The University Of Sheffield.

A Structural and Molecular Atlas for Early Leaf Development in Rice (*Oryza sativa*)

By:

Naomi Cox

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Science
Department Animal and Plant Sciences

July 2020

Abstract

Rice is one of the most important crops in the world, and several global projects are underway to improve its photosynthetic performance to secure yields. Such projects rely on fundamental alterations to leaf development to change its anatomy and biochemistry. To date, the majority of studies of rice leaf development have focussed on the spatial development of already maturing leaves, which may differ from temporal development as the leaf forms *de novo*. In this thesis, I focus on this temporal development, and report a morphological and metabolic characterisation of its early stages.

In Chapter 2, I use scanning electron microscopy to document the temporal development of leaf primordia. I have produced a novel image dataset describing the epidermal differentiation pattern of early development. Use of this technique across two lines of domestic and two lines of wild rice has shown a conservation of epidermal patterning during development. Further, I have created a developmental index to allow for accurate developmental staging of leaf primordia too small to judge by eye.

In Chapter 3, I have developed a protocol to examine the metabolite fingerprints of individually dissected leaf primordia using mass spectrometry. This technique can be used to distinguish between metabolic profiles over different stages of leaf development. This work has identified metabolites to act as potential developmental markers, and suggests that flux through the tricarboxylic acid (TCA) cycle is also indicative of developmental stage.

Finally, in Chapter 4 I describe a set of experiments examining the relationship between leaf anatomy and photosynthetic performance in a genetically engineered rice population with variation in individual anatomical parameters. In this study, there is no correlation between photosynthetic performance and any individual structural trait. This highlights the need for careful consideration of genetic constructs, and a greater knowledge of the leaf developmental trajectory.

Collaborations

The work presented in this thesis is my own work, however some data have been collected in collaboration with colleagues.

In Chapter 2, images contributing to two figures were taken by Hannah Jones and James Pitman. This has been clearly indicated in the figure legends.

The work presented in Chapter 3 is entirely my own, but I would like to acknowledge the training and support received by Dr Heather Walker.

The work presented in Chapter 4 contributes to a larger, ongoing project at Sheffield co-ordinated by Dr Jen Sloan. The data is an amalgamation of work by myself, Jen Sloan, and Irma Canicosa. It has been clearly indicated at the start of this chapter which work was performed by myself, and which was performed by my colleagues.

Acknowledgements

Firstly, I gratefully acknowledge the financial support provided by the BBSRC and my CASE partner IRRI.

Thanks go to my supervisors, Andy Fleming and Lisa Smith, for their support, advice and encouragement throughout the project. I am grateful also to Jen Sloan and Heather Walker for invaluable conversations and technical advice. At IRRI, thanks go to Paul Quick, Jacque Dionora, Irma Canicosa, Jia Yin and to the IRRI Education Team for facilitating my placement.

It has been a privilege to be part of D59 for the last four years. The companionship, support, collaboration and cake have made for a truly excellent work environment. I would particularly like to thank Alice Baille, both for her personal and professional advice, and for being a fantastic PhD role model.

I am lucky to have been supported by my wonderful friends and family throughout the whole science adventure. Particular thanks go to my parents, and to Nicole, Phill and Tom.

List of Abbreviations

A	CO_2 Assimilation Rate
AWD	Alternate Wetting and Drying
BSC	Bundle Sheath Cell
C_i	Intercellular CO_2 Concentration
CLV3	CLAVATA3
DESI	Desorption Electrospray Ionisation
ESI	Electrospray Ionisation
EU	European Union
Flv	Flavodiiron
GC-MS	Gas Chromatography Mass Spectrometry
GDC	Glycine Decarboxylase Complex
GM	Genetically Modified
g_m	Mesophyll Conductance
GMC	Guard Mother Cell
GRiSP	Global Rice Science Partnership
g_s	Stomatal Conductance
HPE1	*HIGH PHOTOSYNTHETIC EFFICIENCY1*
HPLC	High Performance Liquid Chromatography
IRGA	Infrared Gas Exchange Analysis
IRRI	International Rice Research Institute
J_{max}	Ribulose 1,5-bisphosphate Regeneration Rate
L1, L2…	Leaf 1, 2…
MALDI	Matric-assisted Laser Desorption/Ionisation
MC	Mesophyll Cell
MS	Mass Spectrometry
MSI	Mass Spectrometry Imaging
MS/MS	Tandem Mass Spectrometry
m/z	Mass Charge Ratio
NMR	Nuclear Magnetic Resonance Spectroscopy
OPLS-DA	Orthogonal Partial Least Squares Discriminant Analysis
PCA	Principal Component Analysis
Pi	Plastochron Number
PIP1;2	*PLASMA MEMBRANE INTRINSIC PROTEIN1;2*
P1, P2…	Plastochron 1, 2…
QTL	Quantitative Trait Locus
RGSV	Rice Grassy Stunt Virus
ROS	Reactive Oxygen Species
Rubisco	Ribulose-1,5-bisphosphate carboxylase/oxygenase
SAM	Shoot Apical Meristem
SC	Subsidiary Cell
SCR	SCARECROW
SE	Secondary Electron
SEM	Scanning Electron Microscope/Microscopy
SHR	SHORTROOT
SMC	Subsidiary Mother Cell
TCA	Tricarboxylic Acid
TEM	Transmission Electron Microscopy
TOF	Time of Flight
TPU	Triose Phosphate Utilisation
V	Vein
VC_{max}	Maximum Rate of Rubisco
WUS	WUSCHEL

Table of Contents

List of Figures

List of Tables

List of Appendices and Supplemental Figures

Chapter 1: General Introduction

1.1 Rice and Global Food Security

Global food security is one of the most important issues facing humanity in the 21st century. The world population already stands at 7.6 billion people, a number that is set to increase to a predicted peak of 10 billion in 2050 (Melorose et al., 2015). To feed a world of 10 billion, it is estimated that current rate of food production needs to double – a goal we are not currently on target to meet (Ray et al., 2013). Strategies to increase crop yields in response to population increases are further complicated by a number of additional factors, including global climate change, limited suitable arable land, and restricted water availability.

Rice is one of the most important global plants: it is the second most widely grown crop, and provides the most calories through direct consumption (Seck et al., 2012). Importantly, rice is a primary foodstuff for half of the world's population, providing up to 70% of daily calories for people in some of the world's poorest countries (Figure 1.1 (Global Rice Science Partnership (GRiSP), 2013)). Two species of rice are cultivated: *Oryza glaberrima* (African rice) and *Oryza sativa* (Asian rice), with Asian rice having many sub-species including *japonica*, *indica*, *aus* and aromatic, resulting in a wide geographic distribution.

