | |
| RUBS | Control | 438 | 0.02036 | 417 | 0.02015 | 438 | 0.02339 | 417 | 0.02784 |
| | Spring | 460 | 0.01249 | | | 460 | 0.01846 | | |
| SCREAM2 | Control | 161 | 0.05344 | 161 | 0.06109 | 161 | 0.06836 | 161 | 0.0903 |
| | Spring | 632 | 0.05962 | | | 632 | 0.07323 | | |
| YODA | Control | 627 | 0.00241 | 534 | 0.00231 | 627 | 0.00336 | 534 | 0.00418 |
| | Spring | 534 | 0.00303 | | | 534 | 0.00451 | | |

Table 5.2 –Results from the π and θ analysis. Groups mean analysis was carried out within each group of origin for each gene. Compiled means the two groups of origin were compared for each gene.

| | | FST | | | Dxy | | | Tajima's D | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Compiled | | | Compiled | | | Groups | | Compiled | |
| Gene | Origin | Number of Bases | FST | | Number of Bases | Dxy | | Number of bases | Tajima D | Number of Bases | Tajima D |
| BCA1 | Control | 301 | 0.10317 | | 301 | 0.0509 | | 301 | -0.89132 | 301 | -0.91671 |
| | Spring | | | | | | | 695 | -0.6816 | | |
| CDKB1;1 | Control | 475 | 0.09826 | | 475 | 0.01722 | | 712 | -1.67251 | 475 | -2.0327 |
| | Spring | | | | | | | 477 | -1.5699 | | |
| EPF1 | Control | 834 | -0.01091 | | 834 | 0.03898 | | 837 | 0.65216 | 834 | 0.23324 |
| | Spring | | | | | | | 835 | 0.41134 | | |
| ERECTA | Control | 185 | 0.34525 | | 185 | 0.04049 | | 185 | 1.05474 | 185 | 1.30972 |
| | Spring | | | | | | | 185 | 0.23361 | | |
| LHC_I16 | Control | 326 | -0.01551 | | 326 | 0.01524 | | 326 | 0.53108 | 326 | 0.20906 |
| | Spring | | | | | | | 484 | -0.46867 | | |
| MYB88 | Control | 322 | 0.1828 | | 322 | 0.01704 | | 322 | 0.00225 | 322 | -1.09772 |
| | Spring | | | | | | | 322 | -1.6541 | | |
| RIB5P | Control | 403 | 0.01249 | | 403 | 0.01293 | | 445 | -0.01327 | 403 | -0.325 |
| | Spring | | | | | | | 403 | -1.00406 | | |
| RUBINTER | Control | 377 | 0.02044 | | 377 | 0.02994 | | 521 | -1.80222 | 377 | -2.0971 |
| | Spring | | | | | | | 405 | -1.00776 | | |
| RUBS | Control | 417 | 0.27663 | | 417 | 0.023 | | 438 | -0.54053 | 417 | -1.02516 |
| | Spring | | | | | | | 460 | -1.37803 | | |
| SCREAM2 | Control | 161 | 0.02762 | | 161 | 0.06455 | | 161 | -0.94347 | 161 | -1.27962 |
| | Spring | | | | | | | 632 | -1.01352 | | |
| YODA | Control | 534 | 0.06094 | | 534 | 0.00238 | | 627 | -1.00422 | 534 | -1.36913 |
| | Spring | | | | | | | 534 | -1.18769 | | |

Table 5.3 –Results from the FST, Dxy and Tajima's D analysis. Groups mean analysis was carried out within each group of origin for each gene. Compiled means the two groups of origin were compared for each gene.

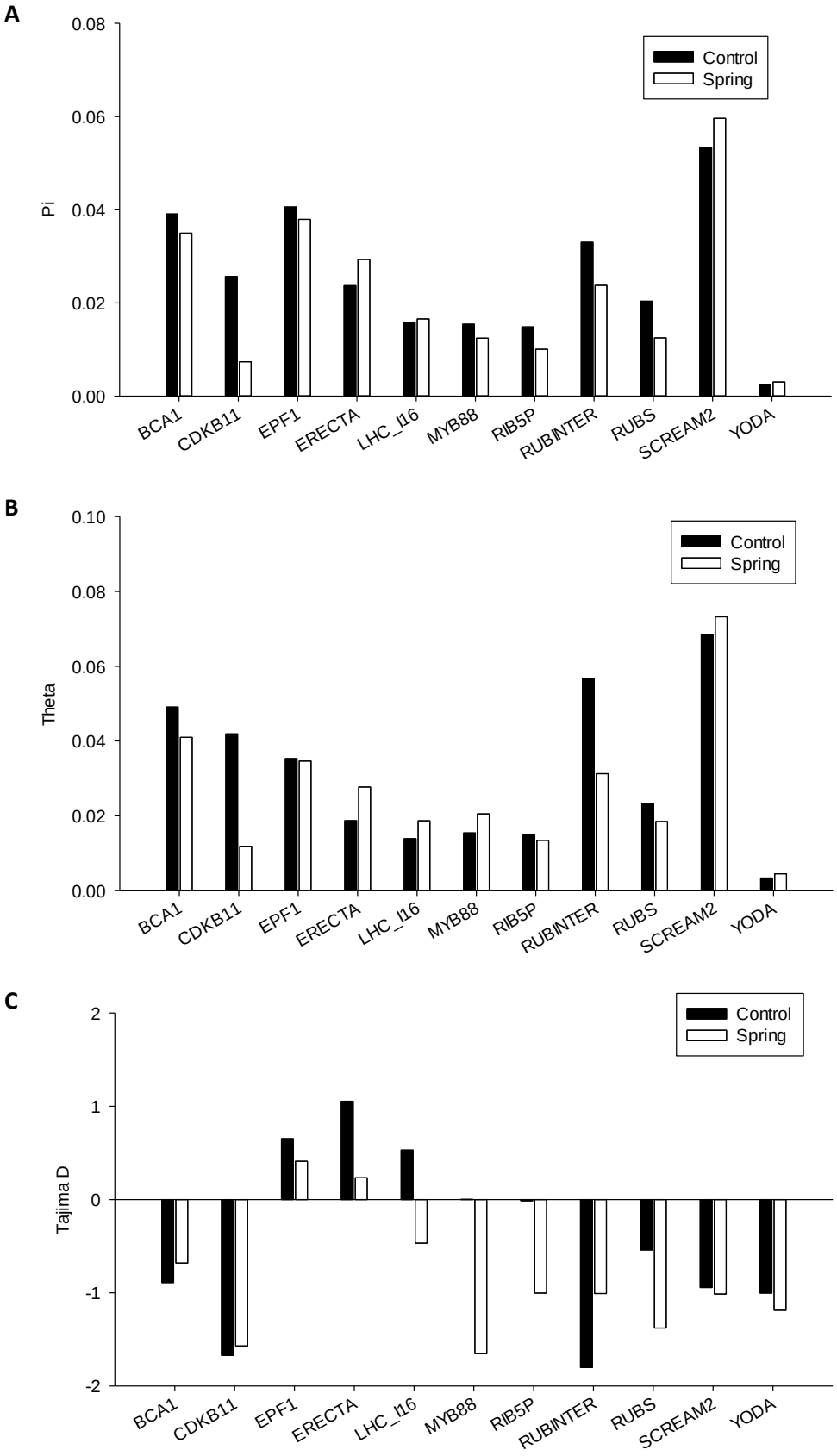



Figure 5.1 – The comparisons between spring and control for each gene for (**A**) π, (**B**) θ and (**C**) Tajima's D.

A paired t-test was carried out to investigate if there was a significant difference between control and spring plants for π, θ and Tajima's D (Table 5.4). Again this could not be calculated for FST or Dxy due the comparison being between spring and control itself.

| | Control Average | Spring Average | Paired T-Test P Value |
|---|---|---|---|
| **Pi (π)** | 0.0259 | 0.0225 | 0.1438 |
| **Theta (θ)** | 0.0310 | 0.0269 | 0.3035 |
| **Tajims's D** | -0.4207 | -0.8473 | 0.0703 |

Table 5.4 – The averages for π, θ and Tajima's D for control and spring for the genes of interest, with a paired, two-tailed t-test to see if the difference between control and spring is significant for each.

The average was calculated for π, θ and Tajima's D for all of the RNA-Seq contigs, and a paired t-test was again carried out to investigate if there was a significant difference between control and spring (Table 5.5).

| | Control Average | Spring Average | Paired T-Test P Value |
|---|---|---|---|
| **Pi (π)** | 0.0127 | 0.0103 | **<0.0001**** |
| **Theta (θ)** | 0.0191 | 0.0122 | **<0.0001**** |
| **Tajims's D** | -1.2462 | -0.6036 | **<0.0001**** |

Table 5.5 – The averages for π, θ and Tajima's D for control and spring for all RNA-Seq contigs, with a paired, two-tailed t-test to see if the difference between control and spring is significant for each.

Although no statistical test can be carried out on FST and Dxy, they give an indication of the amount of genetic differentiation between two populations. The average FST and Dxy was calculated from the RNA-Seq data (Table 5.6).

| | FST | Dxy |
|---|---|---|
| **RNA-Seq Data** | 0.0467 (SD ± 0.0637) | 0.0112 (SD ± 0.0105) |
| **Genes of Interest** | 0.1001 (SD ± 0.1202) | 0.0284 (SD ± 0.0186) |

Table 5.6 – The average for FST and Dxy for the RNA-Seq data set with standard deviation (SD).

To compare the sequences to the whole RNA-Seq data set the distribution of the total number of contigs was plotted next to the distribution of the genes of interest, to see if they differed. The percentage of the total count across a range was calculated and compared for each statistic (Figure 5.2-5.6). A Mann-Whitney U statistic was carried out to see if the distributions were significantly different (Table 5.7).

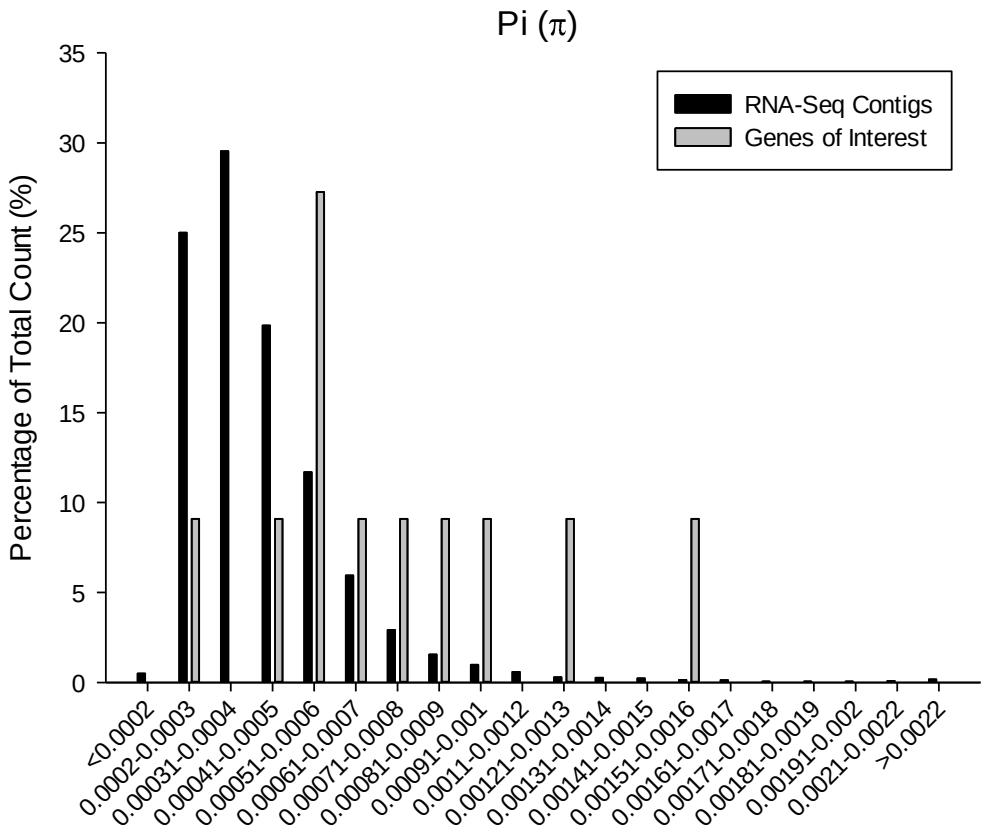



Figure 5.2 – The percentage of the total count of both the RNA-Seq contigs and the genes of interest for π.

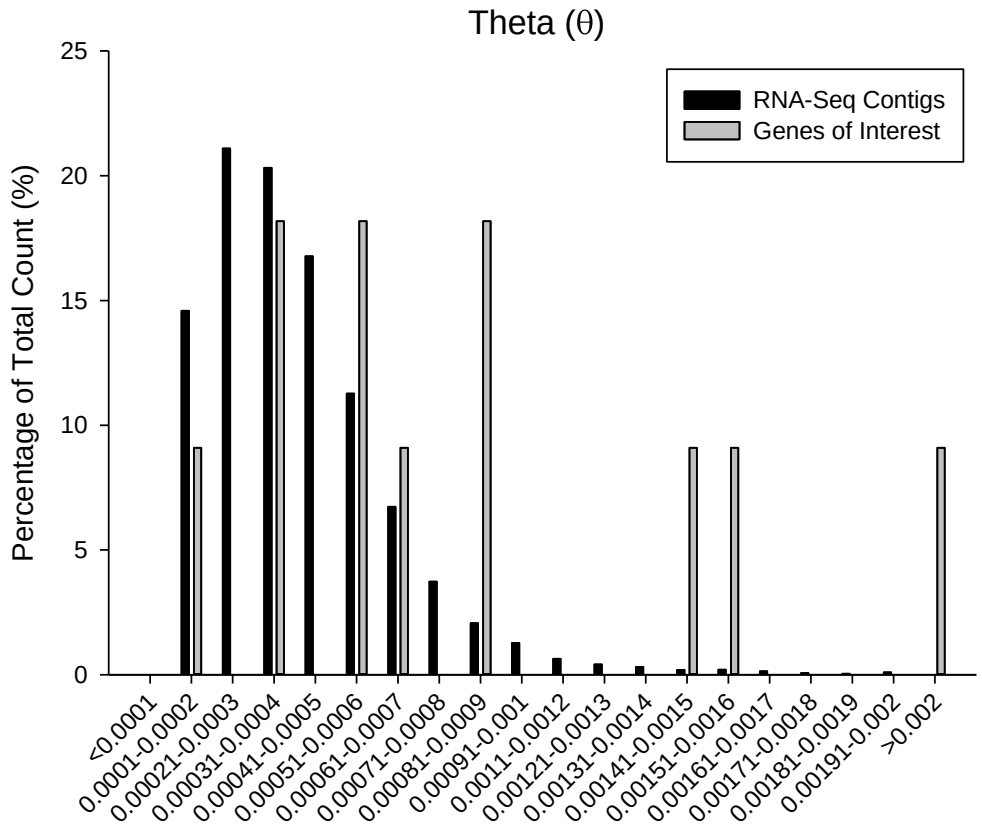



Figure 5.3 – The percentage of the total count of both the RNA-Seq contigs and the genes of interest for θ.

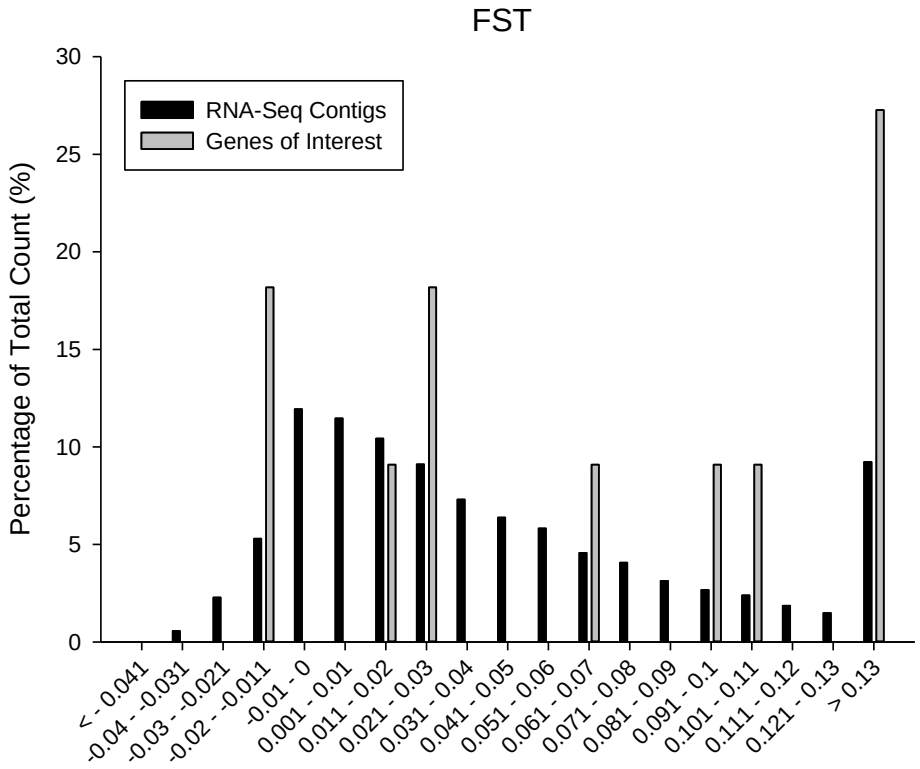



Figure 5.4 – The percentage of the total count of both the RNA-Seq contigs and the genes of interest for FST.

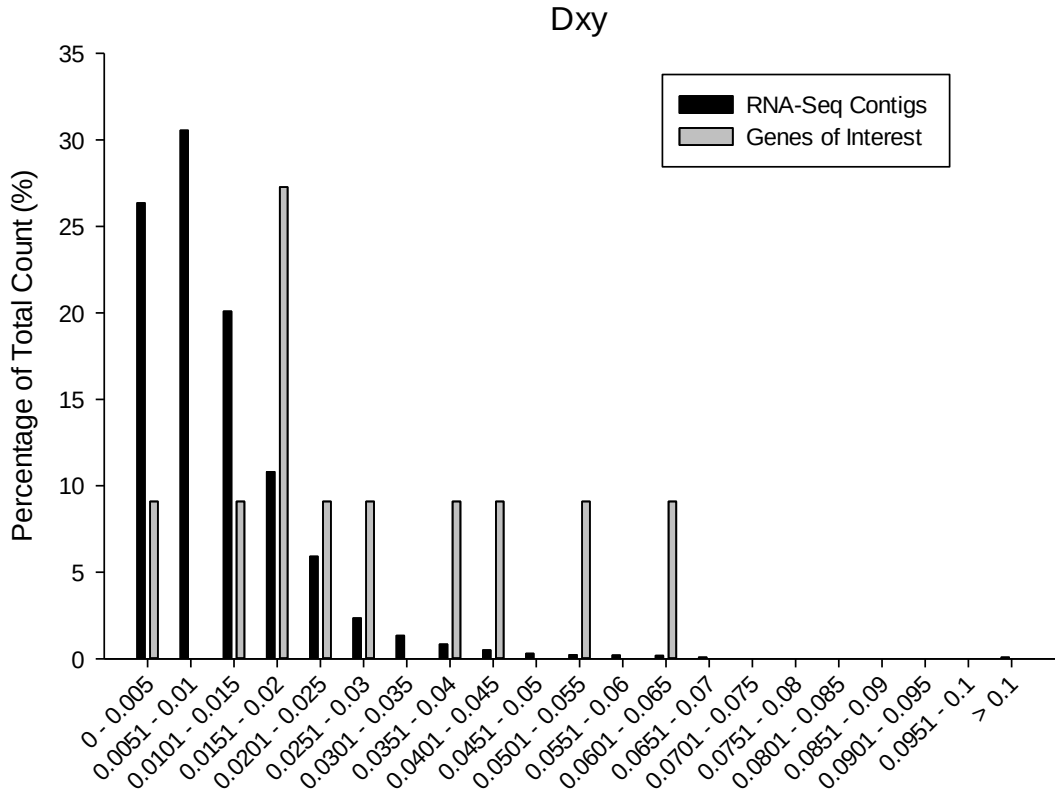



Figure 5.5 – The percentage of the total count of both the RNA-Seq contigs and the genes of interest for Dxy.