Rice yields benefited from substantial yield increases during the Green Revolution, particularly due to the introduction of semi-dwarf varieties such as IR8 and IR64 (Evenson and Gollin, 2003), and the widespread adoption of modern farming technologies such as nitrogen-based fertilisers and pesticides. In recent years, however, the upward trend has slowed, and in some cases, yields are beginning to plateau (Ray et al., 2013). Thus, we must explore all available avenues to achieve the required dramatic increase in the face of the rapidly encroaching 2050 deadline.

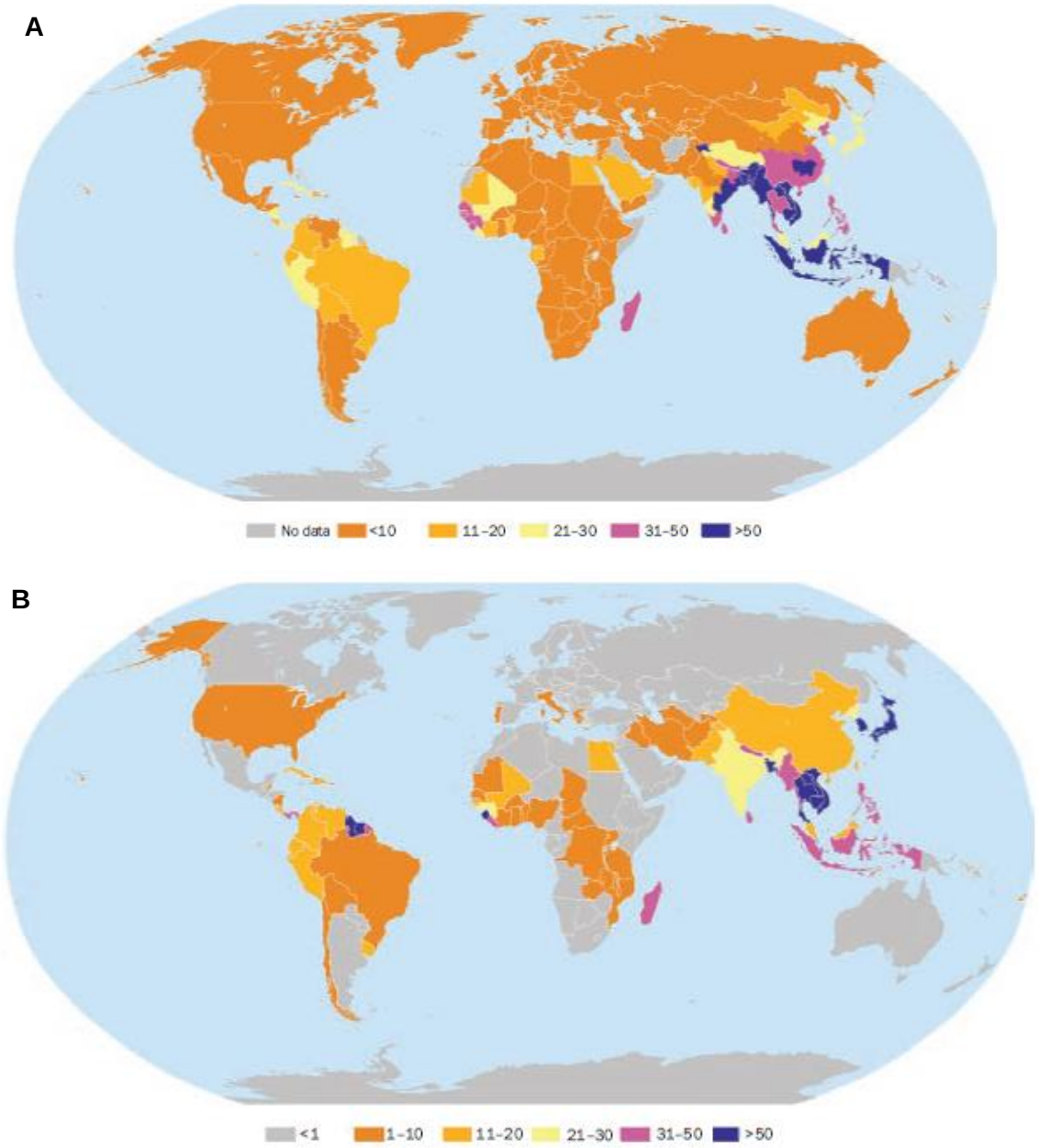


Figure 1.1 – Global distribution of rice production and caloric intake.

Maps show the distribution of percent of daily calories coming from rice (A) and harvested crop area coming from rice (B).

Figure from the International Rice Research Institute (IRRI) Rice Almanac (GRiSP, 2013).

1.2 Improvements to Rice

1.2.1 Conventional Breeding and Wild Rice as a Genetic Resource

Rice improvement through conventional breeding differs from most other widely grown crops in that rice is not a “cash crop”. For example, most rice farms in the Philippines are small, family-owned farms of less than 2 hectares (in comparison, the average-sized maize farm in Iowa, USA is 135 hectares). This, combined with vast climatic differences and varied taste preferences, means that the varieties grown by farmers are rarely uniform within a community, let alone on a country-wide scale. This poses a challenge to breeders, who cannot focus on yield alone; varieties that do not meet the taste, texture and appearance preferences of their target demographic will not be received well, no matter how high yielding.

Both international institutes, such as the International Rice Research Institute (IRRI), and national institutes, such as PhilRice in the Philippines, have their own breeding programs, each focusing on the requirements of their country. In many areas, rice breeding remains fairly traditional, with breeders and farmers keen to introgress elite lines with native or wild varieties with desirable traits. One notable exception is the elite breeding program at IRRI, recently redesigned by Joshua Cobb, which uses an evolutionary selection approach within only the elite lines to improve yield, resulting in higher yields and time reduction from a nine-year breeding cycle to a six- to seven-year breeding cycle (Cobb et al., 2019; Collard et al., 2019).

Both inbred and hybrid rice varieties are grown globally, with hybrid varieties having higher yields (Haque et al., 2015). The production of these lines is reliant on cytoplasmic male sterility or genic male sterility (Li et al., 2007). As there are limited available lines for this breeding, introgression of stress-resistance traits still remains necessary, which further compounds the issue of lengthy breeding cycles (De Guzman and Oard, 2018).

While *O. sativa* and *O. glaberrima* are the only cultivated rice species, they belong to a genus of around 25 *Oryza* species (GRiSP, 2013). Amongst the *Oryza* genus, genomes can be split into 11 groups, six diploid and five tetraploid. Species are either diploid or tetraploid, with both domesticated species possessing diploid AA genomes. Wild rice germplasm is considered a promising genetic tool, with, for example, various wild lines possessing quantitative trait loci (QTLs) conferring resistance to various biotic and abiotic stresses (Zhang and Xie, 2016). This diversity has been exploited for some time to create inter-specific hybrids by backcrossing wild rice that performs favourably for a given trait, including resistance to bacterial blast and blight (Sitch et al., 1992) and resistance to brown

planthopper (Ishii et al., 1994) into a domestic line. It also has improved resistance to abiotic stresses such as salinity (Menguer et al., 2017; Quan et al., 2018). Germplasm has further been explored for genes conferring other useful traits such as rice sheath blight resistance (Prasad et al., 2008).