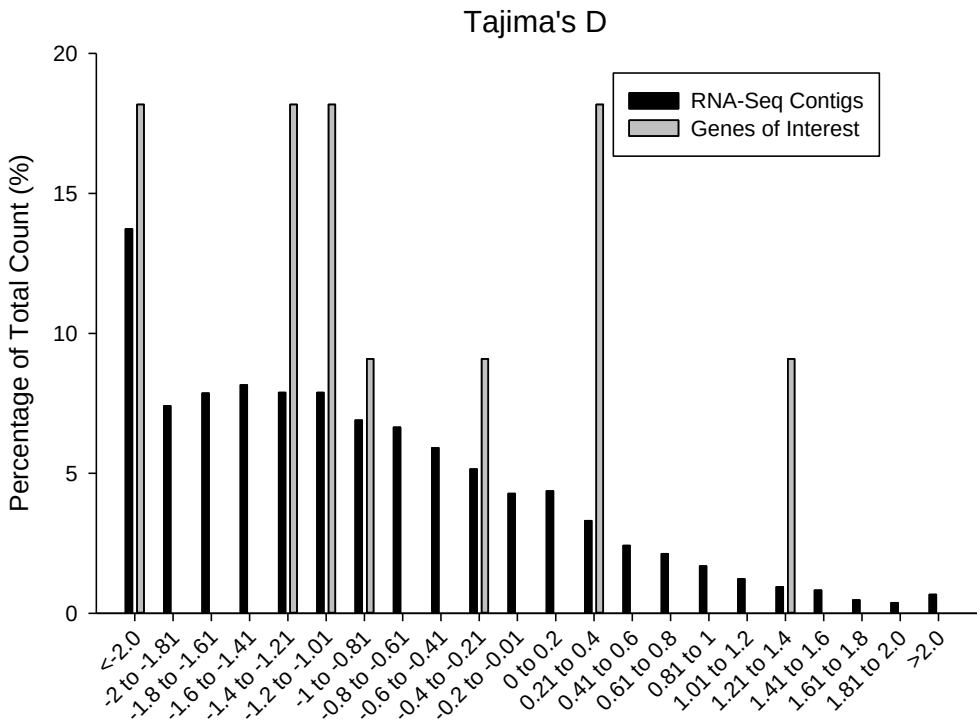



Figure 5.6 – The percentage of the total count of both the RNA-Seq contigs and the genes of interest for Tajima's D.

| | **Mann-Whitney U** |
|---|---|
| **FST** | 0.1900 |
| **Dxy** | **0.0001**** |
| **Tajima's D** | 0.0900 |
| **π** | **0.0001**** |
| **θ** | **0.0001**** |

Table 5.7 – The Mann-Whitney U statistic to show if the distributions of the RNA-Seq data and genes of interest are significantly different.

It was shown that 7 of the 11 genes of interest contained introns, after being searched against the *Arabidopsis* genome. These sequences were then converted into amino acid sequences and the coding proteins were analysed for fixed differences. No fixed amino acid differences were found, and the variation was only contained in one of the sequences when present.

## 5.4 Discussion

The use of population genetics allows for a different aspect of *P. lanceolata* to be investigated, to potentially identify reasons for why we have seen the previous phenotypic and physiological differences on a molecular level. The polymorphism measurements give a good basic indication of the variation in the sequences, with more polymorphisms indicative of increased variation. Across the whole transcriptome there was a significantly higher level of polymorphism in the control plants compared to the spring plants ($\pi$ shows a 23.30% greater value of $\pi$ ($P < 0.0001$), $\theta$ shows a 56.56% greater value of $\theta$ ($P < 0.0001$)). This reduction in polymorphism suggests that the spring population is derived from the control population (a bottleneck), because when a subpopulation deviates away, only a percentage of the variation of the original population is transferred. Bottlenecks can often be coupled with a selective sweep occurring in the sub-population to any environmental stimuli the new population is under; which would also decrease variation. This could be expected in the spring, as even though the control population is close to the spring, it contains many different environmental stimuli. Lower levels of polymorphism have been associated in response to stresses and bottlenecks in other species, and even 'polymorphism deserts' have been noted due to biotic or abiotic stresses in early ancestors of Rice (Krishnan S, Waters *et al.* 2014), where no polymorphism occurs in one population compared to another. Low genetic diversity has also been seen in domesticated species, when compared to their natural counterparts, because of breeding selection for desirable traits (Nabholz, Sarah *et al.* 2014). Tajima's D statistic was used to confirm this hypothesis, and a significant negative decrease was found in the spring plants compared to the control plants (106.46% lower ($P < 0.0001$)). A negative Tajima's D is indicative of a bottle neck, furthering the conclusion that the spring plants are derived from the control plants. For the genes of interest alone, the same pattern was found, but the differences were not significant. This is probably due to the low number of replications compared to the whole RNA-Seq data set. When percentage of the total count for RNA-Seq was plotted the genes of interest trended to the higher proportion of $\pi$ and $\theta$, which would suggest that we have correctly targeted where more variation is occurring. The distribution was also shown to be significantly different for $\pi$ and $\theta$. Introns could also be present in the genes of interest because full length DNA sequences were used. The presence of introns in 7 of the 11 genes of interest could explain the higher levels of polymorphism in the genes of interest, and are therefore not comparable with the RNA-Seq data, but the spring and control plants can be compared within the genes of interest. The comparison showed the control plants did possess higher polymorphism in most cases, with only a selection (ERECTA, LHC_I16, SCREAM2 and YODA) showing higher polymorphism in the spring plants, which agrees with the idea of a bottleneck. The Tajima's D of the individual genes of interest, showed the spring plants had lower Tajima's D

values in most cases, with only BCA1,CDKB1;1 and RUBINTER being higher, again agreeing with the hypothesis that a bottle neck has occurred in the spring plants in these genes. None of the these genes were in the top and bottom 5% of Tajima's D values of the RNA-Seq data set, which could suggest that these changes are negligible.

Using FST and Dxy we were able to explore further into the genetic diversity of the genes of interest. The very low FST and Dxy values of the RNA-Seq indicate there is no evidence for genetic diversity between the two populations. If there was selection in one of the populations you would expect there to be some genetic variation, suggesting there is no evidence for selection towards elevated $CO_2$. High FST values are often notable between species with large geographical differences (Lundqvist and Andersson 2001), and at distances between sites of many kilometres (sites of different altitudes) using wind pollinated species, low values have been recorded (Chapman, Hiscock *et al.* 2013). With only a distance of approximately 100 meters separating the two sites a low FST is not unexpected. The top 5% was calculated from the RNA-Seq data set for both FST and Dxy, and these contained 489 and 488 respectively. For FST three of the DNA sequenced genes fell into the top 5% of the RNA-Seq data (ERECTA,MYB88 and RUBS), and five fell into the top 5% for Dxy (EPF1, ERECTA, RUBINTER, RUBS and SCREAM2). The genes that fall into the top 5% can be considered as significant, and indicate that the genetic difference occurring between the sequences can be explained by population structure. Given that only 0.5 of the 10 candidate genes would be expected in the top 5% of the distribution, having 3 and 5 suggests higher values of FST and Dxy respectively in the genes of interest relative to the RNA-Seq data. Genes that are differentially expressed have previously been shown to trend to higher FST and Dxy values (Chapman, Hiscock *et al.* 2013), and as differentially expressed genes were selected, could explain this trend. The relationship could be due to linkage disequilibrium between signatures of divergence in promoters controlling gene expression, and the coding region of the genes. The Mann-Whitney U showed that the distribution of Dxy was significantly different ($P < 0.0001$) trending towards the higher values, again suggesting a greater proportion of high Dxy values in genes of interest. Genes which have lower polymorphism in the spring plants, and fall into the top 5% of FST and Dxy are of interest because there is evidence of higher selection pressures due to lower polymorphism, which are due solely to the genetic diversity within the spring population. Selection could be occurring in these genes, potentially explaining the observed phenotypic differences.

Any variations within the sequences have the chance to be non-synonymous, in other words change the amino acid associated with that nucleotide where variation is occurring. No fixed

differences between the spring and control plants were found in the sequences of the RNA-Seq data, which is unusual in such a large a data set. Also, no fixed differences were found in the genes from the DNA sequencing. However, there was variation in these sequences, and so the DNA sequences that were in the top 5% of FST and Dxy were translated into amino acid sequences to see if any of the genetic variation caused changes in the amino acids in a high percentage of population, which would suggest a move towards an allelic difference. The coding regions of the genes were scanned, but no differences were observed between the two populations that were present in more than one individual.

As the two populations are relatively recently derived (thousands of years), these results are not totally unexpected, and amino acid changes were identified in individuals in some genes, which could be the start of the selection process occurring. It is also now accepted that a large proportion of phenotypic differences are not controlled via molecular divergence, and rather regulation of gene expression is a major driver (Carroll 2005). This was highlighted by King and Wilson (2014), when the sequences of the chimpanzee and human were compared, and the amount of molecular divergence was far too low to represent the level of anatomical differences between the two species (King and Wilson 2014). Although this is still a premature assumption, most divergence appears not to be in the protein coding regions, but rather in promoter or structural mutations (Hoekstra and Coyne 2007). This could suggest that the large gene expression differences we have identified may alone be responsible for the phenotypes we have identified, and show the start of speciation. Identifying mutations in these upstream regions causing differential gene expression is difficult, as it could be genetic differences in the promoters themselves, or transcription factors which bind to the promoters. As the *P. lanceolata* genome remains unannotated, identifying transcription factors would be difficult as they can be elusive even in annotated genomes. This also does not mean the plants are not under an epigenetic effect, and in fact could be in part what is causing the differential expression; but this cannot be confirmed using the genetic sequence alone.

Although population statistics give a good indication of several parameters, they are not strict measures and so can only be used to infer conclusions. These analyses have given an overview of the genetic differentiation that has occurred between the two populations, and although no concrete conclusions can be made from the data, several suggestions can be made giving an insight into the genetic structure of the two populations. No fixed differences between the sequences suggest that the differences are not due to adaptations, as the sequences would

contain fixed allelic differences, indicating the gene expression differences are playing a major role.

# General Discussion

## 6.1 General Discussion

There is currently no study which provides a holistic review of the adaptive responses to multi-generational exposure to elevated $CO_2$. This is an important aspect of understanding how plants will respond to our future climate, and of direct relevance to global challenges including food, water and energy security, as well as the conservation and management of biodiversity over the coming decades. A range of responses to multi-generational elevated atmospheric $CO_2$ have been researched individually, but these have produced often conflicting results (Van Gardingen, Grace *et al.* 1997). Due to single responses being measured, it is not always apparent why these differences are being observed. This research has produced an overview of multi-generational plant responses to elevated $CO_2$, enabling linkages between phenotypic responses, and the underlying genetics and genomics for the responses seen (Figure 6.1).

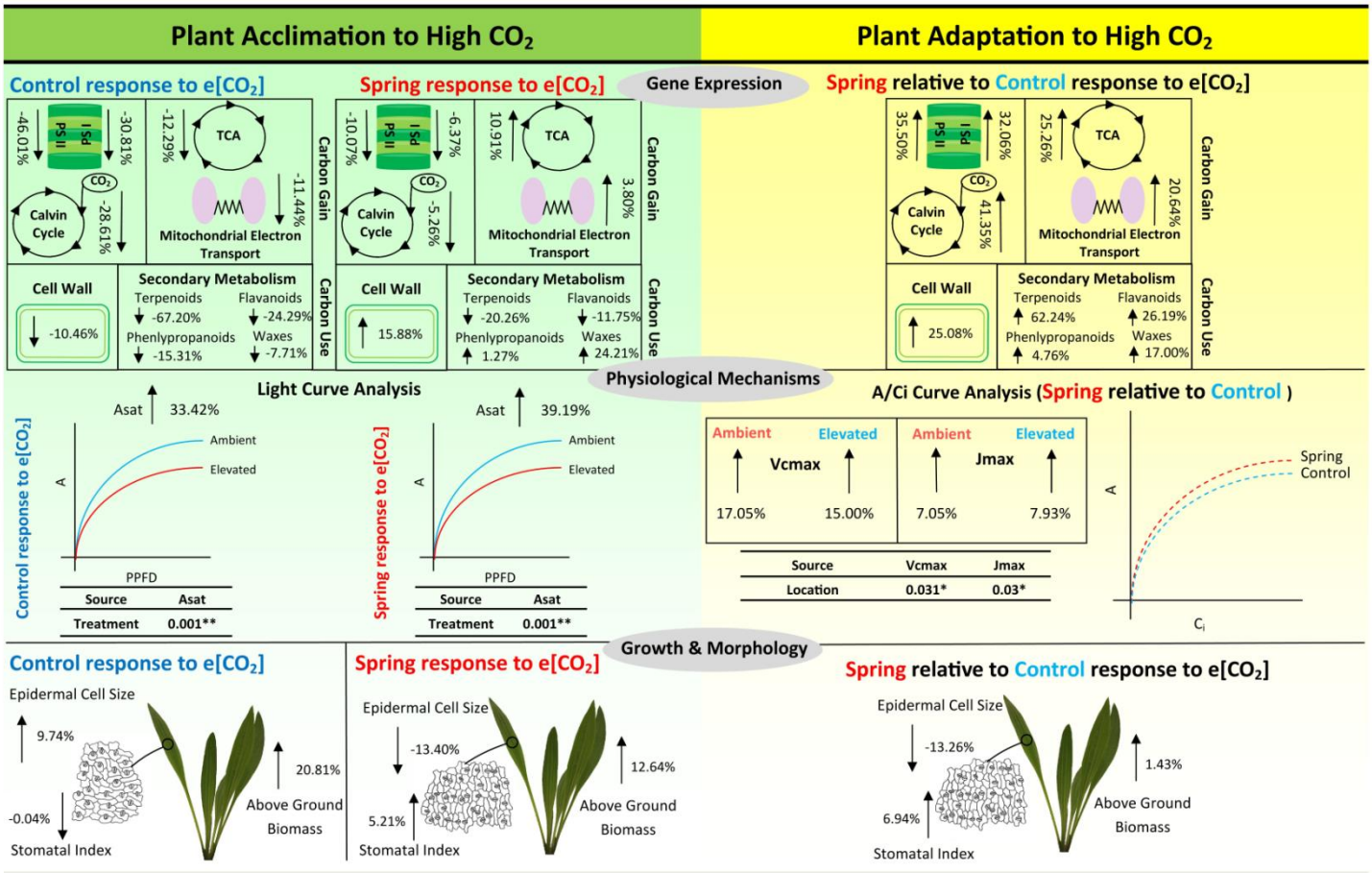



Figure 6.1 - Plant Acclimation and Adaptation to future high $CO_2$. A schematic overview of plant acclimation and adaptation to elevated $CO_2$, considering gene expression, physiological mechanisms and growth responses following short (acclimation) and long-term (adaptation) exposure. For gene expression categories, all annotated genes were averaged, within particular function categories, as defined in MapMan including Calvin Cycle, light reactions of photosynthesis, TCA and mitochondrial electron transport, cell wall and secondary metabolism. For physiological analysis, leaves were subjected to short-term changes in PAR and $CO_2$ to generate response curves. Growth and morphological responses were quantified in the controlled environment experiment described in Chapter 2.

Global gene expression showed that control plants expressed numerous significant changes to elevated $CO_2$, whereas the changes in the spring plants were magnitudes lower. There were a total of 1739 differentially expressed genes between locations and only 117 between treatments. This suggests that the major genomic differences are pre-adaptation to the location of origin-here, either spring or control. The selection of genes identified are indicative of a suite of genes that could control adaptation to elevated $CO_2$ within the spring, as they no longer respond to the stimuli of $CO_2$.

On a phenotypic level there are stomatal differences which appear counter-intuitive (an increase in stomatal index in spring plants in response to elevated $CO_2$), but appear to have no negative affect on plant growth or physiological mechanisms. Spring plants did show a significant decrease in leaf size compared to control plants in elevated $CO_2$ ($P < 0.05$), although above ground biomass did not exhibit a significant difference. The stomatal pathway was elucidated in *P. lanceolata* using RNA-Seq to provide gene expression across the pathway. Two genes which act during the asymmetric divisions (SCRM2 and CDBK1;1 ; refer to Figure 3.17) of stomatal patterning show interesting expression patterns. They are both positive regulators, and both show an increase in expression in elevated $CO_2$ in the spring plants, and no change in the control plants in response to elevated $CO_2$ (both with a significant ($P < 0.05$) effect between the spring and control site). This indicates that the extra proliferation of stomata in the spring plants may be occurring during this stage in the stomatal patterning pathway.

Although the number of stomata is important, the efficiency of stomatal opening and closing determines a large proportion of their functionality. Using stomatal response data we analysed how the stomata are functioning in both elevated and ambient $CO_2$, and also how quickly they can respond to changes in $CO_2$ concentration. The $C_i$, stomatal conductance and assimilation was consistently higher in the spring plants grown in elevated $CO_2$ (SE) than all other groups, coupled with consistently lower water use efficiency (iWUE). This suggests the SE plants have enhanced flux of $CO_2$ and $H_2O$ through the stomata, partially resulting from increased stomatal numbers, but also enhanced function. The guard cell length showed a significant treatment effect ($P < 0.05$), with both control and spring plants increasing guard cell size in elevated $CO_2$. This suggests the spring plants not only have more stomata, but also larger stomata, which again explains the increased flux through the stomata. While SE had consistently higher $C_i$, stomatal conductance and assimilation, control plants grown in ambient $CO_2$ (CA) had consistently lower values than all other groups, except for iWUE which was the highest. CE and SA were not as consistent, and varied when ambient conditions were increased to elevated $CO_2$. Measurements were conducted

in both ambient and elevated $CO_2$ for all four groups (CA, CE, SA and SE). When measured in ambient $CO_2$, CE had a higher assimilation rate than SA, but when measured in elevated $CO_2$ they have very similar rates. For iWUE measured in ambient $CO_2$, SA and CE have a similar iWUE, whereas measured in elevated $CO_2$; SA has a higher iWUE than CE. This suggests the spring plants are displaying an adaptive response to elevated $CO_2$, as they outperform the control plants under these conditions even after a growth at ambient $CO_2$. When $CO_2$ concentrations were increased from ambient to elevated $CO_2$, spring and control plants reacted in a similar manner, both increasing $C_i$ at similar rates. This would suggest that stomatal opening and closing does not differ between the spring and control plants, but rather aperture is the point of control to explain differences between the two groups of plants. Interestingly one of the two genes that were significantly different in both locations in response to elevated $CO_2$was an early auxin induced protein (AT3G04730), which is interesting as auxin is known to antagonise effects of ABA on stomatal opening and closure (Tanaka, Sano *et al.* 2006). The gene is expressed in CA but not in CE, and in SE but not in SA, so this could be responsible in part for control of the stomatal aperture when the $CO_2$ concentration was increased when measuring the SA and CE plants.

Stomata alone could not be responsible for all of the phenotypic changes observed underpinning growth, and thus, biochemical pathways were also investigated at the level of gene expression. These pathways were elucidated using RNA-Seq, which highlighted extensive differences in gene expression across both the photosynthesis and respiration pathways. The former showed a decrease in expression in the control plants (average 35.14% decrease across the pathway) in response to elevated $CO_2$, whereas the spring plants showed a much lower change (average 7.23% decrease across the pathway) and even increases in some genes across the pathway. Light curve (AQ) analysis produced only a significant treatment effect ($P < 0.01$) suggesting that light harvesting for assimilation did not differ between populations even with the major gene expression changes. The A/Ci curve on the other hand did show a significant location effect for both Jmax and Vcmax ($P < 0.05$), with spring plants having higher values for both (Vcmax; SE/CE =15% increase, SA/CA =17% increase. Jmax; SE/CE =8% increase, SA/CA =7% increase). This suggests the spring plants are more efficient and have a higher capacity to harvest $CO_2$. This could be associated with the increased number of stomata, but in general, this result is not confirmed by the literature which generally shows a decline in Vcmax with increasing atmospheric $CO_2$ and less investment in RuBisCO, and thus requires further investigation. It is possible that some conformation change in the RuBsiCO protein may have occurred in the spring plants, but this was not noted at the level of gene expression, where RuBisCO gene expression was decreased in SE/SA and CE/CA. More research is required to understand this response in greater detail.

The respiration pathway exhibited a similar pattern of gene expression as the photosynthesis pathway, but unlike the photosynthesis pathway, an increase in respiration is the most common previously documented age-dependant response (Markelz, Vosseller *et al.* 2014). Control plants showed a down regulation across the pathway (average decrease of 11.87% across the pathway) whereas the spring plants showed an up regulation (average of 7.36% across the pathway). Increased reparatory gene expression in the spring plants could provide energy required for additional growth, and also investment in secondary carbon chemistry, leading to spring plants of increased size, but no significant differences in biomass was noted between locations.

The genetic diversity between samples provided both some interesting conclusions, and ideas which need to be further investigated. The higher amount of polymorphism in the control plants compared to the spring plants ($\pi$ shows a 23.30% greater $\pi$ ($P < 0.0001$), $\theta$ shows a 56.56% greater $\theta$ ($P < 0.0001$)) is consistent with the hypothesis that the spring population is derived from a more widespread population, comprising of the control site. This was confirmed using Tajima's D, which produced a more negative value in the spring plants compared to control plants (106.46% lower ($P < 0.0001$)), indicative of a bottleneck, which again suggests the spring population is derived from the control population. With the sequencing of more individuals, the size and age of the bottleneck may be identified. A low FST and Dxy value from the RNA-Seq data indicated there was no evidence for genetic divergence between the two populations. If there was selection in one of the populations it is expected that there would be some divergence between populations, suggesting there is no evidence for selection towards elevated $CO_2$. Using DNA sequences of genes of interest (from the stomatal and photosynthesis pathway), we were able to show that there may be evidence for genetic divergence between populations specifically in the genes of interest. The top 5% was calculated from the RNA-Seq data set for both FST and Dxy, and these contained 489 and 488 genes respectively. For FST three of the eleven DNA sequenced genes fell into the top 5% of the RNA-Seq data (ERECTA,MYB88 and RUBS), and for Dxy five of the eleven fell into the top 5% (EPF1, ERECTA, RUBINTER, RUBS and SCREAM2). The genes that fall into this top 5% can be considered as significant, and indicate that the genetic difference occurring between the sequences can be explained by population structure in these cases. Given that only 0.5 of the 10 candidate genes would be expected in the top 5% of the distribution, having 3 and 5 (FST and Dxy respectively) suggests there may be evidence for genetic divergence in the genes of interest relative to the RNA-Seq data. No fixed differences were found in either the genes of interest, or the RNA-Seq data. If there were any sequence based adaptations one might expect to see some fixed differences. However the presumed recent divergence of the spring and control populations might manifest as increased, but not fixed, divergence, as seen here.

Due to this, the genetic diversity provides limited insight into the observed gene expression differences, and follow up experiments will be necessary to ultimately define the cause of the phenotype and gene expression differentiation. The data suggests the response may not be due strictly to adaptation, as there are no fixed site differences between the control and spring plants. We cannot rule out that there are fixed differences in upstream regulators and transcription factors that were not sequenced, as *P. lanceolata* has no reference genome, and so we do not know the completeness of the transcriptome for *P. lanceolata* used here. Another explanation could be the changes are due to epigenetic effects, as these changes are not seen at the DNA sequence, and can be heritable (Bond and Baulcombe 2014). To test this methylation sequencing could be carried out on a selection of sequences to see if any evidence of epigenetic patterning is present. These epigenetic changes could be parental (as in the epigenetic changes are heritable or occurred in the seeds), and so another way to test whether epigenetic motifs are present would be to use seeds which haven't been grown in their original environment. This should be tested by growing seed collected from the SA and CE plants from the chamber experiments, and then growing those seeds in both ambient and elevated $CO_2$ concentrations. If any epigenetic changes towards elevated $CO_2$ are heritable, they may be lost after a generation in ambient $CO_2$, and therefore changes may not be expressed in the next generation if epigenetic effects are being expressed (Daxinger and Whitelaw 2010).

It is difficult to generalise the genetic responses as we have only looked into one species, and this is a limitation, but also provides a great basis to move forward as the multi-generational genetics research is so novel. To follow this research up it would be a natural progression to investigate new species using the same protocol to see if the responses are consistent, as studies within the spring have shown inconsistencies in some traits (Miglietta, Raschi *et al.* 1998). This would confirm whether these inconsistencies are due to other variables, or if plants are just responding in different ways to multi-generational $CO_2$. It would also be interesting to investigate a suite of species with as many varied aspects as possible, such as different chromosome numbers (polyploid, diploid etc.) and different photosynthetic mechanisms (C4 and CAM plants), as well as different breeding systems. The variability of species utilising C3 photosynthesis in response to short term elevated $CO_2$ has been noted (Wullschelger 1993), and so it would be interesting to see if this variability carried forward to multi-generational exposure to elevated $CO_2$. Looking into C4 and CAM plants you would expect different responses as they have already adapted to have a more efficient photosynthetic system, but it would be interesting to see how they react compared to the C3 species.

Another limitation is the use of one spring site, which could be argued that potentially has variables which may differ from other natural springs which are causative of the responses we see. The site in Bossoleto was chosen due to work previously carried out indicating it was the spring site in Italy with the lowest variation of variables from the control site (Bettarini, Grifoni et al. 1999). The main variable of concern is $H_2S$, as this can directly impact plant physiology (Duan, Ma et al. 2015). The concentration in Bossoleto was one of the lowest measured, and the levels within the spring were below detection (Bettarini, Grifoni *et al.* 1999). The temperature at the site was elevated during the day compared to temperature of the control site (Bettarini, Grifoni *et al.* 1999), but this is due to the greenhouse gas effect caused by the elevated $CO_2$, and so cannot be avoided when using natural springs. An increase in humidity inside the spring has been associated with these increases in temperature (Van Gardingen, Grace *et al.* 1997), but whether this off-sets the heightened temperature to affect iWUE is yet to be elucidated. The flora within the Bossoleto spring is of a broad taxonomic range (Selvi 1999), more so than the other natural springs in Italy. This is one of the reasons why the site was originally chosen, but also means it could contain different bacteria, or attract different animal species which could affect plant growth. A positive to the increased flora is that you would expect conditions in the spring to be less harsh if they harbour more diverse species, and so harsh variables could be reduced. Several C3 grassland species were consistently present at the edge of all springs across Italy, with niches of varying species only present as one moves closer to the $CO_2$ vents of each spring (Selvi 1999). Even so, other springs should be investigated using the same species to see how consistent these responses are, as there could be other factors such as bacteria or animal life within the springs which have not been noted which could be affecting the plants. This should be carried out in springs across the world where possible, as these are the most likely to have the most diverse environments. Using a chamber experiment hopefully negates any potential variables, but being aware of any other diverse conditions is important as they could have adaptive properties which are carried through to the chamber experiment and affect the $CO_2$ response.

Other considerations within the spring include the gradient of $CO_2$ from the vent to the outside of the spring. Where possible seeds were collected from places equidistant from the vent, but there may be places where the $CO_2$ concentration differs. There was a family effect in numerous phenotypes from both chamber experiments ($P<0.05$), which could be a consequence of seeds being collected from different concentrations within the spring. On the other hand, family effects have previously been noted in studies using *P. lanceolata* (Klus, Kalisz *et al.* 2001), suggesting that this family effect might not be related to the gradient and could just be an effect of the variability of *P. lanceolata*, but it would be interesting to take seed from several points across the gradient

to see if there was any correlation. Another reason for the possible family effect is the potential of the spring plants being derived from the control population at different times. This could lead to the spring plants being exposed to elevated $CO_2$ for different periods of time, leading to variation between the plants. It is impossible from the number of reps we have to determine if any plants have derived earlier than other, but if more reps could be analysed a pattern may become apparent which correlates with changes in the phenotype. This also raises another potential concern, which is we do not know how long the plants have been present in the spring and so cannot confirm the samples we took have in fact had multi-generational exposure to elevated $CO_2$. Seeds could have been transferred between the spring and control site at any point in time, as animals freely move between the two sites. The population genetics indicated that the spring population was derived from the control population, a result which would not have occurred if they had not had a degree of multi-generational exposure, but we cannot confirm that they have been present in the spring for hundreds of years, although it is expected that this is the case.

The A/Ci curves produced some interesting results, which are novel and have not previously been reported. Confirmation of these results with more replication and possibly with other species would be valuable. Another follow up would be to carry out low oxygen curves to see if the efficacy for $CO_2$ differs when there is less oxygen that could bind to RuBisCO.

Although there are potential issues, these appear to have not affected this set of results, and all follow up experiments discussed do not negate from the results we have, rather could be used to further these findings. In conclusion, this study provides the first transcriptomic analysis of multi-generational effects of elevated atmospheric $CO_2$ on native vegetation. We have identified an interesting phenotype, and have provided reasoning for this phenotype at the level of gene expression. We have also provided an insight into how plants could potentially react to prolonged multi-generational elevated $CO_2$ concentrations on a much broader level, investigating the biochemical and physiological pathways, and finding that the previously documented responses to short term elevated $CO_2$ may not be true for long term exposure. We were unable to confirm whether the traits are due to adaptive or epigenetic reasons, but have provided a reliable platform to base future work on to elucidate this area further.

# 7.1 Appendix

The results from the chromosome analysis can be seen in Table 7.1. The results showed the four samples sent for genome size estimation all had the same genome size (the slight differences are negligible; if the plants had experienced a ploidy event the genome size would be double). All of the plants were 2x ploidy so it can be assumed that all plants have the same genome size.

| | Ploidy | Genome Size ( DNA pg/2C) |
|---|---|---|
| S1 | 2x | 2.61 |
| S2 | 2x | 2.56 |
| C1 | 2x | 2.61 |
| C2 | 2x | 2.63 |
| S1 | 2x | - |
| S2 | 2x | - |
| S3 | 2x | - |
| S4 | 2x | - |
| S5 | 2x | - |
| S6 | 2x | - |
| S7 | 2x | - |
| S8 | 2x | - |
| S9 | 2x | - |
| S10 | 2x | - |
| S11 | 2x | - |
| C1 | 2x | - |
| C2 | 2x | - |
| C3 | 2x | - |
| C4 | 2x | - |
| C5 | 2x | - |
| C6 | 2x | - |
| C7 | 2x | - |
| C8 | 2x | - |
| C9 | 2x | - |
| C10 | 2x | - |
| C11 | 2x | - |

*Table 7.1 – Table showing the results of the chromosome analysis. S refers to a plant originating from the spring, and C refers to a plant originating from the control site. The ploidy is the number of sets of chromosomes the plant has.*

The genes used in Figure 3.22 are listed in Table 7.2, along with the associated statistics. Statistics were carried out on the sub-set of genes associated with photosynthesis.

| Gene Description | AT Number | ((CE-CA)/CA)*100 (Percentage Change (%)) | ((SE-SA)/SA)*100 (Percentage Change (%)) | Statistics (Two way ANOVA) | | |
|---|---|---|---|---|---|---|
| | | | | Location | Treatment | Location*Treatment |
| TCA / org. transformation.TCA.pyruvate DH.E1 | at1g24180 | -38.724 | 34.623 | 0.012* | 0.398 | 0.005** |
| TCA / org. transformation.TCA.pyruvate DH.E1 | at5g50850 | -34.066 | -17.375 | 0.197 | 0.398 | 0.472 |
| TCA / org. transformation.TCA.pyruvate DH.E1 | at1g59900 | -11.057 | 5.837 | 0.024* | 0.488 | 0.069 |
| TCA / org. transformation.TCA.pyruvate DH.E1 | at1g01090 | -8.134 | 11.391 | 0.351 | 0.586 | 0.026* |
| TCA / org. transformation.TCA.pyruvate DH.E1 | at1g30120 | -0.945 | 27.660 | 0.008** | 0.199 | 0.097 |
| TCA / org. transformation.TCA.aconitase | at2g05710 | -14.548 | 11.173 | 0.136 | 0.627 | 0.191 |
| TCA / org. transformation.TCA.aconitase | at4g35830 | -26.632 | -11.722 | 0.116 | 0.166 | 0.339 |
| gluconeogenesis.Malate DH | at5g09660 | -36.620 | -10.800 | 0.064 | 0.141 | 0.290 |
| gluconeogenesis.Malate DH | at2g22780 | -20.274 | 5.181 | 0.149 | 0.314 | 0.075 |
| gluconeogenese/ glyoxylate cycle.malate synthase | at5g03860 | -2.250 | 79.661 | 0.192 | 0.362 | 0.309 |
| gluconeogenese/ glyoxylate cycle.PEPCK | at5g65690 | -49.118 | -13.398 | 0.341 | 0.104 | 0.225 |
| gluconeogenese/ glyoxylate cycle.PEPCK | at4g37870 | -14.476 | 37.178 | 0.010** | 0.413 | 0.128 |
| gluconeogenese/ glyoxylate cycle.pyruvate dikinase | at4g15530 | -51.909 | -34.569 | 0.334 | 0.170 | 0.469 |
| TCA / org. transformation.TCA.succinate dehydrogenase | at5g66760 | -9.680 | 4.291 | 0.012* | 0.510 | 0.199 |
| TCA / org. transformation.TCA.succinate dehydrogenase | at5g40650 | -23.728 | 15.077 | 0.071 | 0.541 | 0.091 |
| TCA / org. transformation.TCA.malate DH | at1g53240 | -21.509 | 24.483 | 0.003** | 0.584 | 0.005** |
| mitochondrial electron transport / ATP synthesis.NADH-DH.localisation not clear | at5g11770 | -32.326 | 14.201 | 0.003** | 0.211 | 0.005** |
| mitochondrial electron transport / ATP synthesis.NADH-DH.localisation not clear | at5g52840 | -28.539 | 21.027 | 0.017* | 0.479 | 0.005** |
| mitochondrial electron transport / ATP synthesis.NADH-DH.localisation not clear | at5g18800 | -15.406 | -14.774 | 0.281 | 0.189 | 0.561 |
| mitochondrial electron transport / ATP synthesis.cytochrome c reductase | at5g13440 | -12.998 | 12.378 | 0.003** | 0.658 | 0.059 |
| mitochondrial electron transport / ATP synthesis.cytochrome c reductase | at4g32470 | -41.924 | 12.693 | 0.038* | 0.039* | 0.005** |
| mitochondrial electron transport / ATP synthesis.NADH-DH.type II.internal matrix | at2g29990 | -17.455 | 54.194 | 0.337 | 0.455 | 0.361 |

| | | | | | | |
|---|---|---|---|---|---|---|
| mitochondrial electron transport / ATP synthesis.NADH-DH.type II.external | at4g05020 | -36.977 | -29.075 | 0.138 | 0.087 | 0.534 |
| mitochondrial electron transport / ATP synthesis.NADH-DH.type II.external | at4g21490 | -8.258 | -4.777 | 0.027* | 0.280 | 0.510 |
| mitochondrial electron transport / ATP synthesis.alternative oxidase | at5g64210 | -37.006 | 1.048 | 0.162 | 0.323 | 0.238 |
| mitochondrial electron transport / ATP synthesis.cytochrome c oxidase | at1g28140 | -30.828 | 13.746 | 0.222 | 0.141 | 0.005** |
| mitochondrial electron transport / ATP synthesis.cytochrome c oxidase | at3g15640 | -27.048 | 7.878 | 0.065 | 0.326 | 0.064 |
| mitochondrial electron transport / ATP synthesis.cytochrome c | at4g10040 | -26.280 | 31.227 | 0.137 | 0.616 | 0.334 |
| mitochondrial electron transport / ATP synthesis.cytochrome c | at5g40810 | -12.606 | 1.718 | 0.081 | 0.311 | 0.127 |
| mitochondrial electron transport / ATP synthesis.cytochrome c | at1g15220 | -24.400 | -0.247 | 0.070 | 0.349 | 0.291 |
| PS.lightreaction.photosystem II.LHC-II | at3g08940 | -41.494 | 2.662 | 0.236 | 0.002** | 0.003** |
| PS.lightreaction.photosystem II.LHC-II | at5g54270 | -36.383 | -10.014 | 0.177 | 0.002** | 0.022* |
| PS.lightreaction.photosystem II.LHC-II | at2g34420 | -47.612 | -11.438 | 0.370 | 0.002** | 0.003** |
| PS.lightreaction.photosystem II.LHC-II | at2g34430 | -47.721 | -10.429 | 0.428 | 0.002** | 0.003** |
| PS.lightreaction.photosystem II.LHC-II | at3g27690 | -33.029 | -21.385 | 0.003** | 0.018* | 0.196 |
| PS.lightreaction.photosystem II.LHC-II | at2g05100 | -37.966 | -8.882 | 0.272 | 0.002** | 0.032* |
| PS.lightreaction.photosystem II.LHC-II | at1g29910 | -51.086 | -29.143 | 0.411 | 0.002** | 0.191 |
| PS.lightreaction.photosystem II.LHC-II | at4g10340 | -29.673 | -11.992 | 0.092 | 0.002** | 0.157 |
| PS.lightreaction.photosystem II.LHC-II | at1g29920 | -47.605 | -10.647 | 0.430 | 0.002** | 0.003** |
| PS.lightreaction.photosystem II.LHC-II | at2g05070 | -41.671 | 9.322 | 0.258 | 0.006** | 0.003** |
| PS.lightreaction.photosystem II.LHC-II | at1g15820 | -41.820 | 6.316 | 0.102 | 0.068 | 0.111 |

| | | | | | | |
|---|---|---|---|---|---|---|
| PS.lightreaction.other electron carrier (ox/red).plastocyanin | at1g76100 | -30.525 | 1.292 | 0.023* | 0.004** | 0.009** |
| PS.lightreaction.ATP synthase.delta chain | at4g09650 | -38.629 | -0.485 | 0.073 | 0.005** | 0.022* |
| PS.lightreaction.photosystem I.LHC-I | at3g47470 | -24.444 | -13.866 | 0.319 | 0.003** | 0.260 |
| PS.lightreaction.photosystem I.LHC-I | at1g61520 | -30.920 | -10.226 | 0.126 | 0.006** | 0.165 |
| PS.lightreaction.photosystem I.LHC-I | at1g19150 | -20.766 | -11.531 | 0.003** | 0.018* | 0.405 |
| PS.lightreaction.photosystem I.PSI polypeptide subunits | at1g31330 | -29.654 | -2.844 | 0.102 | 0.012* | 0.093 |
| PS.lightreaction.photosystem I.PSI polypeptide subunits | at5g64040 | 51.852 | -44.378 | 0.283 | 0.173 | 0.045* |
| PS.lightreaction.photosystem I.PSI polypeptide subunits | at2g20260 | -31.978 | -2.648 | 0.074 | 0.018* | 0.115 |
| PS.lightreaction.photosystem I.PSI polypeptide subunits | at1g55670 | -37.249 | 3.760 | 0.087 | 0.009** | 0.016* |
| PS.lightreaction.photosystem I.PSI polypeptide subunits | at1g08380 | -25.231 | -11.700 | 0.157 | 0.011* | 0.271 |
| PS.lightreaction.photosystem I.PSI polypeptide subunits | at4g02770 | -35.540 | -8.719 | 0.298 | 0.002** | 0.046* |
| PS.lightreaction.photosystem I.PSI polypeptide subunits | at4g12800 | -34.088 | -4.071 | 0.074 | 0.004** | 0.039 |
| PS.calvin cyle.rubisco small subunit | at1g67090 | -29.895 | -5.737 | 0.03* | 0.008** | 0.111 |
| PS.calvin cyle.transketolase | at2g45290 | -15.510 | 34.261 | 0.228 | 0.162 | 0.093 |
| PS.calvin cyle.Rib5P Isomerase | at3g04790 | -35.812 | 8.263 | 0.003** | 0.017* | 0.005** |

Table 7.2 - Summary of genes which are displayed in Figure 3.22. A two-way ANOVA was conducted to identify differentially expressed genes, with location (spring and control) and treatment (ambient and elevated $CO_2$) as factors, where * P≤0.05; ** P≤0.01.