However, like all routes relying on conventional breeding, there are several drawbacks to using wild lines only in traditional breeding programmes. Firstly, breeding in a new trait is time consuming: traditional rice breeding programs have breeding cycles of around nine years (Collard et al., 2019). Secondly, as breeding is reliant on natural variation that occurs at meiosis, it is hard to monitor which traits are being passed along. While marker-assisted selection does provide a tool to speed up conventional breeding and make it more precise (Collard and Mackill, 2008), this focusses on the trait of interest. It cannot monitor all of the other genetic exchange taking place. Importantly, breeding reliant on monogenic traits, as crosses involving wild rice often are, has the disadvantage of evolution making the gene introgression redundant. One of the classic “success stories” of creating an improved rice line through backcrossing with a wild line is the introduction of resistance to Rice Grassy Stunt Virus (RGSV) – which can decimate entire rice crops – from *Oryza nivara* (Khush and Ling, 1974) into *Oryza sativa*. However, shortly after the introduction of resistant cultivars into agriculture, newer, more virulent strains of the virus evolved that were also infective to the newly-released resistant strains (Hibino et al., 1985). Thus, while wild rice germplasm does provide a promising gene bank, exploiting this through conventional breeding alone is unlikely to be sufficient for significantly improving cultivated rice.

One final issue to consider concerning conventional breeding of rice crops is that for each trait newly introgressed into an elite variety, the same process must be repeated for each variety in circulation. Obviously, genetic engineering approaches face the same issue; however, it takes substantially less time to introduce change with modern genetic techniques than with conventional breeding.

1.2.2 Genetic Approaches

Advances in genetic engineering technology since the 1980s mean there are now a wide variety of options to modify crop plants (Prado et al., 2014). Entire genes either from rice (cis-genic) or from other species such as maize (transgenic) can be transformed into rice through biolistics, or, more commonly, *Agrobacterium*-mediated transformation. Golden Gate cloning (Weber et al., 2011) has recently become popular in plants. It allows for the transformation of multiple genes in the same transformation event, which is a promising strategy for multigenic traits. More recently, CRISPR-Cas9 has become a popular tool for genome editing, as it does not require the transfer of genetic material to the marketable

product. A plant that has undergone genome editing by CRISPR is thus indistinguishable from a plant that was produced through conventional breeding, making it an attractive tool in many countries where transgenic crops are prohibited.

Genetically Modified (GM) crops are now grown in 28 countries and take up 10% of global arable land (Royal Society, 2016). There are commercial GM varieties available of 10 crops: maize, canola, cotton and soybean are widely grown, while alfalfa, aubergine, papaya, potato, squash and sugar beet are grown on a much smaller scale (Royal Society, 2016). While GM crops are grown in countries on every inhabited continent, there are still many countries and regions with restrictions on growth, import, or both, the most notable being the European Union (EU). Over half of its member states ban cultivation of GM crops, though the EU imports 30 million tonnes of GM soy and corn every year for animal feed, making it the largest importer of GM crops in the world (Tagliabue, 2017). Despite mixed uptake, there have been several commercial and scientific success stories of lines developed with improved biotic and abiotic stress tolerance (such as Bt-maize, cotton and aubergine), and improved nutrition (such as golden rice) (Kamthan et al., 2016).

Despite the advantages of genetic technology for crop improvement, there are also several important scientific considerations. One consideration when using this technology is the insertion site of the DNA. In both biolistics and *Agrobacterium*-mediated transformation, the DNA is integrated randomly into the genome. In rice, “safe harbours” for gene insertion have been discovered (Yin *et al*., poster proceedings): regions of the genome where new genetic material can be inserted with no negative consequences to the plant, although this does require time-consuming screening. Regions of the genome can also now be edited very precisely using CRISPR.

Although legislation differs across the world, there are several advantages to genetic modification over conventional breeding. Firstly, genetic engineering is significantly faster: for the introduction of a monogenic trait, it takes six to nine years to introgress this through conventional breeding, but only two generations to an F_2 homozygote (taking less than a year) using genetic technology. Secondly, it is more specific: while in genetic engineering it is known exactly what genetic material is being altered, crossing between lines is reliant on genetic recombination, hence not only the gene of interest is affected in a conventional cross. Genetic engineering can also allow for the disruption of haplotypes where undesirable alleles are often inherited in parallel with alleles of interest due to linkage disequilibrium. Finally, many ambitious projects concerning multigenic traits are impossible to implement using only conventional breeding methods.

1.2.3 Genetic Modification to Alter Photosynthesis

Of the many avenues to explore to improve rice yields, one of the most promising is photosynthesis. Most crops perform photosynthesis at a tiny fraction of their theoretical potential, rice being one of the worst performing of the major crops (Fitzgerald and Resurreccion, 2009). Ultimately, photosynthesis is responsible for producing all the energy required for plant growth, and all the energy that enters the food chain through grain consumption. It has thus been identified as one of the most important targets for crop improvement (Zhu et al., 2010).

Put simply, photosynthesis refers to the set of reactions whereby plants use light energy to convert water and carbon dioxide to sugars and oxygen (Figure 1.2). The key reactions in photosynthesis take place in the chloroplasts where light is captured by photosynthetic pigments, primarily chlorophyll, providing the energy for the light-dependent reactions that produce ATP and NADPH via electron transport. These molecules can then be used for CO_2 fixation into triose phosphates via the Calvin-Benson cycle.

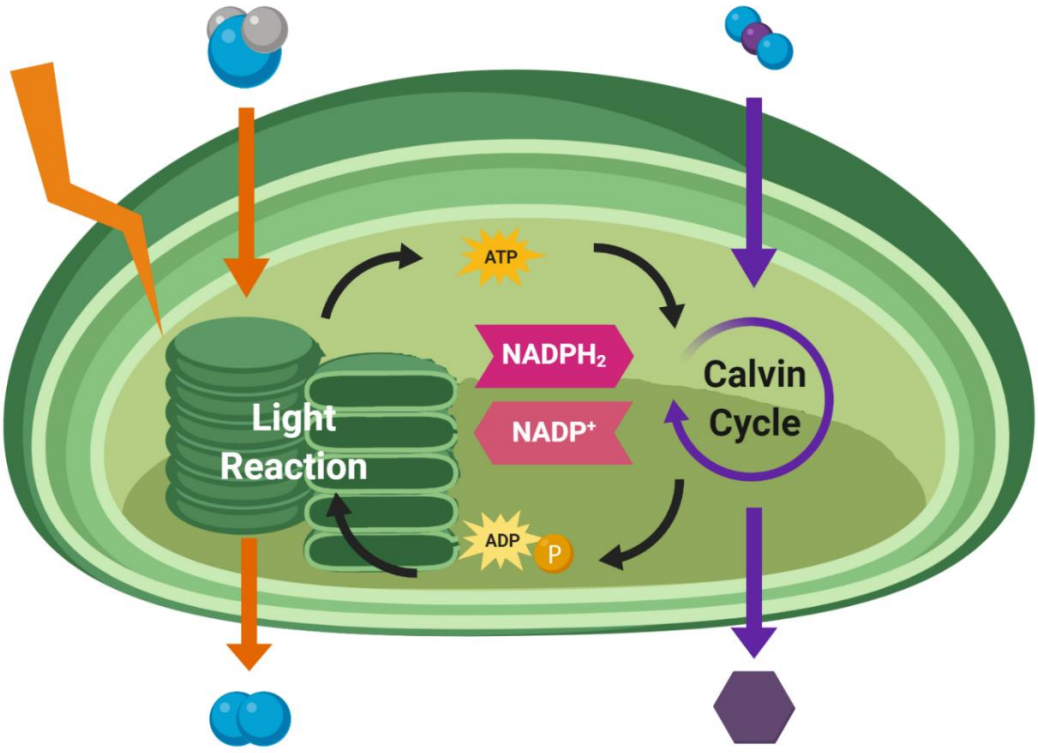


Figure 1.2 – A simplified model of photosynthesis

The majority of the reactions of photosynthesis take place in the chloroplast. The light reactions occur in the thylakoid membranes, and uses water to produce oxygen, as well as ATP and $NADPH_2$ for the Calvin-Benson cycle. The Calvin-Benson cycle, which is light-independent and occurs in the stroma, uses this ATP and $NADPH_2$ to fix carbon dioxide into sugars. In this diagram, Oxygen, Hydrogen and Carbon are represented by blue, grey

and purple spheres respectively, and generic sugars by the purple hexagon.
Figure made using biorender.com.

As photosynthesis is a complex pathway, there are several areas we can target to improve it. These are summarised in Figure 1.3. I have grouped these broadly into overall canopy structure, leaf-level physical changes including leaf structure and organelle dynamics, and biochemical changes including electron transport chain efficiency and carbon fixation. In reality, however, these areas are interlinked, and often affecting one will bring about change in another.

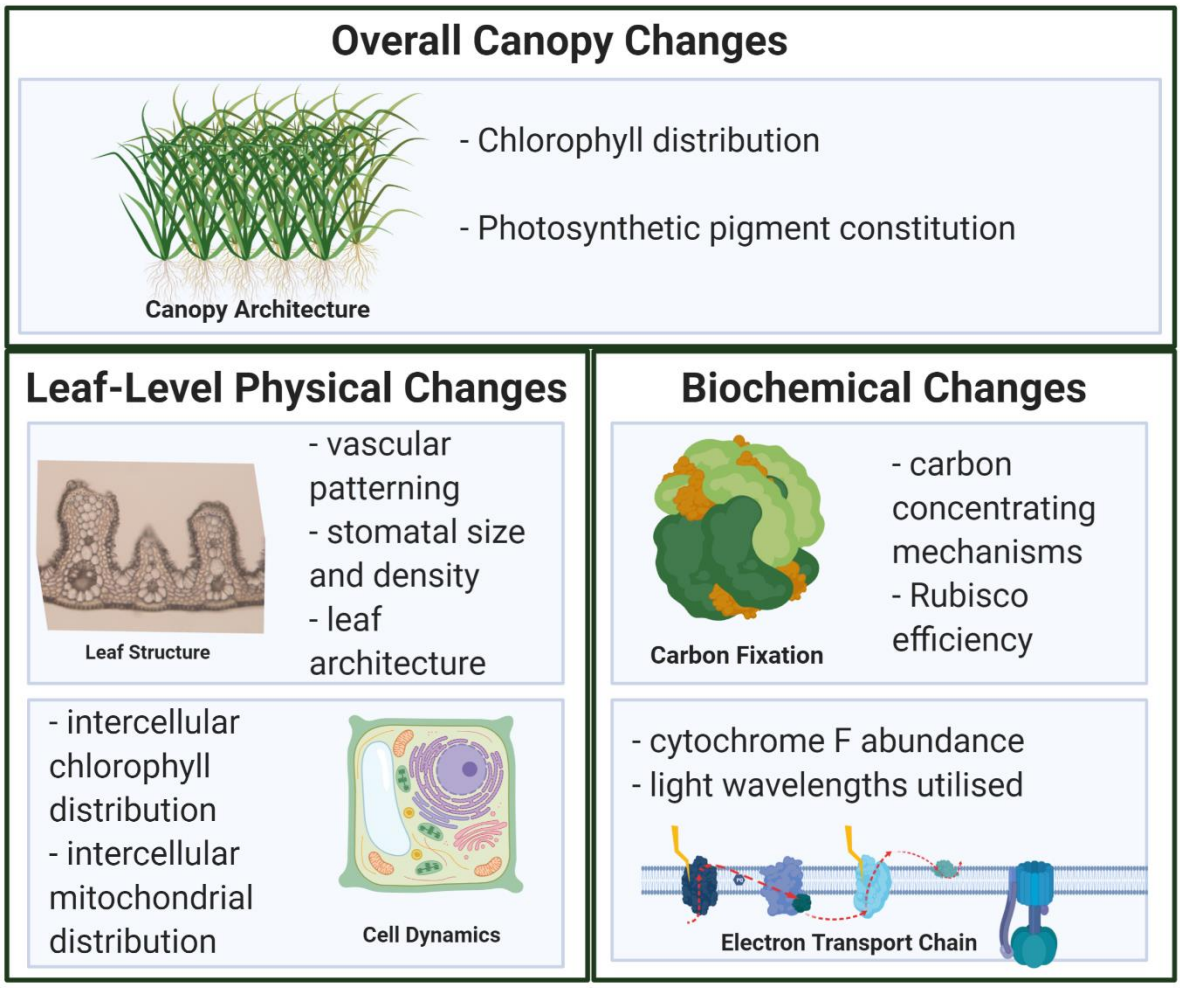


Figure 1.3 – Potential areas for photosynthetic improvement
Summary of some available targets for photosynthetic improvement in crops, grouped by canopy changes, leaf level anatomical changes and biochemical changes.
Figure created with biorender.com.

1.2.3.1 Biochemical Canopy Level Changes

Considering that the majority of energy and carbon required for grain production is produced by the last emerging leaf – known as the flag leaf, many efforts to improve photosynthesis are targeting the upper canopy of the rice crop in the hopes of securing improved yields. Several approaches in this area are examining the effects of changing the location, size or distribution of chlorophyll and other photosynthetic pigments.