# 8.1 References

Acharya, B. and S. Assmann (2009). Hormone interactions in stomatal function. Plant Molecular Biology **69**(4): 451-462.

Adhikari, B., D. Wall and B. Adams (2009). Desiccation survival in an Antarctic nematode: molecular analysis using expressed sequenced tags. BMC Genomics **10**(1): 69.

Ainsworth, E. A., P. A. Davey, C. J. Bernacchi, O. C. Dermody, E. A. Heaton, D. J. Moore, P. B. Morgan, S. L. Naidu, H.S. Yoo Ra, X.G. Zhu, P. S. Curtis and S. P. Long (2002). A meta-analysis of elevated [$CO_2$] effects on soybean (*Glycine max*) physiology, growth and yield. Global Change Biology **8**(8): 695-709.

Ainsworth, E. A. and S. P. Long (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**(2): 351-372.

Ainsworth, E. A. and A. Rogers (2007). The response of photosynthesis and stomatal conductance to rising [$CO_2$]: mechanisms and environmental interactions. Plant, Cell & Environment **30**(3): 258-270.

Ainsworth, E. A., A. Rogers, H. Blum, J. Nösberger and S. P. Long (2003). Variation in acclimation of photosynthesis in Trifolium repens after eight years of exposure to Free Air $CO_2$ Enrichment (FACE). Journal of Experimental Botany **54**(393): 2769-2774.

Ainsworth, E. A., A. Rogers, L. O. Vodkin, A. Walter and U. Schurr (2006). The effects of elevated $CO_2$ concentration on soybean gene expression. An analysis of growing and mature leaves. Plant Physiology **142**(1): 135-147.

Allen, G. J., S. P. Chu, C. L. Harrington, K. Schumacher, T. Hoffmann, Y. Y. Tang, E. Grill and J. I. Schroeder (2001). A defined range of guard cell calcium oscillation parameters encodes stomatal movements. Nature **411**(6841): 1053-1057.

Allen, G. J., K. Kuchitsu, S. P. Chu, Y. Murata and J. I. Schroeder (1999). Arabidopsis abi1-1 and abi2-1 phosphatase mutations reduce abscisic acid–induced cytoplasmic calcium rises in guard cells. The Plant Cell Online **11**(9): 1785-1798.

Andalo C., B. Godelle, and M. Monsseau (1999). Are *Arabidopsis thaliana* from a natural $CO_2$ spring adapted to elevated $CO_2$. Ecosystem Responses To $CO_2$: The MAPLE Project Results**:** 158-167.

Anders, S. and W. Huber (2010). Differential expression analysis for sequence count data. Genome Biology **11**(10): R106.

Andersson, I. and A. Backlund (2008). Structure and function of Rubisco. Plant Physiology and Biochemistry **46**(3): 275-291.

Andrews, T. J. (1988). Catalysis by cyanobacterial ribulose-bisphosphate carboxylase large subunits in the complete absence of small subunits. Journal of Biological Chemistry **263**(25): 12213-12219.

Arp, W. J. (1991). Effects of source-sink relations on photosynthetic acclimation to elevated $CO_2$. Plant, Cell & Environment **14**(8): 869-875.

Artimo, P., M. Jonnalagedda, K. Arnold, D. Baratin, G. Csardi, E. de Castro, S. Duvaud, V. Flegel, A. Fortier, E. Gasteiger, A. Grosdidier, C. Hernandez, V. Ioannidis, D. Kuznetsov, R. Liechti, S. Moretti, K. Mostaguir, N. Redaschi, G. Rossier, I. Xenarios and H. Stockinger (2012). ExPASy: SIB bioinformatics resource portal. Nucleic Acids Research **40**(W1): W597-W603.

Assmann, S. M. (1999). The cellular basis of guard cell sensing of rising $CO_2$. Plant, Cell & Environment **22**(6): 629-637.

Assmann, S. M., L. Simoncini and J. I. Schroeder (1985). Blue light activates electrogenic ion pumping in guard cell protoplasts of Vicia faba. Nature **318**(6043): 285-287.

Baker, M. (2012). De novo genome assembly: what every biologist should know. Nature Methods **9**(4): 333-337.

Balaguer, L., E. Manrique, A. de los Rios, C. Ascaso, K. Palmqvist, M. Fordham and J. D. Barnes (1999). Long-term responses of the green-algal lichen *Parmelia caperata* to natural $CO_2$ enrichment. Oecologia **119**(2): 166-174.

Bandres, E., E. Cubedo, X. Agirre, R. Malumbres, R. Zarate, N. Ramirez, A. Abajo, A. Navarro, I. Moreno, M. Monzo and J. Garcia-Foncillas (2006). Identification by real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. Molecular Cancer **5**(1): 29.

Bannister, A. J. and T. Kouzarides (2011). Regulation of chromatin by histone modifications. Cell Research **21**(3): 381-395.

Barber, R. D., D. W. Harmer, R. A. Coleman and B. J. Clark (2005). GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. Physiological Genomics **21**(3): 389-395.

Barghini, E., L. Natali, R. M. Cossu, T. Giordani, M. Pindo, F. Cattonaro, S. Scalabrin, R. Velasco, M. Morgante and A. Cavallini (2014). The Peculiar Landscape of Repetitive Sequences in the Olive (*Olea europaea* L.) Genome. Genome Biology and Evolution **6**(4): 776-791.

Bassett, I. J. and C. W. Crompton (1968). Pollen morphology and chromosome numbers of the family Plantaginaceae in North America. Canadian Journal of Botany **46**(4): 349-361.

Battisti, D. S. and R. L. Naylor (2009). Historical warnings of future food insecurity with unprecedented seasonal heat. Science **323**(5911): 240-244.

Batzoglou, S., D. B. Jaffe, K. Stanley, J. Butler, S. Gnerre, E. Mauceli, B. Berger, J. P. Mesirov and E. S. Lander (2002). ARACHNE: A whole-genome shotgun assembler. Genome Research **12**(1): 177-189.

Bazzaz, F. A., M. Jasieński, S. C. Thomas and P. Wayne (1995). Microevolutionary responses in experimental populations of plants to $CO_2$-enriched environments: parallel results from two model systems. Proceedings of the National Academy of Sciences **92**(18): 8161-8165.

Beerling, D. (1993). Changes in the stomatal density of *Betula nana* leaves in reponse to increases in atmospheric carbon dioxide concentrations. Special Papers in Palaeontology **49**: 181-187.

Beerling, D. J. and W. G. Chaloner (1992). Stomatal density as an indicator of atmospheric $CO_2$ concentration. The Holocene **2**(1): 71-78.

Beerling, D. J. and W. G. Chaloner (1993). Stomatal density responses of Egyptian *Olea europaea* L. leaves to $CO_2$ change since 1327 BC. Annuals of Botany **71**(5): 431-435.

Beerling, D. J. and W. G. Chaloner (1994). Atmospheric $CO_2$ changes since the last glacial maximum: evidence from the stomatal density record of fossil leaves. Review of Palaeobotany and Palynology **81**(1): 11-17.

Beerling, D. J., W. G. Chaloner, B. Huntley, J. A. Pearson and M. J. Tooley (1993). Stomatal density responds to the glacial cycle of environmental change. Proceedings of the Royal Society of London. Series B: Biological Sciences **251**(1331): 133-138.

Beerling, D. J. and F. I. Woodward (1997). Changes in land plant function over the Phanerozoic: reconstructions based on the fossil record. Botanical Journal of the Linnean Society **124**(2): 137-153.

Berg, L. (2007). Introductory botany: plants, people, and the environment. Cengage Learning.

Bergmann, D. C. and F. D. Sack (2007). Stomatal development. Annual Review of Plant Biology, Palo Alto, Annual Reviews **58:** 163-181.

Berk, R. and J. MacDonald (2008). Overdispersion and poisson regression. Journal of Quantitative Criminology **24**(3): 269-284.

Bernacchi, C., P. Morgan, D. Ort and S. Long (2005). The growth of soybean under free air [$CO_2$] enrichment (FACE) stimulates photosynthesis while decreasing *in vivo* Rubisco capacity. Planta **220**(3): 434-446.

Bernacchi, C. J., C. Pimentel and S. P. Long (2003). In vivo temperature response functions of parameters required to model RuBP-limited photosynthesis. Plant, Cell & Environment **26**(9): 1419-1430.

Bernstein, L., P. Bosch, O. Canziani, Z. Chen, R. Christ, O. Davidson, W. Hare, S. Huq, D. Karoly and V. Kattsov (2007). Climate change 2007: synthesis report. An assessment of the Intergovernmental Panel on Climate Change. IPCC **20**: 2011.

Bettarini, I., G. Calderoni, F. Miglietta, A. Raschi and J. Ehleringer (1995). Isotopic carbon discrimination and leaf nitrogen content of *Erica arborea* L. along a $CO_2$ concentration gradient in a $CO_2$ spring in Italy. Tree Physiology **15**(5): 327-332.

Bettarini, I., D. Grifoni, F. Miglietta and A. Raschi (1999). Local greenhouse effects in a $CO_2$ spring in central Italy. Ecosystem Responses To $CO_2$: The MAPLE Project Results**:** 13-23.

Bettarini I., F. P. Vaccari, F. Miglietta and A. Raschi (1999). Stomatal physiology and morphology of calcareous grasslands in a future $CO_2$ enriched world. Ecosystem Responses To $CO_2$: The MAPLE Project Results**:** 39-52.

Bird, A. (2007). Perceptions of epigenetics. Nature **447**(7143): 396-398.

Björkman, O. (1981). Responses to different quantum flux densities. Physiological Plant Ecology I., Springer Berlin Heidelberg. **12/A:** 57-107.

Bloem, M. W. and S. de Pee (2013). Nutrition and human rights: why meeting nutrient needs should be a human right. Advancing the Human Right to Health. Oxford, Oxford University Press.

Boetzer, M., C. V. Henkel, H. J. Jansen, D. Butler and W. Pirovano (2011). Scaffolding pre-assembled contigs using SSPACE. Bioinformatics **27**(4): 578-579.

Bonasio, R., S. Tu and D. Reinberg (2010). Molecular signals of epigenetic states. Science **330**(6004): 612-616.

Bond, D. M. and D. C. Baulcombe (2014). Small RNAs and heritable epigenetic variation in plants. Trends in Cell Biology **24**(2): 100-107.

Bos, M., H. Harmens and K. Vrieling (1986). Gene flow in Plantago I. Gene flow and neighbourhood size in *P. lanceolata*." Heredity **56**(1): 43-54.

Bos, M. and E. Van der Haring (1988). "Gene flow in Plantago II. Gene flow pattern and population structure. A simulation study." Heredity **61**(1): 1-11.

Boudolf, V., R. Barrôco, J. d. A. Engler, A. Verkest, T. Beeckman, M. Naudts, D. Inzé and L. De Veylder (2004). B1-type cyclin-dependent kinases are essential for the formation of stomatal complexes in *Arabidopsis thaliana*. The Plant Cell Online **16**(4): 945-955.

Bowes, G. (1991). Growth at elevated $CO_2$: photosynthetic responses mediated through Rubisco. Plant, Cell & Environment **14**(8): 795-806.

Bowman, S., M. Patel, P. Yerramsetty, C. Mure, A. Zielinski, J. Bruenn and J. Berry (2013). A novel RNA binding protein affects rbcL gene expression and is specific to bundle sheath chloroplasts in C4 plants. BMC Plant Biology **13**(1): 138.

Bryant, J. P., F. S. Chapin III and D. R. Klein (1983). Carbon/nutrient balance of boreal plants in relation to vertebrate herbivory. Oikos: 357-368.

Butler, J., I. MacCallum, M. Kleber, I. A. Shlyakhter, M. K. Belmonte, E. S. Lander, C. Nusbaum and D. B. Jaffe (2008). ALLPATHS: De novo assembly of whole-genome shotgun microreads. Genome Research **18**(5): 810-820.

Cardon, Z. G., J. A. Berry and I. E. Woodrow (1994). Dependence of the extent and direction of average stomatal response in *Zea mays* L. and *Phaseolus vulgaris* L. on the frequency of fluctuations in environmental stimuli." Plant Physiology **105**(3): 1007-1013.

Carrie, C., M. Murcha, E. Giraud, S. Ng, M. Zhang, R. Narsai and J. Whelan (2013). How do plants make mitochondria? Planta **237**(2): 429-439.

Carroll, S. B. (2005). Evolution at two levels: on genes and form. PLoS Biology **3**(7):245.

Castel, S. E. and R. A. Martienssen (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nature Review Genetics **14**(2): 100-112.

Catala, R., J. Ouyang, I. A. Abreu, Y. Hu, H. Seo, X. Zhang and N.-H. Chua (2007). The Arabidopsis E3 SUMO ligase SIZ1 regulates plant growth and drought responses. The Plant Cell Online **19**(9): 2952-2966.

Cavers, P. B., I. J. Bassett and C. W. Crompton (1980). The biology of Canadian weeds: 47. *Plantago lanceolata* L." Canadian Journal of Plant Science **60**(4): 1269-1282.

Chamaillé-Jammes, S. and W. J. Bond (2010). Will global change improve grazing quality of grasslands? A call for a deeper understanding of the effects of shifts from C4 to C3 grasses for large herbivores. Oikos **119**(12): 1857-1861.

Champagne, K. S. and T. G. Kutateladze (2009). Structural insight into histone recognition by the ING PHD fingers. Current Drug Targets **10**(5): 432.

Chang, S., J. Puryear and J. Cairney (1993). A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology Reporter **11**(2): 113-116.

Chapman, M. A., S. J. Hiscock and D. A. Filatov (2013). Genomic divergence during speciation driven by adaptation to altitude. Molecular Biology and Evolution **168**.

Chen, J., F.-H. Wu, W.-H. Wang, C.-J. Zheng, G.-H. Lin, X.-J. Dong, J.-X. He, Z.-M. Pei and H.-L. Zheng (2011). Hydrogen sulphide enhances photosynthesis through promoting chloroplast biogenesis, photosynthetic enzyme expression, and thiol redox modification in *Spinacia oleracea* seedlings. Journal of Experimental Botany **62**(13): 4481-4493.

Chen, L.-Q., C.-S. Li, W. G. Chaloner, D. J. Beerling, Q.-G. Sun, M. E. Collinson and P. L. Mitchell (2001). Assessing the potential for the stomatal characters of extant and fossil Ginkgo leaves to signal atmospheric $CO_2$ change. American Journal of Botany **88**(7): 1309-1315.

Cheng, S.-H., B. d. Moore and J. R. Seemann (1998). Effects of short- and long-term elevated $CO_2$ on the expression of ribulose-1,5-bisphosphate carboxylase/oxygenase genes and carbohydrate accumulation in leaves of *Arabidopsis thaliana* (L.) Heynh. Plant Physiology **116**(2): 715-723.

Chinnusamy, V. and J.-K. Zhu (2009). Epigenetic regulation of stress responses in plants. Current Opinion in Plant Biology **12**(2): 133-139.

Clark, R. M., G. Schweikert, C. Toomajian, S. Ossowski, G. Zeller, P. Shinn, N. Warthmann, T. T. Hu, G. Fu, D. A. Hinds, H. Chen, K. A. Frazer, D. H. Huson, B. Schölkopf, M. Nordborg, G. Rätsch, J. R. Ecker and D. Weigel (2007). Common Sequence Polymorphisms Shaping Genetic Diversity in *Arabidopsis thaliana*. Science **317**(5836): 338-342.

Clemens, S. (2006). Toxic metal accumulation, responses to exposure and mechanisms of tolerance in plants. Biochimie **88**(11): 1707-1719.
Compeau, P. E. C., P. A. Pevzner and G. Tesler (2011). How to apply de Bruijn graphs to genome assembly. Nature Biotechnology **29**(11): 987-991.

Constable, G. A. and H. M. Rawson (1980). Effect of leaf position, expansion and age on photosynthesis, transpiration and water use efficiency of cotton. Functional Plant Biology **7**(1): 89-100.

Cook, A. C., W. C. Oechel, B. Sveinbjornsson, A. Raschi, F. Miglietta, R. Tognetti and P. Van Gardingen (1997). Using Icelandic $CO_2$ springs to understand the long-term effects of elevated atmospheric $CO_2$. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 87-102.

Crafts-Brandner, S. J. and M. E. Salvucci (2000). Rubisco activase constrains the photosynthetic potential of leaves at high temperature and $CO_2$. Proceedings of the National Academy of Sciences **97**(24): 13430-13435.

Crane, P. R., P. Herendeen and E. M. Friis (2004). Fossils and plant phylogeny. American Journal of Botany **91**(10): 1683-1699.

Creelman, R. A. and J. E. Mullet (1995). Jasmonic acid distribution and action in plants: regulation during development and response to biotic and abiotic stress. Proceedings of the National Academy of Sciences **92**(10): 4114-4119.

Croxdale, J. L. (2000). Stomatal patterning in angiosperms. American Journal of Botany **87**(8): 1069-1080.

Cure, J. D. and B. Acock (1986). Crop responses to carbon dioxide doubling: a literature survey. Agricultural and Forest Meteorology **38**(1–3): 127-145.

Curtis, P. S. and X. Wang (1998). A meta-analysis of elevated $CO_2$ effects on woody plant mass, form, and physiology. Oecologia **113**(3): 299-313.

Cutler, J. M., D. W. Rains and R. S. Loomis (1977). The importance of cell size in the water relations of plants. Physiologia Plantarum **40**(4): 255-260.

Czechowski, T., M. Stitt, T. Altmann, M. K. Udvardi and W.R. Scheible (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiology **139**(1): 5-17.

Darbah, J. N. T., T. D. Sharkey, C. Calfapietra and D. F. Karnosky (2010). Differential response of aspen and birch trees to heat stress under elevated carbon dioxide. Environmental Pollution **158**(4): 1008-1014.

Davey, P. A., S. Hunt, G. J. Hymus, E. H. DeLucia, B. G. Drake, D. F. Karnosky and S. P. Long (2004). Respiratory oxygen uptake is not decreased by an instantaneous elevation of [$CO_2$], but is increased with long-term growth in the field at elevated [$CO_2$]. Plant Physiology **134**(1): 520-527.

Davey, P. A., A. J. Parsons, L. Atkinson, K. Wadge and S. P. Long (1999). Does photosynthetic acclimation to elevated $CO_2$ increase photosynthetic nitrogen-use efficiency? A study of three native UK grassland species in open-top chambers. Functional Ecology **13**: 21-28.

Daxinger, L. and E. Whitelaw (2010). Transgenerational epigenetic inheritance: More questions than answers. Genome Research **20**(12): 1623-1628.

de Boer, H. J., E. I. Lammertsma, F. Wagner-Cremer, D. L. Dilcher, M. J. Wassen and S. C. Dekker (2011). Climate forcing due to optimization of maximal leaf conductance in subtropical vegetation under rising $CO_2$. Proceedings of the National Academy of Sciences **108**(10): 4041-4046.

De Kort, H., K. Vandepitte and O. Honnay (2013). A meta-analysis of the effects of plant traits and geographical scale on the magnitude of adaptive differentiation as measured by the difference between QST and FST. Evolutionary Ecology **27**(6): 1081-1097.

De Souza, A. P., M. Gaspar, E. A. Da Silva, E. C. Ulian, A. J. Waclawovsky, M. Y. Nishiyama Jr, R. V. Dos Santos, M. M. Teixeira, G. M. Souza and M. S. Buckeridge (2008). Elevated $CO_2$ increases photosynthesis, biomass and productivity, and modifies gene expression in sugarcane. Plant, Cell & Environment **31**(8): 1116-1127.

Delucia, E., T. Sasek and B. Strain (1985). Photosynthetic inhibition after long-term exposure to elevated levels of atmospheric carbon dioxide. Photosynthesis Research **7**(2): 175-184.

Demidenko, N. V., M. D. Logacheva and A. A. Penin (2011). Selection and validation of reference genes for quantitative real-time PCR in buckwheat (*Fagopyrum esculentum*) based on transcriptome sequence data." PLoS ONE **6**(5): e19434.

Desikan, R., K. Last, R. Harrett-Williams, C. Tagliavia, K. Harter, R. Hooley, J. T. Hancock and S. J. Neill (2006). Ethylene-induced stomatal closure in Arabidopsis occurs via AtrbohF-mediated hydrogen peroxide synthesis. The Plant Journal **47**(6): 907-916.

Dietzel, L., K. Bräutigam and T. Pfannschmidt (2008). Photosynthetic acclimation: state transitions and adjustment of photosystem stoichiometry – functional relationships between short-term and long-term light quality acclimation in plants. FEBS Journal **275**(6): 1080-1088.

Ding, S., Q. Lu, Y. Zhang, Z. Yang, X. Wen, L. Zhang and C. Lu (2009). Enhanced sensitivity to oxidative stress in transgenic tobacco plants with decreased glutathione reductase activity leads to a decrease in ascorbate pool and ascorbate redox state. Plant Molecular Biology **69**(5): 577-592.

Dippery, J. K., D. T. Tissue, R. B. Thomas and B. R. Strain (1995). Effects of low and elevated $CO_2$ on C3 and C4annuals. Oecologia **101**(1): 13-20.

Doheny-Adams, T., L. Hunt, P. J. Franks, D. J. Beerling and J. E. Gray (2012). Genetic manipulation of stomatal density influences stomatal size, plant growth and tolerance to restricted water supply across a growth carbon dioxide gradient. Philisophical Transactions of the Royal Society of Biology **367** (10): 547-555.

Dong, J., C. A. MacAlister and D. C. Bergmann (2009). BASL Controls Asymmetric Cell Division in Arabidopsis. Cell **137**(7): 1320-1330.

Dooley, F. D., S. P. Nair and P. D. Ward (2013). Increased growth and germination success in plants following hydrogen sulfide administration. PLoS ONE **8**(4): 62048.

Dow, G. J., J. A. Berry and D. C. Bergmann (2014). The physiological importance of developmental mechanisms that enforce proper stomatal spacing in *Arabidopsis thaliana*. New Phytologist **201**(4): 1205-1217.

Drake, B. G., J. Azcon-Bieto, J. Berry, J. Bunce, P. Dijkstra, J. Farrar, R. M. Gifford, M. A. Gonzalez-Meler, G. Koch, H. Lambers, J. Siedow and S. Wullschleger (1999). Does elevated atmospheric $CO_2$ concentration inhibit mitochondrial respiration in green plants? Plant, Cell & Environment **22**(6): 649-657.

Drake, B. G., M. A. Gonzalez-Meler and S. P. Long (1997). More efficient plants: A consequence of rising atmospheric $CO_2$? Annual Review of Plant Physiology and Plant Molecular Biology **48**: 609-639.

Drake, P. L., R. H. Froend and P. J. Franks (2013). Smaller, faster stomata: scaling of stomatal size, rate of response, and stomatal conductance. Journal of Experimental Botany **64**(2): 495-505.

Du, Y., L. Zhang, F. Xu, B. Huang, G. Zhang and L. Li (2013). Validation of housekeeping genes as internal controls for studying gene expression during Pacific oyster (*Crassostrea gigas*) development by quantitative real-time PCR. Fish & Shellfish Immunology **34**(3): 939-945.

Du, Z., X. Zhou, Y. Ling, Z. Zhang and Z. Su (2010). AgriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Research **38**(suppl 2): W64-W70.

Duan, B., Y. Ma, M. Jiang, F. Yang, L. Ni and W. Lu (2015). Improvement of photosynthesis in rice (*Oryza sativa* L.) as a result of an increase in stomatal aperture and density by exogenous hydrogen sulfide treatment. Plant Growth Regulation **75**(1): 33-44.

Eamus, D. (1991). The interaction of rising $CO_2$ and temperatures with water use efficiency. Plant, Cell & Environment **14**(8): 843-852.

Eberhard, S., G. Finazzi and F.A. Wollman (2008). The dynamics of photosynthesis. Annual Review of Genetics **42**(1): 463-515.

Ehleringer J.R., Sandquist D.R. and Philips S.L. (1997). Burning coal seams in southern Utah: a natural system for studies of plant responses to elevated $CO_2$. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 56-68.

Fan, G. Z., Q. S. Cai, C. M. Wang, J. M. Wan and J. G. Zhu (2005). QTL for yield and its components responded to elevated $CO_2$ in rice (*Oryza sativa* L.). Acta genetica Sinica **32**(10): 1066-1073.

Fangmeier, A., U. Grüters, P. Högy, B. Vermehren and H. J. Jäger (1997). Effects of elevated $CO_2$, nitrogen supply and tropospheric ozone on spring wheat—II. Nutrients (N, P, K, S, Ca, Mg, Fe, Mn, Zn). Environmental Pollution **96**(1): 43-59.

Faria, T., D. Wilkins, R. T. Besford, M. Vaz, J. S. Pereira and M. M. Chaves (1996). Growth at elevated $CO_2$ leads to down-regulation of photosynthesis and altered response to high temperature in *Quercus suber* L. seedlings. Journal of Experimental Botany **47**(11): 1755-1761.

Farquhar, G. D., S. Caemmerer and J. A. Berry (1980). A biochemical model of photosynthetic $CO_2$ assimilation in leaves of C3 species. Planta **149**(1): 78-90.

Farquhar, G. D., J. R. Ehleringer and K. T. Hubick (1989). Carbon isotope discrimination and photosynthesis. Annual Review of Plant Physiology and Plant Molecular Biology **40**(1): 503-537.

Ferris, R., M. Sabatti, F. Miglietta, R. F. Mills and G. Taylor (2001). Leaf area is stimulated in Populus by free air $CO_2$ enrichment (POPFACE), through increased cell expansion and production. Plant, Cell & Environment **24**(3): 305-315.

Ferris, R. and G. Taylor (1994). Stomatal Characteristics of Four Native Herbs Following Exposure to Elevated $CO_2$. Annuals of Botany **73**(4): 447-453.

Field, C. B., V. R. Barros, K. Mach and M. Mastrandrea (2014). Climate change 2014: impacts, adaptation, and vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Field, C. B., R. B. Jackson and H. A. Mooney (1995). Stomatal responses to increased $CO_2$: implications from the plant to the global scale. Plant, Cell & Environment **18**(10): 1214-1225.

Filatov, D. A. (2009). Processing and population genetic analysis of multigenic datasets with ProSeq3 software. Bioinformatics **25**(23): 3189-3190.

Fleige, S. and M. W. Pfaffl (2006). RNA integrity and the effect on the real-time qRT-PCR performance. Molecular Aspects of Medicine **27**(2–3): 126-139.

Fordham, M. and J. D. Barnes (1999). Growth and photosynthesis capacity in *Agrostis canina* and *Plantago major* adapted to contrasting long-term atmospheric $CO_2$ concentrations. Ecosystem Responses To $CO_2$: The MAPLE Project Results**:** 143-157.

Fordham, M. C., J. D. Barnes, A. Raschi, I. Bettarini, H. G. Griffiths, F. Miglietta, R. Tognetti and P. R. v. Gardingen (1997). The impact of elevated $CO_2$ on the growth of *Agrostis canina* and *Plantago major*; adapted to contrasting $CO_2$ concentrations. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 174-196.

Foyer, C. H., J. Neukermans, G. Queval, G. Noctor and J. Harbinson (2012). Photosynthetic control of electron transport and the regulation of gene expression. Journal of Experimental Botany **63**(4): 1637-1661.

Franks, P. J. and G. D. Farquhar (2007). The mechanical diversity of stomata and its significance in gas-exchange control. Plant Physiology **143**(1): 78-87.

Frenck, G., L. Linden, T. N. Mikkelsen, H. Brix and R. B. Jørgensen (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**(5): 1163-1172.

Fukayama, H., M. Sugino, T. Fukuda, C. Masumoto, Y. Taniguchi, M. Okada, R. Sameshima, T. Hatanaka, S. Misoo, T. Hasegawa and M. Miyao (2011). Gene expression profiling of rice grown in free air $CO_2$ enrichment (FACE) and elevated soil temperature. Field Crops Research **121**(1): 195-199.

Gardiner, C. M., L. A. Guethlein, H. G. Shilling, M. Pando, W. H. Carr, R. Rajalingam, C. Vilches and P. Parham (2001). Different NK cell surface phenotypes defined by the DX9 antibody are due to KIR3DL1 gene polymorphism. The Journal of Immunology **166**(5): 2992-3001.

Gardner, M., S. Williamson, F. Casals, E. Bosch, A. Navarro, F. Calafell, J. Bertranpetit and D. Comas (2007). Extreme individual marker FST values do not imply population-specific selection in humans: the NRG1 example. Human Genetics **121**(6): 759-762.

Gay, A. P. and B. Hauck (1994). Acclimation of *Lolium temulentum* to enhanced carbon-dioxide concentration. Journal of Experimental Botany **45**(277): 1133-1141.

Geber and Dawson (1993). Evolutionary responses of plants to global change. Biotic Interactions and Global Change. P. M. Kareiva, J. G. Kingsolver & R. B. Huey. Sinauer Associates Inc., Sunderland, MA. 176-197

Gerhart, L. M. and J. K. Ward (2010). Plant responses to low [$CO_2$] of the past. New Phytologist **188**(3): 674-695.

Gesch, R. W., I. H. Kang, M. Gallo-Meagher, J. C. V. Vu, K. J. Boote, L. H. Allen and G. Bowes (2003). Rubisco expression in rice leaves is related to genotypic variation of photosynthesis under elevated growth $CO_2$ and temperature. Plant, Cell & Environment **26**(12): 1941-1950.

Gifford, R. M., H. Lambers and J. I. L. Morison (1985). Respiration of crop species under $CO_2$ enrichment. Physiologia Plantarum **63**(4): 351-356.

Gobert, A., S. Isayenkov, C. Voelker, K. Czempinski and F. J. M. Maathuis (2007). The two-pore channel TPK1 gene encodes the vacuolar $K^+$ conductance and plays a role in $K^+$ homeostasis. Proceedings of the National Academy of Sciences **104**(25): 10726-10731.

Gonzalez-Meler, M. A., M. Ribas-Carbo, J. N. Siedow and B. G. Drake (1996). Direct Inhibition of Plant Mitochondrial Respiration by Elevated $CO_2$. Plant Physiology **112**(3): 1349-1355.

Gonzalez-Meler, M. A., L. Taneva and R. J. Trueman (2004). Plant respiration and elevated atmospheric $CO_2$ concentration: cellular responses and global significance. Annals of Botany **94**(5): 647-656.

Göring, H., S. Koshuchowa and C. Deckert (1990). Influence of gibberellic acid on stomatal movement. Biochemie und Physiologie der Pflanzen **186**(5–6): 367-374.

Gowik, U. and P. Westhoff (2011). The path from C3 to C4 photosynthesis. Plant Physiology **155**(1): 56-63.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman and A. Regev (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology **29**(7): 644-652.

Grace, J., P. v. Gardingen, A. Raschi and R. Tognetti (1997). Sites of naturally elevated carbon dioxide. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 1-6.

Gray, J. E. and A. M. Hetherington (2004). Plant development: YODA the stomatal switch. Current Biology **14**(12): R488-R490.

Gray, J. E., G. H. Holroyd, F. M. Van der Lee, A. R. Bahrami, P. C. Sijmons, F. I. Woodward, W. Schuch and A. M. Hetherington (2000). The HIC signalling pathway links $CO_2$ perception to stomatal development. Nature **408**(6813): 713-716.

Griggs, D. J., and M. Noguer (2002). Climate change 2001: the scientific basis. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change. Weather **57**(8): 267-269.

Grill, D., M. Müller, M. Tausz, B. Strnad, A. Wonisch and A. Raschi (2004). Effects of sulphurous gases in two $CO_2$ springs on total sulphur and thiols in acorns and oak seedlings. Atmospheric Environment **38**(23): 3775-3780.

Gupta, P., S. Duplessis, H. White, D. F. Karnosky, F. Martin and G. K. Podila (2005). Gene expression patterns of trembling aspen trees following long-term exposure to interacting elevated $CO_2$ and tropospheric $O_3$. New Phytologist **167**(1): 129-142.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95-98.

Hamilton, D. W. A., A. Hills, B. Köhler and M. R. Blatt (2000). $Ca^{2+}$ channels at the plasma membrane of stomatal guard cells are activated by hyperpolarization and abscisic acid. Proceedings of the National Academy of Sciences **97**(9): 4967-4972.

Han, H. and P. Felker (1997). Field validation of water-use efficiency of the CAM plant *Opuntia ellisianain* south Texas. Journal of Arid Environments **36**(1): 133-148.

Hara, K., R. Kajita, K. U. Torii, D. C. Bergmann and T. Kakimoto (2007). The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. Genes & Development **21**(14): 1720-1725.

Hara, K., T. Yokoo, R. Kajita, T. Onishi, S. Yahata, K. M. Peterson, K. U. Torii and T. Kakimoto (2009). Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR 2 in Arabidopsis leaves. Plant and Cell Physiology **50**(6): 1019-1031.

Harley, P. and T. Sharkey (1991). An improved model of C3 photosynthesis at high $CO_2$: reversed $O_2$ sensitivity explained by lack of glycerate reentry into the chloroplast. Photosynthesis Research **27**(3): 169-178.

Harley, P. C., R. B. Thomas, J. F. Reynolds and B. R. Strain (1992). Modelling photosynthesis of cotton grown in elevated $CO_2$. Plant, Cell & Environment **15**(3): 271-282.

Hashimoto, M., J. Negi, J. Young, M. Israelsson, J. I. Schroeder and K. Iba (2006). Arabidopsis HT1 kinase controls stomatal movements in response to $CO_2$. Nature Cell Biology **8**(4): 391-397.

Hattenschwiler, S., F. Miglietta, A. Raschi and C. Korner (1997). Thirty years of in situ tree growth under elevated $CO_2$: a model for future forest responses? Global Change Biology **3**(5): 463-471.

Havlak, P., R. Chen, K. J. Durbin, A. Egan, Y. Ren, X.-Z. Song, G. M. Weinstock and R. A. Gibbs (2004). The atlas genome assembly system. Genome Research **14**(4): 721-732.

Hepper, F. N. (1990). Pharaoh's flowers. The botanical treasures of Tutankhamun. HMSO.

Herschbach, C., M. Schulte, P. von Ballmoos, C. Brunold and H. Rennenberg (2012). Sulfate and nitrate assimilation in leaves of *Quercus ilex* and *Quercus pubescens* grown near natural $CO_2$

springs in central Italy. Sulfur Metabolism in Plants. L. J. De Kok, L. Tabe, M. Tausz et al., Springer Netherlands. **1:** 237-248.

Hetherington, A. M. and F. I. Woodward (2003). The role of stomata in sensing and driving environmental change. Nature **424**(6951): 901-908.

Himmelbach, A. (1998). Signalling of abscisic acid to regulate plant growth. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **353**(1374): 1439-1444.

Hoekstra, H. E. and J. A. Coyne (2007). The locus of evolution: evo devo and the genetics of adaptation. Evolution **61**(5): 995-1016.

Hofmann, N. R. (2008). They all scream for ICE1/SCRM2: core regulatory units in stomatal development. The Plant Cell Online **20**(7): 1732.

Högy, P., H. Wieser, P. Köhler, K. Schwadorf, J. Breuer, J. Franzaring, R. Muntifering and A. Fangmeier (2009). Effects of elevated $CO_2$ on grain yield and quality of wheat: results from a 3-year free-air $CO_2$ enrichment experiment. Plant Biology **11**: 60-69.

Hosy, E., A. Vavasseur, K. Mouline, I. Dreyer, F. Gaymard, F. Porée, J. Boucherez, A. Lebaudy, D. Bouchez, A. A. Véry, T. Simonneau, J. B. Thibaud and H. Sentenac (2003). The Arabidopsis outward $K^+$ channel GORK is involved in regulation of stomatal movements and plant transpiration. Proceedings of the National Academy of Sciences **100**(9): 5549-5554.

Hu, H., A. Boisson-Dernier, M. Israelsson-Nordstrom, M. Bohmer, S. Xue, A. Ries, J. Godoski, J. M. Kuhn and J. I. Schroeder (2010). Carbonic anhydrases are upstream regulators of $CO_2$-controlled stomatal movements in guard cells. Nature Cell Biology **12**(1): 87-93.

Huang, X., J. Wang, S. Aluru, S. P. Yang and L. Hillier (2003). PCAP: a whole-genome assembly program. Genome Research **13**(9): 2164-2170.

Hudson, R. R., M. Slatkin and W. Maddison (1992). Estimation of levels of gene flow from DNA sequence data. Genetics **132**(2): 583-589.

Huxman, T. E., E. P. Hamerlynck, B. D. Moore, S. D. Smith, D. N. Jordan, S. F. Zitzer, R. S. Nowak, J. S. Coleman and J. R. Seemann (1998). Photosynthetic down-regulation in *Larrea tridentata* exposed to elevated atmospheric $CO_2$: interaction with drought under glasshouse and field (FACE) exposure. Plant, Cell & Environment **21**(11): 1153-1161.

I. Bettarini, F. Miglietta, A. Raschi, R. Tognetti and P. v. Gardingen (1997). Studying morpho-physiological responses of *Scirpus lacustris* from naturally $CO_2$-enriched environments. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 134-147.