Optimising the antenna size in the flag leaf (Ort et al., 2011) could have a small effect on increasing canopy photosynthesis (by 3%), with the added benefit of increasing nitrogen use efficiency by 14% (Song et al., 2017) according to computational simulations. There have also been suggestions regarding altering canopy chlorophyll distribution to improve photosynthesis. Currently, the plant's photosynthetic machinery is saturated at just 25% of full sunlight (Zhu et al., 2008), yet the majority of light is absorbed by the upper canopy. Reducing the chlorophyll content in the upper canopy could be a way to increase nitrogen use efficiency, and consequently plant performance, without a yield penalty (Walker et al., 2018). In addition to modifying existing pigments, an approach which exploits the photosynthetic pigments produced in other kingdoms could also be beneficial. Chloroplasts in higher plants can utilise light of wavelengths from 400 to 700nm – less than half of the solar spectrum. It is predicted that around a fifth more photons could be captured by expanding this upper limit to 750nm through engineering in bacterial photosynthetic pigments (Chen and Blankenship, 2011; Ho et al., 2016).

Recently, a gene, *HIGH PHOTOSYNTHETIC EFFICIENCY 1* (*HPE1*) has been identified in *Arabidopsis thaliana* that optimises light-harvesting pigments (Jin et al., 2016) through decreasing total chlorophyll content and increasing chlorophyll *a*/*b* ratio. Arabidopsis *hpe1* mutants had more rapid electron transport than their wild-type counterparts, and higher photosynthetic efficiency under high light stress. The authors suggest that this could be used as a strategy for improving the light capture in other higher plants, although analogues have not yet been identified in a crop.

1.2.3.2 Leaf-Level Physical Changes

In addition to optimising performance of the flag leaf to influence yield, many studies are interested in leaf structural changes which will impact the photosynthetic performance of the whole plant.

Studies in the *Oryza* genus have examined the physical leaf properties known to be associated with photosynthesis in other model plants. These can be grouped into structural features (including leaf thickness, mesophyll cell size and mesophyll cell surface area), and functional traits (including mesophyll conductance, stomatal conductance and transpiration rate) (Giuliani et al., 2013). While it is known that high mesophyll conductance and high

stomatal conductance correlate positively with the rate of photosynthesis, neither functional trait has been found to correlate positively with any individual structural trait: this suggests that a multi-faceted approach will be required to increase photosynthesis through structural alterations (Giuliani et al., 2013). In contrast to this, more recent work overexpressing the aquaporin *OsPIP1;2* (*PLASMA MEMBRANE INTRINSIC PROTEIN 1;2)* produced plants with mesophyll conductance around 150% higher than wild-type plants (Xu et al., 2019), suggesting that single-gene approaches can still yield beneficial results to functional traits.

Elevated stomatal conductance has been shown to increase the rate of photosynthesis in rice grown in well-watered conditions (Kusumi et al., 2012). However, as freshwater availability decreases, well-irrigated rice systems may become less common – many rice research institutes already advocate for switching to an alternate wetting and drying (AWD) model (CGIAR, 2014).Taken together, it is important to consider the performance of the plant concerning water use efficiency when trying to affect photosynthesis by altering stomatal conductance. While decreasing stomatal conductance by reducing stomatal density consequently reduces photosynthetic rate, these lower density plants perform significantly better under drought conditions; thus, there is no yield penalty in this case (Caine et al., 2018).

1.2.3.3 Biochemistry

Several approaches aiming to elucidate photosynthetic change focus on optimising the biochemical pathways involved. Broadly, this often focusses on either electron transport or carbon fixation.

One major concern when optimising the electron transport chain is ensuring adequate photoprotection. Damage can occur to the photosystems when light is absorbed beyond what is needed for photosynthesis – light levels that plants frequently face outdoors (Zhu et al., 2008). There are several ways plants deal with excess light energy, including heat dissipation, production of reactive oxygen species (ROS), photoinactivation and the xanthophyll cycle (Demmig-Adams and Adams III, 1992). Adequate photoprotection is particularly important, as the extreme results of damage by light include photobleaching and cell death (Murchie and Niyogi, 2011).

One way to further mitigate this photodamage could be to engineer in additional photoprotective proteins. Flavodiirons (Flv) are proteins present across bacteria, archaea and eukarya (although absent in the higher plants) that redirect electrons away from photosystem I under fluctuating light conditions, preventing the damaging effects of excess light (Allahverdiyeva et al., 2013). Introduction of two Flv proteins into *Arabidopsis thaliana* sped up photorecovery under fluctuating light, while having no effect under steady light (Yamamoto et al., 2016). As fluctuating light more closely resembles conditions that crops

encounter in the field, Yamamoto and colleagues hypothesise that replicating this in crop species could increase yield.

The organisms that experience the highest rates of photoprotection are diatoms and algae. In extreme high-light conditions (2000µmols^{-1}m^{-2}), the green alga *Chlorella ohadii* is still able to grow with minimal photo-damage by blocking the majority of linear electron flow, relying on 90% cyclic electron flow for growth (Ananyev et al., 2017). While cyclic electron transport can occur in photosystem II in higher plants, the mechanisms are currently poorly understood (Lysenko et al., 2016). Work in Arabidopsis has identified some of the genes involved in protecting photosystem II from damage under high light conditions, for example *MAINTENANCE OF PSII UNDER HIGH LIGHT 1* (*MPH1*) (Liu And Last, 2015), which appear to have orthologues throughout the plant kingdom. As mutation of such genes causes increased photodamage, it is possible that improvement or overexpression would improve photoprotection. Further, recent work has suggested that salicylic acid may aid in the protection of photosystem II under high light conditions by preventing the disassembly of its complexes (Chen et al., 2020).

A core biochemical target for plant improvement is carbon fixation. Carbon fixation refers to the process by which CO_2 is converted to sugars by the Calvin-Benson cycle. Carbon fixation in cyanobacteria is much more efficient than in higher plants as a result of several biochemical and structural differences, reviewed in Zarzycki et al., 2013. One core reason that carbon fixation can be so inefficient in higher plants is the inefficiency of Ribulose-1,5-bisphosphate carboxylase/oxygenase – commonly known as Rubisco. While it is the primary enzyme for carbon fixation in the Calvin-Benson cycle, Rubisco can also fix O_2, which is subsequently salvaged in a process known as photorespiration. This is both energetically wasteful and produces toxic by-products which must then be recycled. Importantly, levels of photorespiration increase under high temperatures and high light intensities (Peterhansel et al., 2010). It is predicted that in the absence of photorespiration, crop yields could increase between 12% and 55% (Walker et al., 2016). It will not be sufficient to simply substitute plant Rubisco for bacterial, as while the cyanobacterial enzyme is faster and hence more efficient, it is less specific for CO_2 than plant Rubisco. To overcome this, cyanobacteria contain specialised intercellular compartments called carboxysomes, which concentrate CO_2 around Rubisco. There have been some attempts to engineer these structures into C_3 plants (Hanson et al., 2017), but they are unlikely to be a useful solution for multicellular higher plants.