Iba, K. (2002). Acclimative response to temperature stress in higher plants: approaches of gene engineering for temperature tolerance. Annual Review of Plant Biology **53**(1): 225-245.

Ineson, P., M. F. Cotrufo, A. Raschi, R. Tognetti and P. v. Gardingen (1997). Increasing concentrations of atmospheric $CO_2$ and decomposition processes in forest ecosystems. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 242-267.

Jablonski, L. M., X. Wang and P. S. Curtis (2002). Plant reproduction under elevated $CO_2$ conditions: a meta-analysis of reports on 79 crop and wild species. New Phytologist **156**(1): 9-26.

Jager, C. E., G. M. Symons, J. J. Ross and J. B. Reid (2008). Do brassinosteroids mediate the water stress response? Physiologia Plantarum **133**(2): 417-425.

Jeukens, J., S. Renaut, J. St-Cyr, A. W. Nolte and L. Bernatchez (2010). The transcriptomics of sympatric dwarf and normal lake whitefish (*Coregonus clupeaformis* spp., Salmonidae) divergence as revealed by next-generation sequencing. Molecular Ecology **19**(24): 5389-5403.

Jewer, P. C. and L. D. Incoll (1980). Promotion of stomatal opening in the grass *Anthephora pubescens* nees by a range of natural and synthetic cytokinis. Planta **150**(3): 218-221.

Johnson, H., H. W. Polley and H. Mayeux (1993). Increasing $CO_2$ and plant-plant interactions: effects on natural vegetation. $CO_2$ and Biosphere. J. Rozema, H. Lambers, S. C. Geijn and M. L. Cambridge, Springer Netherlands. **14:** 157-170.

Johnson J.D., Michelozzi M., Tognetti R., Raschi A., Miglietta F., Tognetti R. and V. G. P. (1997). Carbon physiology of *Quercus pubescens* Wild. growing at the Bossoleto $CO_2$ spring in central Italy. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 148-164.

Jolliffe, I. (2005). Principal component analysis. Encyclopedia of Statistics in Behavioral Science, John Wiley & Sons, Ltd.

Jones, M. B., J. C. Brown, A. Raschi and F. Miglietta (1995). The effects on *Arbutus unedo* L. of long-term exposure to elevated $CO_2$. Global Change Biology **1**(4): 295-302.

Kanaoka, M. M., L. J. Pillitteri, H. Fujii, Y. Yoshida, N. L. Bogenschutz, J. Takabayashi, J.-K. Zhu and K. U. Torii (2008). "SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to Arabidopsis stomatal differentiation. The Plant Cell Online **20**(7): 1775-1785.

Kaplan, F., W. Zhao, J. T. Richards, R. M. Wheeler, C. L. Guy and L. H. Levine (2012). Transcriptional and metabolic insights into the differential physiological responses of Arabidopsis to optimal and supraoptimal atmospheric $CO_2$. PLoS ONE **7**(8): e43583.

Kaschuk, G., M. Hungria, P. A. Leffelaar, K. E. Giller and T. W. Kuyper (2010). Differences in photosynthetic behaviour and leaf senescence of soybean (*Glycine max* [L.] Merrill) dependent on N2 fixation or nitrate supply. Plant Biology **12**(1): 60-69.

Kass, S. U., D. Pruss and A. P. Wolffe (1997). How does DNA methylation repress transcription? Trends in Genetics **13**(11): 444-449.

Kellogg, E. A. and N. D. Juliano (1997). The structure and function of RuBisCo and their implications for systematic studies. American Journal of Botany **84**(3): 413-428.

Kellomäki, S., K.-Y. Wang and M. Lemettinen (2000). Controlled environment chambers for investigating tree response to elevated $CO_2$ and temperature under boreal conditions. Photosynthetica **38**(1): 69-81.

Kepinski, S. (2007). The anatomy of auxin perception. BioEssays **29**(10): 953-956.

Kim, T.-H., M. Böhmer, H. Hu, N. Nishimura and J. I. Schroeder (2010). Guard cell signal transduction network: advances in understanding abscisic acid, $CO_2$, and $Ca^{2+}$ signaling. Annual Review of Plant Biology **61**(1): 561-591.

Kimura, M. (1969). The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. Genetics **61**(4): 893-903.

Kimura, M. (1984). The neutral theory of molecular evolution. Cambridge University Press.

King, M. C. and A. Wilson (2014). Evolution at two levels in humans and chimpanzees. Essential Readings in Evolutionary Biology **188**(4184): 301.

Klein, R. R. and J. E. Mullet (1986). Regulation of chloroplast-encoded chlorophyll-binding protein translation during higher plant chloroplast biogenesis. Journal of Biological Chemistry **261**(24): 11138-11145.

Klus, D. J., S. Kalisz, P. S. Curtis, J. A. Teeri and S. J. Tonsor (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**(6): 1080-1087.

Knapp, A. K., M. Cocke, E. P. Hamerlynck and C. E. Owensby (1994). Effect of elevated $CO_2$ on stomatal density and distribution in a C4 grass and a C3 forb under field conditions. Annuals of Botany **74**(6): 595-599.

Knoll, A. H., R. K. Bambach, J. L. Payne, S. Pruss and W. W. Fischer (2007). Paleophysiology and end-permian mass extinction. Earth and Planetary Science Letters **256**(3–4): 295-313.

Koch, K. E. (1996). Carbohydrate-modulated gene expression in plants. Annual Review of Plant Physiology and Plant Molecular Biology **47**(1): 509-540.

Kölliker, R., S. Bassin, D. Schneider, F. Widmer and J. Fuhrer (2008). Elevated ozone affects the genetic composition of *Plantago lanceolata* L. populations. Environmental Pollution **152**(2): 380-386.

Konrad, W., A. Roth-Nebelsick and M. Grein (2008). Modelling of stomatal density response to atmospheric $CO_2$. Journal of Theoretical Biology **253**(4): 638-658.

Körner, C., R. Asshoff, O. Bignucolo, S. Hättenschwiler, S. G. Keel, S. Peláez-Riedl, S. Pepin, R. T. W. Siegwolf and G. Zotz (2005). Carbon flux and growth in mature deciduous forest trees exposed to elevated $CO_2$. Science **309**(5739): 1360-1362.

Körner, C. and F. Miglietta (1994). Long term effects of naturally elevated $CO_2$; on mediterranean grassland and forest trees. Oecologia **99**(3): 343-351.

Kouwenberg, L. L. R., J. C. McElwain, W. M. Kürschner, F. Wagner, D. J. Beerling, F. E. Mayle and H. Visscher (2003). Stomatal frequency adjustment of four conifer species to historical changes in atmospheric $CO_2$. American Journal of Botany **90**(4): 610-619.

Kramer, P. J. (1981). Carbon dioxide concentration, photosynthesis, and dry matter production. BioScience **31**(1): 29-33.

Krishnan S, G., D. L. E. Waters and R. J. Henry (2014). Australian wild rice reveals pre-domestication origin of polymorphism deserts in rice genome. PLoS ONE **9**(6): e98843.

Kubiske, M. E. and K. S. Pregitzer (1996). Effects of elevated $CO_2$ and light availability on the photosynthetic light response of trees of contrasting shade tolerance. Tree Physiology **16**(3): 351-358.

Kwak, J. M., Y. Murata, V. M. Baizabal-Aguirre, J. Merrill, M. Wang, A. Kemper, S. D. Hawke, G. Tallman and J. I. Schroeder (2001). Dominant negative guard cell $K^+$ channel mutants reduce inward-rectifying $K^+$ currents and light-induced stomatal opening in Arabidopsis. Plant Physiology **127**(2): 473-485.

Lai, L. B., J. A. Nadeau, J. Lucas, E.-K. Lee, T. Nakagawa, L. Zhao, M. Geisler and F. D. Sack (2005). The Arabidopsis R2R3 MYB proteins FOUR LIPS and MYB88 restrict divisions late in the stomatal cell lineage. The Plant Cell Online **17**(10): 2754-2767.

Laurie, C. C. and L. F. Stam (1994). The effect of an intronic polymorphism on alcohol dehydrogenase expression in *Drosophila melanogaster*. Genetics **138**(2): 379-385.

Le, J., X. G. Liu, K. Z. Yang, X. L. Chen, J.-J. Zou, H. Z. Wang, M. Wang, S. Vanneste, M. Morita, M. Tasaka, Z.-J. Ding, J. Friml, T. Beeckman and F. Sack (2014). Auxin transport and activity regulate stomatal patterning and development. Nature Communications **5**.

Leakey, A. D. B., E. A. Ainsworth, C. J. Bernacchi, A. Rogers, S. P. Long and D. R. Ort (2009). Elevated $CO_2$ effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. Journal of Experimental Botany **60**(10): 2859-2876.

Leakey, A. D. B., C. J. Bernacchi, D. R. Ort and S. P. Long (2006). Long-term growth of soybean at elevated [$CO_2$] does not cause acclimation of stomatal conductance under fully open-air conditions. Plant, Cell & Environment **29**(9): 1794-1800.

Leakey, A. D. B., F. Xu, K. M. Gillespie, J. M. McGrath, E. A. Ainsworth and D. R. Ort (2009). Genomic basis for stimulated respiration by plants growing under elevated carbon dioxide. Proceedings of the National Academy of Sciences.

Leakey, A. D. B., F. Xu, K. M. Gillespie, J. M. McGrath, E. A. Ainsworth and D. R. Ort (2009). Genomic basis for stimulated respiration by plants growing under elevated carbon dioxide. Proceedings of the National Academy of Sciences **106**(9): 3597-3602.

Lebaudy, A., A. A. Véry and H. Sentenac (2007). $K^+$ channel activity in plants: genes, regulations and functions. FEBS Letters **581**(12): 2357-2366.

LeCain, D. R. and J. A. Morgan (1998). Growth, gas exchange, leaf nitrogen and carbohydrate concentrations in NAD-ME and NADP-ME C4 grasses grown in elevated $CO_2$. Physiologia Plantarum **102**(2): 297-306.

Lee, M., Y. Choi, B. Burla, Y. Y. Kim, B. Jeon, M. Maeshima, J. Y. Yoo, E. Martinoia and Y. Lee (2008). The ABC transporter AtABCB14 is a malate importer and modulates stomatal response to $CO_2$. Nature Cell Biology **10**(10): 1217-1223.

Lenz, K. E., G. E. Host, K. Roskoski, A. Noormets, A. Sôber and D. F. Karnosky (2010). Analysis of a Farquhar-von Caemmerer-Berry leaf-level photosynthetic rate model for Populus tremuloides in the context of modeling and measurement limitations. Environmental Pollution **158**(4): 1015-1022.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and G. P. D. P. Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**(16): 2078-2079.

Li, J., H. Jiang and W. Wong (2010). Modeling non-uniformity in short-read rates in RNA-Seq data. Genome Biology **11**(5): R50.

Li, J. And R. Tibshirani (2011). Finding consistent patterns: a nonparametric approach for identifying differential expression in RNA-Seq data. To appear, Statistical Methods in Medical Research.

Li, P., E. A. Ainsworth, A. D. B. Leakey, A. Ulanov, V. Lozovaya, D. R. Ort and H. J. Bohnert (2008). Arabidopsis transcript and metabolite profiles: ecotype-specific responses to open-air elevated [$CO_2$]. Plant, Cell & Environment **31**(11): 1673-1687.

Li, P., H. J. Bohnert and R. Grene (2007). All about FACE – plants in a high-[$CO_2$] world. Trends in Plant Science **12**(3): 87-89.

Li, P., A. Sioson, S. Mane, A. Ulanov, G. Grothaus, L. Heath, T. M. Murali, H. Bohnert and R. Grene (2006). Response diversity of *Arabidopsis thaliana* ecotypes in elevated [$CO_2$] in the field. Plant Molecular Biology **62**(4-5): 593-609.

Li, R., W. Fan, G. Tian, H. Zhu, L. He, J. Cai, Q. Huang, Q. Cai, B. Li, Y. Bai, Z. Zhang, Y. Zhang, W. Wang, J. Li, F. Wei, H. Li, M. Jian, J. Li, Z. Zhang, R. Nielsen, D. Li, W. Gu, Z. Yang, Z. Xuan, O. A. Ryder, F. C.-C. Leung, Y. Zhou, J. Cao, X. Sun, Y. Fu, X. Fang, X. Guo, B. Wang, R. Hou, F. Shen, B. Mu, P. Ni, R. Lin, W. Qian, G. Wang, C. Yu, W. Nie, J. Wang, Z. Wu, H. Liang, J. Min, Q. Wu, S. Cheng, J. Ruan, M. Wang, Z. Shi, M. Wen, B. Liu, X. Ren, H. Zheng, D. Dong, K. Cook, G. Shan, H. Zhang, C. Kosiol, X. Xie, Z. Lu, H. Zheng, Y. Li, C. C. Steiner, T. T.-Y. Lam, S. Lin, Q. Zhang, G. Li, J. Tian, T. Gong, H. Liu, D. Zhang, L. Fang, C. Ye, J. Zhang, W. Hu, A. Xu, Y. Ren, G. Zhang, M. W. Bruford, Q. Li, L. Ma, Y. Guo, N. An, Y. Hu, Y. Zheng, Y. Shi, Z. Li, Q. Liu, Y. Chen, J. Zhao, N. Qu, S. Zhao, F. Tian, X. Wang, H. Wang, L. Xu, X. Liu, T. Vinar, Y. Wang, T.-W. Lam, S.-M. Yiu, S. Liu, H. Zhang, D. Li, Y. Huang, X. Wang, G. Yang, Z. Jiang, J. Wang, N. Qin, L. Li, J. Li, L. Bolund, K. Kristiansen, G. K.-S. Wong, M. Olson, X. Zhang, S. Li, H. Yang, J. Wang and J. Wang (2010). The sequence and de novo assembly of the giant panda genome. Nature **463**(7279): 311-317.

Lieffering, M., H.-Y. Kim, K. Kobayashi and M. Okada (2004). The impact of elevated $CO_2$ on the elemental concentrations of field-grown rice grains. Field Crops Research **88**(2–3): 279-286.

Linder, B. and K. Raschke (1992). A slow anion channel in guard cells, activating at large hyperpolarization, may be principal for stomatal closing. FEBS Letters **313**(1): 27-30.

Liu, W. and D. A. Saint (2002). A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. Analytical Biochemistry **302**(1): 52-59.

Logemann, J., J. Schell and L. Willmitzer (1987). Improved method for the isolation of RNA from plant tissues. Analytical Biochemistry **163**(1): 16-20.

Lohse, G. and R. Hedrich (1992). Characterization of the plasma-membrane H+-ATPase from Vicia faba guard cells. Planta **188**(2): 206-214.

Long, S. P. (1991). Modification of the response of photosynthetic productivity to rising temperature by atmospheric $CO_2$ concentrations: has its importance been underestimated? Plant, Cell & Environment **14**(8): 729-739.

Long, S. P., E. A. Ainsworth, A. Rogers and D. R. Ort (2004). Rising atmospheric carbon dioxide: plants FACE the future. Annual Review of Plant Biology **55**: 591-628.

Long, S. P., N. R. Baker and C. A. Raines (1993). Analysing the responses of photosynthetic $CO_2$ assimilation to long-term elevation of atmospheric $CO_2$ concentration. Vegetation **104-105**(1): 33-45.

Long, S. P. and B. G. Drake (1991). Effect of the long-term elevation of $CO_2$ concentration in the field on the quantum yield of photosynthesis of the C(3) sedge, *Scirpus olneyi*. Plant Physiology **96**(1): 221-226.

Lu, Z., J. W. Radin, E. L. Turcotte, R. Percy and E. Zeiger (1994). High yields in advanced lines of Pima cotton are associated with higher stomatal conductance, reduced leaf area and lower leaf temperature. Physiologia Plantarum **92**(2): 266-272.

Lundqvist, E. and E. Andersson (2001). Genetic diversity in populations of plants with different breeding and dispersal strategies in a free-flowing boreal river system. Hereditas **135**(1): 75-83.

Mousseau M., Z. H. Enoch, J. C. Sabroux, A. Raschi, R. Tognetti and P. v. Gardingen (1997). Controlled degassing of lakes with high $CO_2$ content in Cameroon: an opportunity for ecosystem $CO_2$-enrichment experiments. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 45-55.

MacAlister, C. A., K. Ohashi-Ito and D. C. Bergmann (2007). Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. Nature **445**(7127): 537-540.

MacRobbie, E. A. C. (1998). Signal transduction and ion channels in guard cells. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **353**(1374): 1475-1488.

Mahieu, S., S. Soussou, J.-C. Cleyet-Marel, B. Brunel, L. Mauré, C. Lefèbvre and J. Escarré (2012). Local adaptation of metallicolous and non-metallicolous *Anthyllis vulneraria* populations: their utilization in soil restoration. Restoration Ecology **21**(5): 551-559.

Makino, A. and T. Mae (1999). Photosynthesis and plant growth at elevated levels of $CO_2$. Plant and Cell Physiology **40**(10): 999-1006.

Malone, S. R., H. S. Mayeux, H. B. Johnson and H. W. Polley (1993). Stomatal density and aperture length in four plant species grown across a subambient $CO_2$ gradient. American Journal of Botany **80**(12): 1413-1418.

Manderscheid, R., J. Bender, H. J. Jäger and H. J. Weigel (1995). Effects of season long $CO_2$ enrichment on cereals II. Nutrient concentrations and grain quality. Agriculture, Ecosystems & Environment **54**(3): 175-185.

Manter, D. K. and J. Kerrigan (2004). A/Ci curve analysis across a range of woody plant species: influence of regression analysis parameters and mesophyll conductance. Journal of Experimental Botany **55**(408): 2581-2588.

Marchi, S., R. Tognetti, F. P. Vaccari, M. Lanini, M. Kaligarič, F. Miglietta and A. Raschi (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**(2): 181-194.

Marioni, J. C., C. E. Mason, S. M. Mane, M. Stephens and Y. Gilad (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Research **18**(9): 1509-1517.

Markelz, R. J. C., L. N. Vosseller and A. D. B. Leakey (2014). Developmental stage specificity of transcriptional, biochemical and $CO_2$ efflux responses of leaf dark respiration to growth of *Arabidopsis thaliana* at elevated [$CO_2$]. Plant, Cell & Environment **37**(11): 2542-2552.

Masle, J., S. R. Gilmore and G. D. Farquhar (2005). The ERECTA gene regulates plant transpiration efficiency in Arabidopsis. Nature **436**(7052): 866-870.

Matros, A., S. Amme, B. Kettig, G. H. Buck-Sorlin, U. W. E. Sonnewald and H.-P. Mock (2006). Growth at elevated $CO_2$ concentrations leads to modified profiles of secondary metabolites in tobacco cv. SamsunNN and to increased resistance against infection with potato virus Y. Plant, Cell & Environment **29**(1): 126-137.

Maurer-Stroh, S., N. J. Dickens, L. Hughes-Davies, T. Kouzarides, F. Eisenhaber and C. P. Ponting (2003). The tudor domain 'Royal Family': tudor, plant agenet, chromo, PWWP and MBT domains. Trends in Biochemical Sciences **28**(2): 69-74.

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

McElwain, J., F. J. G. Mitchell and M. B. Jones (1995). Relationship of stomatal density and index of *Salix cinerea* to atmospheric carbon dioxide concentrations in the Holocene. Holocene **5**(2): 216-219.

McElwain, J. C. and W. G. Chaloner (1995). Stomatal density and index of fossil plants track atmospheric carbon dioxide in the Palaeozoic. Annals of Botany **76**(4): 389-395.