Where carboxysomes are not an option, plants have also evolved several ways to overcome the problem of photorespiration: two forms of photosynthesis separate CO_2 from Rubisco, either temporally (CAM) or spatially (C_4). The majority of CAM plants are succulents and

other desert plants. The only species of economic importance is the bromeliad pineapple (*Ananas comosus* L.) (Ming et al., 2015), so it is an unlikely candidate for integration in to rice. C_4 photosynthesis concentrates CO_2 around Rubisco in the bundle sheath cells by capturing CO_2 in the mesophyll, and shuttling it to the bundle sheath via a four-carbon compound (Figure 1.4). As C_4 photosynthesis has evolved independently from C_3 over 60 times, with around 60% of C_4 species being grasses (Gowik and Westhoff, 2011) – including another grass crop maize (*Zea mays* L.) – it is reasonable to believe that creating a C_4 rice will be possible.

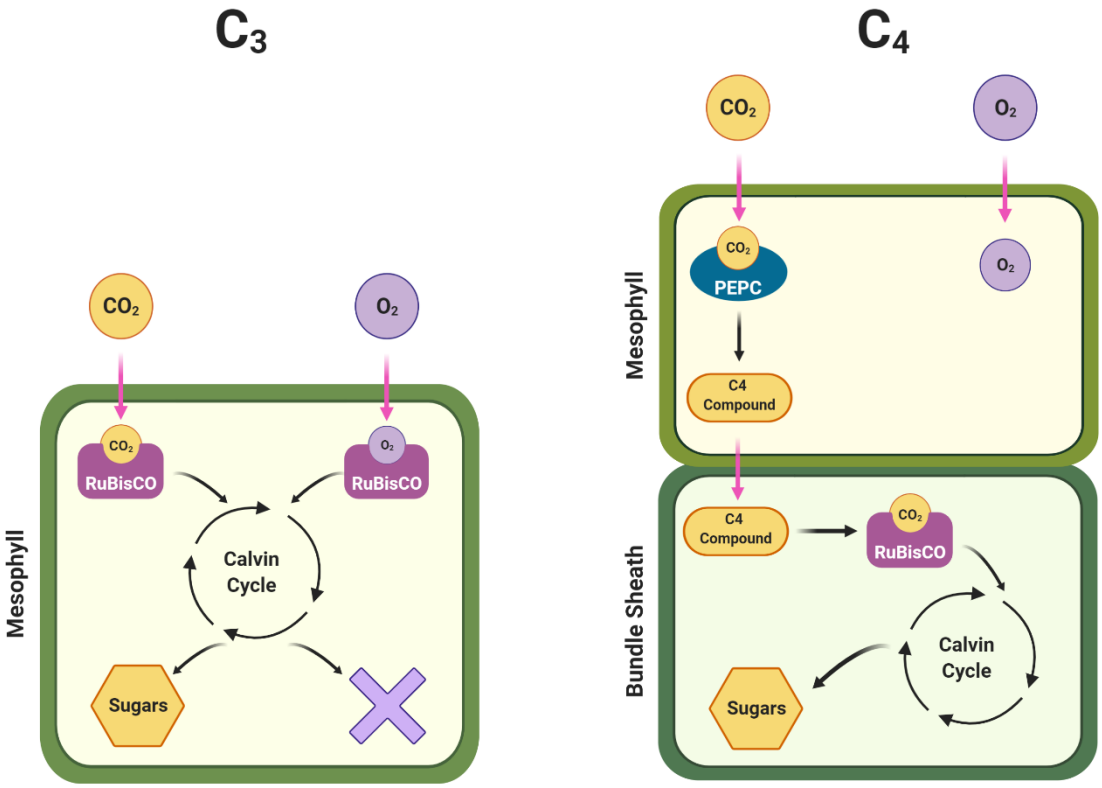


Figure 1.4 – A simplified diagram of C_3 and C_4 photosynthesis

In C_3 photosynthesis, CO_2 enters the mesophyll and is fixed by Rubisco in the Calvin-Benson cycle, where sugars are produced. Rubisco can also fix O_2, resulting in photorespiration.

In C_4 photosynthesis, CO_2 enters the mesophyll cell where it is fixed in a four-carbon compound by PEPC. This four-carbon compound is shuttled to the bundle sheath cell, where CO_2 is then released and fixed by Rubisco for entry into the Calvin-Benson cycle.

Diagram made with biorender.com

1.3 C_4 Rice – Successes and Stumbling Blocks

"The C_4 Rice Project is one of the scientific 'Grand Challenges' of the 21st Century. Researchers from 12 institutions in 8 countries are working together to apply innovative scientific approaches to the development of high yielding rice varieties for smallholder farmers".

c4rice.com

Founded in 2008, the C_4 Rice Project is an international collaborative project aiming to 'upgrade' the C_3 photosynthetic pathway in rice to the more efficient C4 pathway. Several changes, both biochemical and anatomical, will be needed in order to transition rice into a C_4 plant. While in this introduction I will consider them independently, efforts to engineer a C_4 rice will require co-ordinated development between the two.

Firstly, the anatomy of a C_3 and C_4 leaf differs greatly (Langdale, 2011) (Figure 1.5). C_4 plants display a characteristic Kranz anatomy, where veins (V) are surrounded by large bundle sheath cells (BSC) with two mesophyll cells (MC) between each vascular bundle, leading to a pattern of V-BSC-MC-MC-BSC-V across the middle of the leaf. In contrast, in C_3 plants such as rice the bundle sheath cells are much smaller and there are many more mesophyll cells between each vein, leading to a pattern of V-BSC-MC-MC-MC-MC-MC-MC-MC-BSC-V or similar.
Further, bundle sheath cells in C_4 plants have more chloroplasts by volume than in C_3 plants, which seem to use their bundle sheath cells predominantly for starch accumulation (Leegood, 2008).

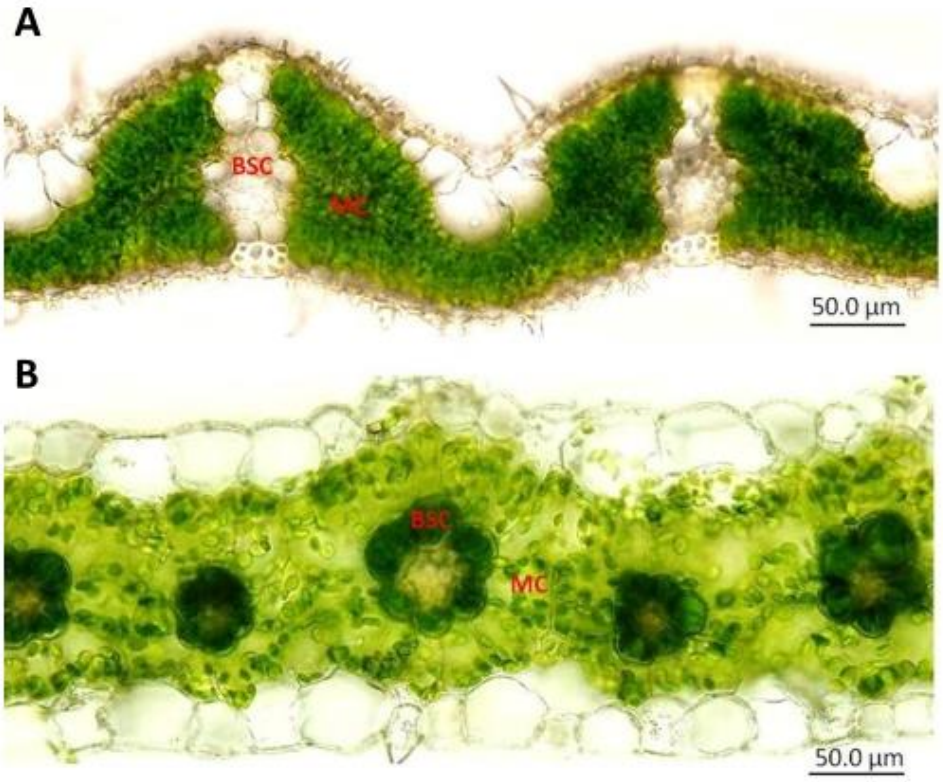


Figure 1.5 – Anatomical differences between C_3 and C_4 leaves

Shown are representative leaf cross sections of a C_3 plant (A – *Oryza sativa* (rice)) and a C_4 plant (B – *Setaria viridis*). BSC – Bundle Sheath Cell, MC – Mesophyll Cell.

Figure from Karki et al., 2013.

A core part of the initial work in the C_4 rice project has been genetic screens of rice and maize in order to identify candidate genes that may be vital in the transition from C_3 to C_4 photosynthesis (Li et al., 2010; Majeran et al., 2010; Wang et al., 2014, 2013). We are approaching a comprehensive understanding of the genetic basis of C_4 photosynthesis in maize. However, efforts to engineer these genes into rice to induce C_4 anatomy or biochemistry have so far had limited success. For example, recent attempts to individually constitutively overexpress 60 maize genes in rice resulted in either no significant change from wild type, or a negative effect on growth (Wang et al., 2017a).

As the C_4 rice project has progressed, it has become increasingly apparent that the introduction of Kranz anatomy into the rice leaf is vital. The development of the characteristic Kranz anatomy can be divided broadly into three stages (Langdale, langdalelab.com). Firstly, the procambium, from which the Kranz tissue develops (Brown, 1975) is initiated; secondly, the mesophyll and bundle sheath cells differentiate (reviewed in Nelson and Langdale, 1989); and thirdly, chloroplasts of distinct types develop in the bundle sheath and mesophyll cells (Furbank and Taylor, 1995) allowing C_4 cycle integration.

One possible approach to integrate Kranz anatomy into rice is to study the evolutionary history of C_4 development from C_3, and mimic these steps in rice using genetic engineering. For example, one well-supported model suggests that plants that perform C_2 photosynthesis – which concentrates CO_2 around Rubisco through a photorespiratory glycine shuttle (Tolbert, 1997) – are intermediates between C_3 and C_4. It has thus been suggested that C_3 plants evolved to C_4 via C_2, with a "proto-Kranz" anatomy between C_3 and C_2, and a "C_4-like" anatomy between C_2 and C_4 (Figure 1.6) (Sage et al., 2014). The anatomical traits that vary between C_3 and C_4 grasses have been well studied (Lundgren et al., 2014), and support the idea of this accelerated evolutionary approach of developing C_4 rice through mimicking steps found in nature. Indeed, Wang and colleagues have succeeded in producing rice with proto-Kranz anatomy (Wang et al., 2017b), the first evolutionary step towards C_4.

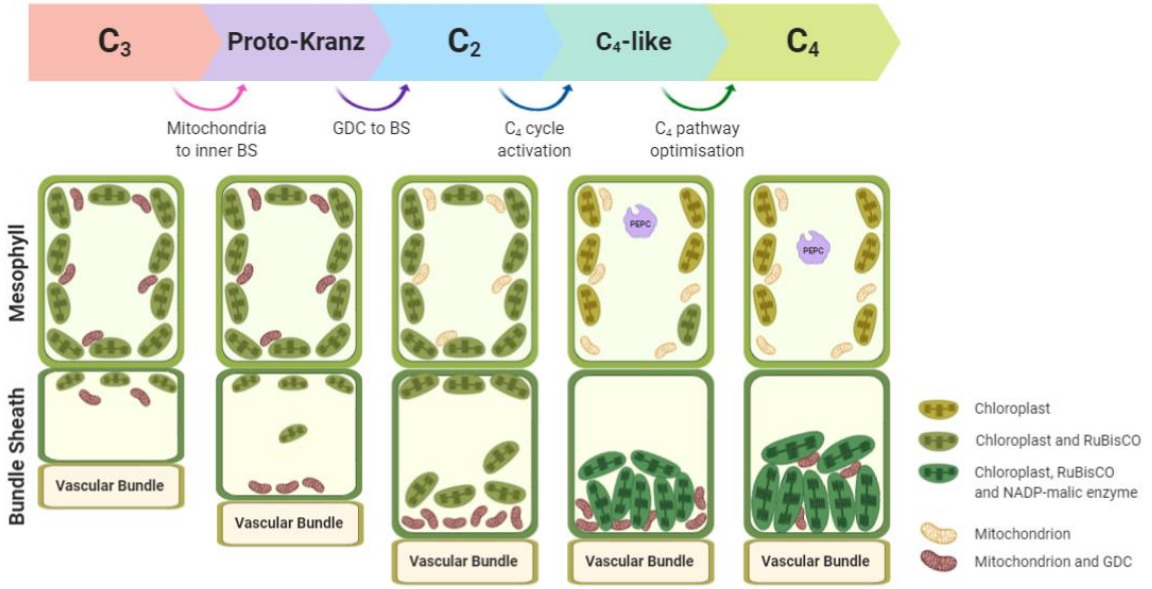


Figure 1.6 – Transition from C_3 to C_4 anatomy

Evolutionary model for the transition from C_3 to C_4 anatomy as proposed by the Sage lab.
In C_3 photosynthesis, the majority of photosynthetic activity occurs in the mesophyll. Between C_3 and proto-Kranz, the BSCs enlarge and the mitochondria shift towards the inner membrane. In C_2 photosynthesis, glycine decarboxylase activity has shifted to the BSC, beginning to concentrate CO_2. In C_4-like photosynthesis, PEPC is expressed in the MC, shuttling CO_2 to the BSC, however the glycine decarboxylase activity remains to rescue CO_2 lost due to photorespiration from photosynthesising chloroplasts in the MC. Finally, in C_4 photosynthesis the pathway is optimised, and CO_2 concentration in the BSC is achieved through PEPC activity.
BSC – Bundle Sheath Cells, GDC – Glycine Decarboxylase Complex, MC – Mesophyll Cells.
Original diagram, simplified from Sage et al., 2014.
Made with biorender.com

While there have been suggestions that engineering a C_2 rice would be more beneficial than C_4 (Bellasio et al., 2019), this argument is reliant on model data, and does concede that C_4 is

more efficient in high temperatures and light intensities. Engineering C_4 rice remains an important challenge for securing the rice supply chain given the increasing problems facing crop plants due to ever-higher temperatures (Fitzgerald and Resurreccion, 2009), and the fact that many rice-producing countries – particularly some of the poorest – are located in the tropics, where high temperatures and light intensities are the norm (Global Rice Science Partnership (GRiSP), 2013).