McGrath, J. M. and D. B. Lobell (2013). Reduction of transpiration and altered nutrient allocation contribute to nutrient decline of crops grown in elevated $CO_2$ concentrations. Plant, Cell & Environment **36**(3): 697-705.

Medlyn, B. E., F. W. Badeck, D. G. G. De Pury, C. V. M. Barton, M. Broadmeadow, R. Ceulemans, P. De Angelis, M. Forstreuter, M. E. Jach, S. Kellomäki, E. Laitat, M. Marek, S. Philippot, A. Rey, J. Strassemeyer, K. Laitinen, R. Liozon, B. Portier, P. Roberntz, K. Wang and P. G. Jstbid (1999).

Effects of elevated [$CO_2$] on photosynthesis in European forest species: a meta-analysis of model parameters. Plant, Cell & Environment **22**(12): 1475-1495.

Medlyn, B. E., C. V. M. Barton, M. S. J. Broadmeadow, R. Ceulemans, P. De Angelis, M. Forstreuter, M. Freeman, S. B. Jackson, S. Kellomäki, E. Laitat, A. Rey, P. Roberntz, B. D. Sigurdsson, J. Strassemeyer, K. Wang, P. S. Curtis and P. G. Jarvis (2001). Stomatal conductance of forest species after long-term exposure to elevated $CO_2$ concentration: a synthesis. New Phytologist **149**(2): 247-264.

Meloni, M. and G. Testa (2014). Scrutinizing the epigenetics revolution. BioSocieties **9**(4): 431-456.

Melotto, M., W. Underwood and S. Y. He (2008). Role of stomata in plant innate immunity and foliar bacterial diseases. Annual Review of Phytopathology **46**(1): 101-122.

Melotto, M., W. Underwood, J. Koczan, K. Nomura and S. Y. He (2006). Plant stomata function in innate immunity against bacterial invasion. Cell **126**(5): 969-980.

Miglietta, F., Raschi, A., Resti, R. and Badiani, M. (1993). Growth and onto-morphogenesis of soybean (*Glycine max* Merril) in an open, naturally $CO_2$-enriched environment. Plant, Cell & Environment (16): 909-918.

Miglietta, F., A. Raschi, I. Bettarini, R. Resti and F. Selvi (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**(7): 873-878.

Miglietta, F., A. Raschi, C. Körner and F. P. Vaccari (1998). Isotope discrimination and photosynthesis of vegetation growing in the Bossoleto $CO_2$ spring. Chemosphere **36**(4–5): 771-776.

Mishra, G., W. Zhang, F. Deng, J. Zhao and X. Wang (2006). A bifurcating pathway directs abscisic acid effects on stomatal closure and opening in Arabidopsis. Science **312**(5771): 264-266.

Mitchell, R. A. C., V. J. Mitchell, S. P. Driscoll, J. Franklin and D. W. Lawlor (1993). Effects of increased $CO_2$ concentration and temperature on growth and yield of winter wheat at two levels of nitrogen application. Plant, Cell & Environment **16**(5): 521-529.

Monje, O. and B. Bugbee (1998). Adaptation to high $CO_2$ concentration in an optimal environment: radiation capture, canopy quantum yield and carbon use efficiency. Plant, Cell & Environment **21**(3): 315-324.

Moore, B. D., S. H. Cheng, D. Sims and J. R. Seemann (1999). The biochemical and molecular basis for photosynthetic acclimation to elevated atmospheric $CO_2$. Plant, Cell & Environment **22**(6): 567-582.

Morison, J. I. L. (1998). Stomatal response to increased $CO_2$ concentration. Journal of Experimental Botany **49**(Special Issue): 443-452.

Mortazavi, A., B. A. Williams, K. McCue, L. Schaeffer and B. Wold (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature Methods **5**(7): 621-628.

Munemasa, S., K. Oda, M. Watanabe-Sugimoto, Y. Nakamura, Y. Shimoishi and Y. Murata (2007). The coronatine-insensitive 1 mutation reveals the hormonal signaling interaction between

abscisic acid and methyl jasmonate in Arabidopsis guard cells. Specific impairment of ion channel activation and second messenger production. Plant Physiology **143**(3): 1398-1407.

Myers, S. S., A. Zanobetti, I. Kloog, P. Huybers, A. D. B. Leakey, A. J. Bloom, E. Carlisle, L. H. Dietterich, G. Fitzgerald, T. Hasegawa, N. M. Holbrook, R. L. Nelson, M. J. Ottman, V. Raboy, H. Sakai, K. A. Sartor, J. Schwartz, S. Seneweera, M. Tausz and Y. Usui (2014). Increasing $CO_2$ threatens human nutrition. Nature **510**(7503): 139-142.

Nabholz, B., G. Sarah, F. Sabot, M. Ruiz, H. Adam, S. Nidelet, A. Ghesquière, S. Santoni, J. David and S. Glémin (2014). Transcriptome population genomics reveals severe bottleneck and domestication cost in the African rice (*Oryza glaberrima*). Molecular Ecology **23**(9): 2210-2227.

Nagalakshmi, U., Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein and M. Snyder (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. Science **320**(5881): 1344-1349.

Narum, S. R. and J. E. Hess (2011). Comparison of FST outlier tests for SNP loci under selection. Molecular Ecology Resources **11**: 184-194.

Naudts, K., J. Van den Berge, E. Farfan, P. Rose, H. AbdElgawad, R. Ceulemans, I. A. Janssens, H. Asard and I. Nijs (2014). Future climate alleviates stress impact on grassland productivity through altered antioxidant capacity. Environmental and Experimental Botany **99**(0): 150-158.

Negi, J., O. Matsuda, T. Nagasawa, Y. Oba, H. Takahashi, M. Kawai-Yamada, H. Uchimiya, M. Hashimoto and K. Iba (2008). $CO_2$ regulator SLAC1 and its homologues are essential for anion homeostasis in plant cells. Nature **452**(7186): 483-486.

Nei, M. and W. H. Li (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences **76**(10): 5269-5273.

Newingham, B. A., C. H. Vanier, T. N. Charlet, K. Ogle, S. D. Smith and R. S. Nowak (2013). No cumulative effect of 10 years of elevated [$CO_2$] on perennial plant biomass components in the Mojave desert. Global Change Biology **19**(7): 2168-2181.

Nijs, I., R. Ferris, H. Blum, G. Hendrey and I. Impens (1997). Stomatal regulation in a changing climate: a field study using Free Air Temperature Increase (FATI) and Free Air $CO_2$ Enrichment (FACE). Plant, Cell & Environment **20**(8): 1041-1050.

Nimmo, H. G. (2000). The regulation of phosphoenolpyruvate carboxylase in CAM plants. Trends in Plant Science **5**(2): 75-80.

Ohashi-Ito, K. and D. C. Bergmann (2006). Arabidopsis FAMA controls the final proliferation/differentiation switch during stomatal development. The Plant Cell Online **18**(10): 2493-2505.

Oleksyk, T. K., M. W. Smith and S. J. O'Brien (2010). Genome-wide scans for footprints of natural selection. Philisophical Transanctions of the Royal Soceity of Biology **365** (10): 185-205.

Onoda, Y., T. Hirose and K. Hikosaka (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**(3): 475-484.

Onoda, Y., T. Hirose and K. Hikosaka (2009). Does leaf photosynthesis adapt to $CO_2$-enriched environments? An experiment on plants originating from three natural $CO_2$ springs. New Phytologist **182**(3): 698-709.

Osada, N., Y. Onoda and K. Hikosaka (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**(1): 41-52.

Paoletti, E., F. Miglietta, A. Raschi, R. Tognetti, F. Manes, P. Grossoni and P. v. Gardingen (1997). Stomatal numbers in holm oak (*Quercus ilex* L.) leaves grown in naturally and artificially $CO_2$-enriched environments. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 197-208.

Paul, M. J. and C. H. Foyer (2001). Sink regulation of photosynthesis. Journal of Experimental Botany **52**(360): 1383-1400.

Pearcy, R., E. D. Schulze and R. Zimmermann (2000). Measurement of transpiration and leaf conductance. Plant Physiological Ecology Springer Netherlands**:** 137-160.

Penninckx, I. A., K. Eggermont, F. R. Terras, B. P. Thomma, G. W. De Samblanx, A. Buchala, J. P. Métraux, J. M. Manners and W. F. Broekaert (1996). Pathogen-induced systemic activation of a plant defensin gene in Arabidopsis follows a salicylic acid-independent pathway. The Plant Cell Online **8**(12): 2309-2323.

Penuelas, J. and R. Matamala (1990). Changes in N and S leaf content, stomatal density and specific leaf area of 14 plant species during the last three centuries of $CO_2$ increase. Journal of Experimental Botany **41**(9): 1119-1124.

Pérez-López, U., A. Robredo, M. Lacuesta, C. Sgherri, A. Muñoz-Rueda, F. Navari-Izzo and A. Mena-Petite (2009). The oxidative stress caused by salinity in two barley cultivars is mitigated by elevated $CO_2$. Physiologia Plantarum **135**(1): 29-42.

Peterson, C. L. and M.-A. Laniel (2004). Histones and histone modifications. Current Biology **14**(14): R546-R551.

Pettersson, R., H. S. J. Lee and P. G. Jarvis (1993). The effect of $CO_2$ concentration on barley. $CO_2$ and biosphere Springer Netherlands. **14:** 462-463.

Pettersson, R. and A. J. McDonald (1994). Effects of nitrogen supply on the acclimation of photosynthesis to elevated $CO_2$. Photosynthesis Research **39**(3): 389-400.

Pevzner, P. A. and H. Tang (2001). Fragment assembly with double-barreled data. Bioinformatics **17**(suppl 1): S225-S233.

Pevzner, P. A., H. Tang and G. Tesler (2004). De novo repeat classification and fragment assembly. Genome Research **14**(9): 1786-1796.

Pevzner, P. A., H. Tang and M. S. Waterman (2001). An Eulerian path approach to DNA fragment assembly. Proceedings of the National Academy of Sciences **98**(17): 9748-9753.

Phillips, T. (2008). The role of methylation in gene expression. Nature Education **1**(1): 116.

Pillitteri, L. J. and J. Dong (2013). Stomatal development in Arabidopsis. The Arabidopsis Book/American Society of Plant Biologists: **1**.

Pillitteri, L. J., D. B. Sloan, N. L. Bogenschutz and K. U. Torii (2007). Termination of asymmetric cell division and differentiation of stomata. Nature **445**(7127): 501-505.

Pleijel, H., J. Gelang, E. Sild, H. Danielsson, S. Younis, P.-E. Karlsson, G. Wallin, L. Skärby and G. Selldén (2000). Effects of elevated carbon dioxide, ozone and water availability on spring wheat growth and yield. Physiologia Plantarum **108**(1): 61-70.

Pohjanvirta, R., M. Niittynen, J. Lindén, P. C. Boutros, I. D. Moffat and A. B. Okey (2006). Evaluation of various housekeeping genes for their applicability for normalization of mRNA expression in dioxin-treated rats. Chemico-Biological Interactions **160**(2): 134-149.

Polley, H. W., H. B. Johnson, H. S. Mayeux and S. R. Malone (1993). Physiology and growth of wheat across a subambient carbon-dioxide gradient. Annuals of Botany **71**(4): 347-356.

Prior, S. A., G. B. Runion, H. H. Rogers and H. A. Torbert (2008). Effects of atmospheric $CO_2$ enrichment on crop nutrient dynamics under no-till conditions. Journal of Plant Nutrition **31**(4): 758-773.

Rae, A. M., P. J. Tricker, S. M. Bunn and G. Taylor (2007). Adaptation of tree growth to elevated $CO_2$: quantitative trait loci for biomass in Populus. New Phytologist **175**(1): 59-69.

Raines, C. A. (2003). The Calvin cycle revisited. Photosynthesis Research **75**(1): 1-10.

Rajasekaran, L. R. and T. J. Blake (1999). New plant growth regulators protect photosynthesis and enhance growth under drought of jack pine seedlings. Journal of Plant Growth Regulation **18**(4): 175-181.

Raven, J. A. (2014). Speedy small stomata? Journal of Experimental Botany eru032.

Regier, N. and B. Frey (2010). Experimental comparison of relative RT-qPCR quantification approaches for gene expression studies in poplar. BMC Molecular Biology **11**(1): 57.

Reich, P. B. and S. E. Hobbie (2013). Decade-long soil nitrogen constraint on the $CO_2$ fertilization of plant biomass. Nature Climate Change **3**(3): 278-282.

Renger, G. and T. Renger (2008). Photosystem II: The machinery of photosynthetic water splitting. Photosynthesis Research **98**(1-3): 53-80.

Rey, A. and P. G. Jarvis (1998). Long-term photosynthetic acclimation to increased atmospheric $CO_2$ concentration in young birch (*Betula pendula*) trees. Tree Physiology **18**(7): 441-450.

Roach, D. A. (2003). Age-specific demography in Plantago: variation among cohorts in a natural plant population. Ecology **84**(3): 749-756.

Roberts, A., C. Trapnell, J. Donaghey, J. Rinn and L. Pachter (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biology **12**(3): R22.

Robinson, E. A., G. D. Ryan and J. A. Newman (2012). A meta-analytical review of the effects of elevated $CO_2$ on plant–arthropod interactions highlights the importance of interacting environmental and biological variables. New Phytologist **194**(2): 321-336.

Robinson, M. and A. Oshlack (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology **11**(3): R25.

Robinson, M. D., D. J. McCarthy and G. K. Smyth (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26**(1): 139-140.

Rogers, A. and S. W. Humphries (2000). A mechanistic evaluation of photosynthetic acclimation at elevated $CO_2$. Global Change Biology **6**(8): 1005-1011.

Rogers, H. H., C. M. Peterson, J. N. McCrimmon and J. D. Cure (1992). Response of plant roots to elevated atmospheric carbon dioxide. Plant, Cell & Environment **15**(6): 749-752.

Rotter, A., C. Camps, M. Lohse, C. Kappel, S. Pilati, M. Hren, M. Stitt, P. Coutos-Thevenot, C. Moser, B. Usadel, S. Delrot and K. Gruden (2009). Gene expression profiling in susceptible interaction of grapevine with its fungal pathogen *Eutypa lata*: extending MapMan ontology for grapevine. BMC Plant Biology **9**(1): 104.

Rowley, J. W., A. J. Oler, N. D. Tolley, B. N. Hunter, E. N. Low, D. A. Nix, C. C. Yost, G. A. Zimmerman and A. S. Weyrich (2011). Genome-wide RNA-seq analysis of human and mouse platelet transcriptomes. Blood **118**(14): e101-e111.

Royer, D. L. (2001). Stomatal density and stomatal index as indicators of paleoatmospheric $CO_2$ concentration. Review of Palaeobotany and Palynology **114**(1–2): 1-28.

Ryle, G. J. A. and J. Stanley (1992). Effect of elevated $CO_2$ on stomatal size and distribution in perenial ryegrass. Annals of Botany **69**(6): 563-565.

Sachs, T. (2005). Pattern formation in plant tissues. Cambridge University Press: **25**.

Sagar, G. R. and J. L. Harper (1964). *Plantago Major* L., *P. Media* L. and *P. Lanceolata* L. Journal of Ecology **52**(1): 189-221.

Sage, R. (1994). Acclimation of photosynthesis to increasing atmospheric $CO_2$: The gas exchange perspective. Photosynthesis Research **39**(3): 351-368.

Sage, R., T. Sharkey and R. Pearcy (1990). The effect of leaf nitrogen and temperature on the $CO_2$ response of photosynthesis in the C3 Dicot *Chenopodium album* L. Functional Plant Biology **17**(2): 135-148.

Sage, R. F. (1995). Was low atmospheric $CO_2$ during the Pleistocene a limiting factor for the origin of agriculture? Global Change Biology **1**(2): 93-106.

Sage, R. F. and C. D. Reid (1992). Photosynthetic acclimation to sub-ambient $CO_2$ (20 PA) in the C3 annual *Phaseolus Vulgaris* L. Photosynthetica **27**(4): 605-617.

Salisbury, E. J. (1928). On the causes and ecological significance of stomatal frequency, with special reference to the woodland flora. Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character **216**: 1-65.

Salzberg, S. L., A. M. Phillippy, A. Zimin, D. Puiu, T. Magoc, S. Koren, T. J. Treangen, M. C. Schatz, A. L. Delcher, M. Roberts, G. Marçais, M. Pop and J. A. Yorke (2012). GAGE: A critical evaluation of genome assemblies and assembly algorithms. Genome Research **22**(3): 557-567.

Sanger, F. and A. R. Coulson (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology **94**(3): 441-448.

Santakumari, M. and R. A. Fletcher (1987). Reversal of triazole-induced stomatal closure by gibberellic acid and cytokinins in Commelina benghalensis. Physiologia Plantarum **71**(1): 95-99.

Saxe, H., D. S. Ellsworth and J. Heath (1998). Tree and forest functioning in an enriched $CO_2$ atmosphere. New Phytologist **139**(3): 395-436.

Scheet, P. and M. Stephens (2006). A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase The American Journal of Human Genetics **78**(4): 629-644.

Schlüter, U., M. Muschak, D. Berger and T. Altmann (2003). Photosynthetic performance of an Arabidopsis mutant with elevated stomatal density (sdd1-1) under different light regimes. Journal of Experimental Botany **54**(383): 867-874.

Schneider, C. A., W. S. Rasband and K. W. Eliceiri (2012). NIH image to ImageJ: 25 years of image analysis. Nature Methods **9**(7): 671-675.

Scholefield, P. A., K. J. Doick, B. M. J. Herbert, C. N. S. Hewitt, J. P. Schnitzler, P. Pinelli and F. Loreto (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**(4): 393-401.

Schroeder, J. I. and S. Hagiwara (1989). Cytosolic calcium regulates ion channels in the plasma membrane of *Vicia faba* guard cells. Nature **338**(6214): 427-430.

Schulte M., F. G. Raiesi, H. Papke, K. Butterbach-Bahl, N. Van Breemen and H. Rennenberg (1997). $CO_2$ concentration and atmospheric trace gas mixing ratio around natural $CO_2$ vents in different Mediterranean forests in central Italy. Ecosystem Responses To $CO_2$: The MAPLE Project Results: 168-188.

Schwartz, A., W.-H. Wu, E. B. Tucker and S. M. Assmann (1994). Inhibition of inward $K^+$ channels and stomatal response by abscisic acid: an intracellular locus of phytohormone action. Proceedings of the National Academy of Sciences **91**(9): 4019-4023.

Seemann, J. R. and T. D. Sharkey (1986). Salinity and nitrogen effects on photosynthesis, ribulose-1,5-bisphosphate carboxylase and metabolite pool sizes in *Phaseolus vulgaris* L." Plant Physiology **82**(2): 555-560.

Selvi, F. B. (1999). Geothermal biotopes in central-western Italy from a botanical view point. Ecosystem Responses To $CO_2$: The MAPLE Project Results**:** 1-12.

Seneweera, S. and J. Conroy (1997). Growth, grain yield and quality of rice (*Oryza sativa* L.) in response to elevated $CO_2$ and phosphorus nutrition. Plant Nutrition for Sustainable Food Production and Environment. T. Ando, K. Fujita, T. Mae et al., Springer Netherlands. **78:** 873-878.

Shapiro, M. D., M. E. Marks, C. L. Peichel, B. K. Blackman, K. S. Nereng, B. Jonsson, D. Schluter and D. M. Kingsley (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature **428**(6984): 717-723.

Sharkey, T. D., C. J. Bernacchi, G. D. Farquhar and E. L. Singsaas (2007). Fitting photosynthetic carbon dioxide response curves for C3 leaves. Plant, Cell & Environment **30**(9): 1035-1040.

Shipley, B. and D. Meziane (2002). The balanced-growth hypothesis and the allometry of leaf and root biomass allocation. Functional Ecology **16**(3): 326-331.

Shpak, E. D., J. M. McAbee, L. J. Pillitteri and K. U. Torii (2005). Stomatal patterning and differentiation by synergistic interactions of receptor kinases. Science **309**(5732): 290-293.

Sigman, D. M. and E. A. Boyle (2000). Glacial/interglacial variations in atmospheric carbon dioxide. Nature **407**(6806): 859-869.

Simmons, G. M., M. E. Kreitman, W. F. Quattlebaum and N. Miyashita (1989). Molecular analysis of the alleles of alcohol dehydrogenase along a cline in *Drosophila melanogaster*. I. Maine, North Carolina, and Florida. Evolution **43**(2): 393-409.

Simpson, J. T., K. Wong, S. D. Jackman, J. E. Schein, S. J. M. Jones and İ. Birol (2009). ABySS: a parallel assembler for short read sequence data. Genome Research **19**(6): 1117-1123.

Snaith, P. J. and T. A. Mansfield (1982). Stomatal sensitivity to abscisic acid: can it be defined? Plant, Cell & Environment **5**(4): 309-311.

Song, Y., Q. Chen, D. Ci and D. Zhang (2013). Transcriptome profiling reveals differential transcript abundance in response to chilling stress in *Populus simonii*. Plant Cell Reports **32**(9): 1407-1425.

Spreitzer, R. J. (2003). Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. Archives of Biochemistry and Biophysics **414**(2): 141-149.

Sreenivasulu, N., B. Usadel, A. Winter, V. Radchuk, U. Scholz, N. Stein, W. Weschke, M. Strickert, T. J. Close, M. Stitt, A. Graner and U. Wobus (2008). Barley grain maturation and germination: metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. Plant Physiology **146**(4): 1738-1758.

Štajner, N., S. Cregeen and B. Javornik (2013). Evaluation of reference genes for RT-qPCR expression studies in Hop (*Humulus lupulus* L.) during infection with vascular pathogen *Verticillium albo-atrum*. PLoS ONE **8**(7): e68228.

Stamm, P., P. Ravindran, B. Mohanty, E. Tan, H. Yu and P. Kumar (2012). Insights into the molecular mechanism of RGL2-mediated inhibition of seed germination in *Arabidopsis thaliana*. BMC Plant Biology **12**(1): 179.

Steinger, T., A. Stephan and B. Schmid (2007). Predicting adaptive evolution under elevated atmospheric $CO_2$ in the perennial grass *Bromus erectus*. Global Change Biology **13**(5): 1028-1039.

Streit, K., R. T. W. Siegwolf, F. Hagedorn, M. Schaub and N. Buchmann (2014). Lack of photosynthetic or stomatal regulation after 9 years of elevated [$CO_2$] and 4 years of soil warming in two conifer species at the alpine treeline. Plant, Cell & Environment **37**(2): 315-326.

Stylinski, C. D., W. C. Oechel, J. A. Gamon, D. T. Tissue, F. Miglietta and A. Raschi (2000). Effects of lifelong [$CO_2$] enrichment on carboxylation and light utilization of *Quercus pubescens* Willd. examined with gas exchange, biochemistry and optical techniques. Plant, Cell & Environment **23**(12): 1353-1362.

Sugano, S. S., T. Shimada, Y. Imai, K. Okawa, A. Tamai, M. Mori and I. Hara-Nishimura (2010). Stomagen positively regulates stomatal density in Arabidopsis. Nature **463**(7278): 241-244.

Sun, Z.-W. and C. D. Allis (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature **418**(6893): 104-108.

Suzuki, M. M. and A. Bird (2008). DNA methylation landscapes: provocative insights from epigenomics. Nature Review Genetics **9**(6): 465-476.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**(3): 585-595.

Takahata, N. and M. Nei (1985). Gene genealogy and variance of interpopulational nucleotide differences. Genetics **110**(2): 325-344.

Takebe, I., Y. Otsuki and S. Aoki (1968). Isolation of tobacco mesophyll cells in intact and active state. Plant and Cell Physiology **9**(1): 115-124.

Tanaka, Y., T. Sano, M. Tamaoki, N. Nakajima, N. Kondo and S. Hasezawa (2006). Cytokinin and auxin inhibit abscisic acid-induced stomatal closure by enhancing ethylene production in Arabidopsis. Journal of Experimental Botany **57**(10): 2259-2266.

Tang, F. C., C. Barbacioru, Y. Z. Wang, E. Nordman, C. Lee, N. L. Xu, X. H. Wang, J. Bodeau, B. B. Tuch, A. Siddiqui, K. Q. Lao and M. A. Surani (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nature Methods **6**(5): 377-U386.

Tans, P. and R. Keeling (2014). NOAA/ESRL (www.esrl.noaa.gov/gmd/ccgg/trends/). Scripps Institution of Oceanography.

Tarazona, S., F. García-Alcalde, J. Dopazo, A. Ferrer and A. Conesa (2011). Differential expression in RNA-seq: A matter of depth. Genome Research **21**(12): 2213-2223.

Taylor, G., N. R. Street, P. J. Tricker, A. Sjödin, L. Graham, O. Skogström, C. Calfapietra, G. Scarascia-Mugnozza and S. Jansson (2005). The transcriptome of Populus in elevated $CO_2$. New Phytologist **167**(1): 143-154.

Taylor, G., P. J. Tricker, F. Z. Zhang, V. J. Alston, F. Miglietta and E. Kuzminsky (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**(1): 177-185.

Teng, N., B. Jin, Q. Wang, H. Hao, R. Ceulemans, T. Kuang and J. Lin (2009). No detectable maternal effects of elevated $CO_2$ on *Arabidopsis thaliana* over 15 generations. PloS ONE **4**(6): e6035.

Thimm, O., O. Bläsing, Y. Gibon, A. Nagel, S. Meyer, P. Krüger, J. Selbig, L. A. Müller, S. Y. Rhee and M. Stitt (2004). MapMan: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. The Plant Journal **37**(6): 914-939.

Thomas, J. C., P. A. Godfrey, M. Feldgarden and D. A. Robinson (2012). Candidate targets of balancing selection in the genome of *Staphylococcus aureus*. Molecular Biology and Evolution **29**(4): 1175-1186.

Thomashow, M. F. (1999). Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Annual Review of Plant Physiology and Plant Molecular Biology **50**(1): 571-599.

Thompson, G. B. and B. G. Drake (1994). Insects and fungi on a C-3 Sedge and a C-4 grass exposed to elevated atmospheric $CO_2$ concentrations in open-top chambers in the field. Plant Cell and Environment **17**(10): 1161-1167.

Tognetti, R., A. Longobucco, F. Miglietta and A. Raschi (1999). Water relations, stomatal response and transpiration of *Quercus pubescens* trees during summer in a Mediterranean carbon dioxide spring. Tree Physiology **19**(4-5): 261-270.

Tognetti, R., A. Raschi and M. B. Jones (2000). Seasonal patterns of tissue water relations in three Mediterranean shrubs co-occurring at a natural $CO_2$ spring. Plant, Cell & Environment **23**(12): 1341-1351.

Tognetti R., A. Longobucco and A. Raschi (1999). Adaptation by *Plantago Lanceolata* and *Tussilago Farfara* To $CO_2$ Enrichment. Ecosystem Responses To $CO_2$: The MAPLE Project Results**:** 53-81.

Tonsor, S. J. (1985). Intrapopulational variation in pollen-mediated gene flow in *Plantago lanceolata* L. Evolution **39**(4): 775-782.

Torii, K. U. (2012). Mix-and-match: ligand–receptor pairs in stomatal development and beyond. Trends in Plant Science **17**(12): 711-719.

Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. Van Baren, S. L. Salzberg, B. J. Wold and L. Pachter (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology **28**(5): 511-515.

Tuba, Z., K. Szente and J. Koch (1994). Response of photosynthesis, stomatal conductance, water use efficiency and production to long-term elevated $CO_2$ in winter wheat. Journal of Plant Physiology **144**(6): 661-668.

Tusher, V. G., R. Tibshirani and G. Chu (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proceedings of the National Academy of Sciences **98**(9): 5116-5121.

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm and S. G. Rozen (2012). Primer3—new capabilities and interfaces. Nucleic Acids Research **40**(15): e115.

Urbanczyk-Wochniak, E., B. Usadel, O. Thimm, A. Nunes-Nesi, F. Carrari, M. Davy, O. Bläsing, M. Kowalczyk, D. Weicht, A. Polinceusz, S. Meyer, M. Stitt and A. Fernie (2006). Conversion of MapMan to allow the analysis of transcript data from solanaceous species: effects of genetic and environmental alterations in energy metabolism in the leaf. Plant Molecular Biology **60**(5): 773-792.

Vahisalu, T., H. Kollist, Y.-F. Wang, N. Nishimura, W.-Y. Chan, G. Valerio, A. Lamminmaki, M. Brosche, H. Moldau, R. Desikan, J. I. Schroeder and J. Kangasjarvi (2008). SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. Nature **452**(7186): 487-491.

Van de Water, P., S. Leavitt and J. Betancourt (1994). Trends in stomatal density and 13C/12C ratios of *Pinus flexilis* needles during last glacial-interglacial cycle. Science **264**(5156): 239-243.

Van Der Burgh, J., H. Visscher, D. L. Dilcher and W. M. Kürschner (1993). Paleoatmospheric signatures in neogene fossil leaves." Science **260**(5115): 1788-1790.

Van Gardingen, P. R., J. Grace, C. E. Jeffree, S. H. Byari, A. Raschi, F. Miglietta and I. Bettarini (1997). Long-term effects of enhanced $CO_2$ concentrations on leaf gas exchange: research opportunities using $CO_2$ springs. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 69-86.

Van Lun, M., J. S. Hub, D. Van der Spoel and I. Andersson (2014). $CO_2$ and $O_2$ distribution in Rubisco suggests the small subunit functions as a $CO_2$ reservoir. Journal of the American Chemical Society **136**(8): 3165-3171.

Van Oosten, J. J. and R. Besford (1996). Acclimation of photosynthesis to elevated $CO_2$ through feedback regulation of gene expression: climate of opinion. Photosynthesis Research **48**(3): 353-365.

Van Oosten, J. J., D. Wilkins and R. T. Besford (1994). Regulation of the expression of photosynthetic nuclear genes by $CO_2$ is mimicked by regulation by carbohydrates: a mechanism for the acclimation of photosynthesis to high $CO_2$? Plant, Cell & Environment **17**(8): 913-923.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe and F. Speleman (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology **3**(7): 11-31.

Vaquero, A., A. Loyola and D. Reinberg (2003). The constantly changing face of chromatin. Science of Aging Knowledge Environment **2003**(14).

Vijay, N., J. W. Poelstra, A. Künstner and J. B. W. Wolf (2013). Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive in silico assessment of RNA-seq experiments. Molecular Ecology **22**(3): 620-634.

Villablanca, F. X., G. K. Roderick and S. R. Palumbi (1998). Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. Molecular Ecology **7**(5): 547-560.

Vu, J. C. V., L. H. Allen, K. J. Boote and G. Bowes (1997). Effects of elevated $CO_2$ and temperature on photosynthesis and Rubisco in rice and soybean. Plant, Cell & Environment **20**(1): 68-76.

Wagner, G., K. Kin and V. Lynch (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory in Biosciences **131**(4): 281-285.

Wakeley, J. (1996). The variance of pairwise nucleotide differences in two populations with migration. Theoretical population biology **49**(1): 39-57.

Wang, L., Z. Feng, X. Wang, X. Wang and X. Zhang (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics **26**(1): 136-138.

Wang, Z., M. Gerstein and M. Snyder (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nature Review Genetics **10**(1): 57-63.

Wang, Z., C. Zang, J. A. Rosenfeld, D. E. Schones, A. Barski, S. Cuddapah, K. Cui, T.-Y. Roh, W. Peng, M. Q. Zhang and K. Zhao (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. Nature Genetics **40**(7): 897-903.

Ward, J. K. and B. R. Strain (1997). Effects of low and elevated $CO_2$ partial pressure on growth and reproduction of *Arabidopsis thaliana* from different elevations. Plant, Cell & Environment **20**(2): 254-260.

Warwick, K. R. and G. Taylor (1995). Contrasting effects of tropospheric ozone on five native herbs which coexist in calcareous grassland. Global Change Biology **1**(2): 143-151.

Weir, B. S. and C. C. Cockerham (1984). Estimating F-statistics for the analysis of population structure. Evolution: 1358-1370.

Weis, E. (1984). Short term acclimation of Spinach to high temperatures. Plant Physiology **74**(2): 402-407.

Wieneke, S., D. Prati, R. Brandl, J. Stöcklin, and H. Auge (2004). Genetic variation in *Sanguisorba minor* after 6 years in situ selection under elevated $CO_2$. Global Change Biology **10**(8): 1389-1401.

Willmer, C. and M. Fricker (1996). Stomata. Springer **2**.

Willmer, C. M. and J. E. Pallas (1974). Stomatal movements and ion fluxes within epidermis of *Commelina communis* L. Nature **252**(5479): 126-127.

Wilson, D. and J. P. Cooper (1969). Effect of light intensity and $CO_2$ on apparant photosynthesis and its relationship with leaf anatomy in genotypes of *Lolium Perenne* L. New Phytologist **68**(3): 627-644.

Wolfe, D. W., R. M. Gifford, D. Hilbert and Y. Luo (1998). Integration of photosynthetic acclimation to $CO_2$ at the whole-plant level. Global Change Biology **4**(8): 879-893.

Wong, S. C. (1979). Elevated atmospheric partial pressure of $CO_2$ and plant growth. Oecologia **44**(1): 68-74.

Woodward, F. I. (1987). Stomatal numbers are sensitive to increases in $CO_2$ from pre-industrial levels. Nature **327**(6123): 617-618.

Woodward, F. I. (1998). Do plants really need stomata? Journal of Experimental Botany **49**(Special Issue): 471-480.

Woodward, F. I. and F. A. Bazzaz (1988). The responses of stomatal density to $CO_2$ partial-pressure. Journal of Experimental Botany **39**(209): 1771-1781.

Woodward, F. I., D. J. Beerling, F. Miglietta, A. Raschi, R. Tognetti and P. v. Gardingen (1997). Plant $CO_2$ responses in the long term: plants from $CO_2$ springs in Florida and tombs in Egypt. Plant Responses to Elevated $CO_2$: Evidence from Natural Springs: 103-113.

Woodward, F. I. and C. K. Kelly (1995). The Influence of $CO_2$ Concentration on Stomatal Density. New Phytologist **131**(3): 311-327.

Woodward, F. I. (1999). Adaptation by *Plantago lanceolata* and *Tussilago farfara* to $CO_2$ enrichment. Ecosystem Responses To $CO_2$: The MAPLE Project Results**:** 133-142.

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution **19**(3): 395-420.

Wu, C. A., D. B. Lowry, A. M. Cooley, K. M. Wright, Y. W. Lee and J. H. Willis (2007). Mimulus is an emerging model system for the integration of ecological and genomic studies. Heredity **100**(2): 220-230.

Wulff, R. D. and H. M. Alexander (1985). Intraspecific variation in the response to $CO_2$ & enrichment in seeds and seedlings of *Plantago lanceolata* L." Oecologia **66**(3): 458-460.

Wullschelger, S. D. (1993). Biochemical Limitations to carbon assimilation in C3 plants—a retrospective analysis of the A/Ci curves from 109 species. Journal of Experimental Botany **44**(5): 907-920.

Xiao-Ping, S. and S. Xi-Gui (2006). Cytokinin- and auxin-induced stomatal opening is related to the change of nitric oxide levels in guard cells in broad bean. Physiologia Plantarum **128**(3): 569-579.

Xu, D. Q. (1994). Stomatal and nonstomatal acclimation to a $CO_2$-enriched atmosphere. Biotronics **23**: 1-9.

Xu, H., A. Shida, F. Futatsuya and A. Kumura (1994). Effects of epibrassinolide and abscisic acid on Sorghum plants growing under soil water deficit II. Physiological basis for drought resistance induced by exogenous epibrassinolide and abscisic acid. Japanese Journal of Crop Science **63**(4): 676-681.

Xu, Z. and G. Zhou (2008). Responses of leaf stomatal density to water status and its relationship with photosynthesis in a grass. Journal of Experimental Botany **59**(12): 3317-3325.

Yang, K., M. Jiang and J. Le (2014). A new loss-of-function allele 28y reveals a role of ARGONAUTE1 in limiting asymmetric division of stomatal lineage ground cell. Journal of Integrative Plant Biology **56**(6): 539-549.

Yang, M. and F. D. Sack (1995). The too many mouths and four lips mutations affect stomatal production in Arabidopsis. The Plant Cell Online **7**(12): 2227-2239.

Young, J. J., S. Mehta, M. Israelsson, J. Godoski, E. Grill and J. I. Schroeder (2006). $CO_2$ signaling in guard cells: calcium sensitivity response modulation, a $Ca^{2+}$ independent phase, and $CO_2$ insensitivity of the gca2 mutant. Proceedings of the National Academy of Sciences **103**(19): 7506-7511.

Young, M. D., D. J. McCarthy, M. J. Wakefield, G. K. Smyth, A. Oshlack and M. D. Robinson (2012). Differential expression for RNA sequencing (RNA-Seq) data: mapping, summarization, statistical analysis, and experimental design bioinformatics for high throughput sequencing. N. Rodríguez-Ezpeleta, M. Hackenberg and A. M. Aransay, Springer New York**:** 169-190.

Zemp, N., A. Minder and A. Widmer (2014). Identification of internal reference genes for gene expression normalization between the two sexes in dioecious white campion." PLoS ONE **9**(3): 92-93.

Zerbino, D. R. and E. Birney (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Research **18**(5): 821-829.

Zhang, G., G. Guo, X. Hu, Y. Zhang, Q. Li, R. Li, R. Zhuang, Z. Lu, Z. He, X. Fang, L. Chen, W. Tian, Y. Tao, K. Kristiansen, X. Zhang, S. Li, H. Yang, J. Wang and J. Wang (2010). Deep RNA sequencing at single base-pair resolution reveals high complexity of the rice transcriptome. Genome Research **20**(5): 646-654.

Zhang, S. and Q. L. Dang (2013). $CO_2$ elevation improves photosynthetic performance in progressive warming environment in white birch seedlings. F1000Research **2**(13).

Zhang, X., L. Zhang, F. Dong, J. Gao, D. W. Galbraith and C.-P. Song (2001). Hydrogen peroxide is involved in abscisic acid-induced stomatal closure in *Vicia faba*. Plant Physiology **126**(4): 1438-1448.

Zhao, Q.-Y., Y. Wang, Y.-M. Kong, D. Luo, X. Li and P. Hao (2011). Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study. BMC Bioinformatics **12**(Suppl 14): S2.

Zinta, G., H. AbdElgawad, M. A. Domagalska, L. Vergauwen, D. Knapen, I. Nijs, I. A. Janssens, G. T. S. Beemster and H. Asard (2014). Physiological, biochemical, and genome-wide transcriptional analysis reveals that elevated $CO_2$ mitigates the impact of combined heat wave and drought stress in *Arabidopsis thaliana* at multiple organizational levels. Global Change Biology **20**(12): 3670-3685.

Zotz, G., S. Pepin and C. Körner (2005). No down-regulation of leaf photosynthesis in mature forest trees after three years of exposure to elevated $CO_2$. Plant Biology **7**(4): 369-374.