A further consideration for engineering such drastic changes in leaf structure and biochemistry is the mode of engineering. The majority of transgenic lines created so far have used a maize ubiquitin promoter. However, when altering fundamental processes such as early leaf development, constitutive expression of a gene can perturb plant growth (Sun et al., 2017). While both bundle sheath and mesophyll cell-specific promoters from maize can be expressed in rice (Peterhansel, 2011), we still do not have a clear picture of when in development the anatomy of a rice leaf is established, and it is possible that by expressing genes using these cell-specific promoters, they are not acting sufficiently early in development to have a marked effect i.e. their temporal pattern of expression may not be ideal or sufficient.

This last point identifies a surprising gap in our knowledge. Although there have been numerous studies providing detailed accounts of the earliest stages of leaf development in many “standard” plant species (e.g. Arabidopsis (reviewed in Kalve et al., 2014) and maize (e.g. Wang et al., 2013), and similar evolutionary development studies (for example in *Cardamine hirsuta* (Hay and Tsiantis, 2016) - a close relative of *Arabidopsis thaliana*), detailed information on the earliest stages of rice leaf development remains at best fragmented or superficial. Furthermore, it is becoming increasingly clear that both monocot and dicot leaves are influenced by both spatial and temporal cues (reviewed in Nelissen et al., 2016), while many classical studies of rice leaf development consider mainly, or exclusively spatial cues.

A broad characterisation for temporal rice leaf development does exist (Itoh et al., 2005, described fully in Chapter 2), with development split into six stages termed P1, where the leaf primordium is initially formed, to P6, which represents a fully differentiated leaf, as shown in Figure 1.7. The key focus of this thesis is the transition between the P3 stage, where the primordium cells are still undifferentiated, and the P4 stage, where the leaf blade rapidly elongates. There are sizeable structural and biochemical changes during this transition, including the establishment of photosynthetic competence (van Campen et al., 2016), yet information on this differentiation of naïve tissue remains scarce. Gaining knowledge of the relative morphological stages at which metabolic transitions occur may aid

in targeting gene expression for C_4 anatomy and biochemistry to an appropriate developmental stage.

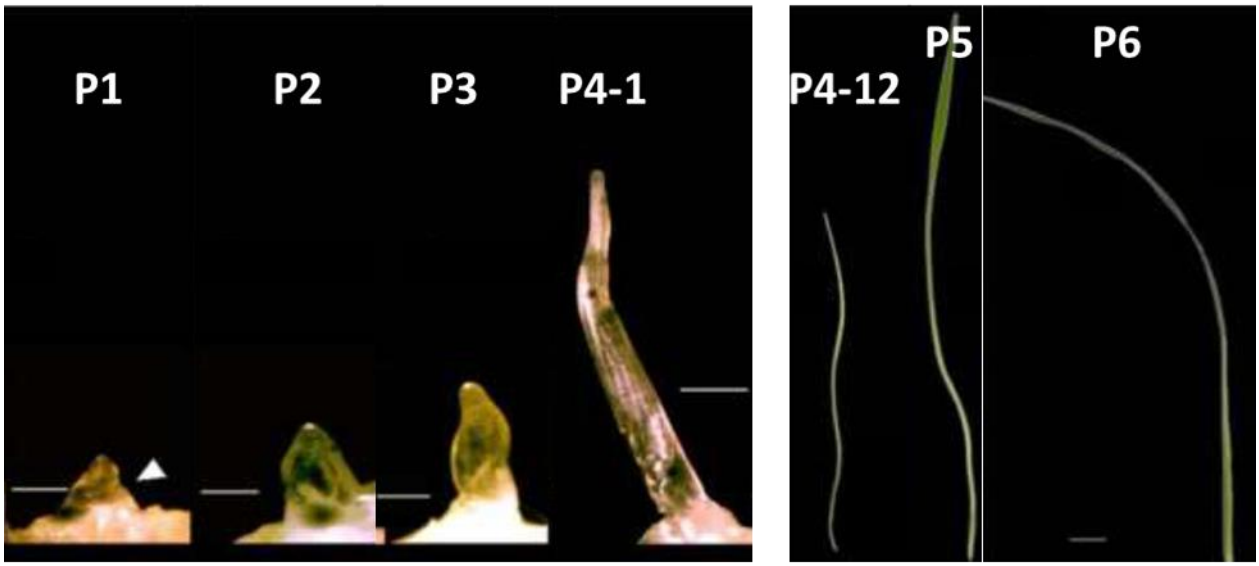


Figure 1.7 – Stages of rice leaf development

Light microscopy images of the P1 to P6 stages of rice leaf development. Scale bars are 125µm for P1 to P4-1; 1cm for P4-12 to P6. Images taken by Julia van Campen, included in van Campen, thesis, 2016.

1.4 Project Summary

While there have been advances in studying the biochemical and anatomical changes required for the switch from C_3 to C_4, an emerging problem is that rice is recalcitrant to manipulation – often there is low genetic penetration, or a shift back to a C_3 state (Paul Quick, personal communication). It has been hypothesised that this observation could be explained by a strict temporal trajectory in the rice leaf from P1 to P6. Possibly, rice leaf structure and function is already set before transgenes are acting to change this.
Thanks to a previous CASE project between the Fleming lab and IRRI, we now know that the transition to photosynthetic competence occurs between the P3 and P4-1 developmental stages (van Campen et al., 2016); however, very little is known about this transition.

In this thesis, I present an integrated atlas for morphological and metabolic changes in early leaf development in rice to fill a fundamental knowledge gap in this area. Further, I hope that that this information can be used both to better select phases for genetic engineering in projects aimed at introducing changes to the rice leaf, in particular to improve photosynthetic efficiency. The project uses a combination of microscopy and biochemical approaches to document this key transition phase in rice leaf development with respect to the formation of a functional leaf structure. I also describe an analysis of the relationship between leaf structure

and photosynthetic function in a population of rice plants with varying leaf anatomy, which were produced by the C_4 rice lab at IRRI.

1.5 Aims

1. Produce a morphological atlas describing the early stages of rice leaf development, focussing on the stages where previous data have indicated that photosynthetic competence is established.
2. By comparing data obtained from a range of rice cultivars and species, investigate the extent to which a single atlas may be used as a generic tool to describe early rice leaf development, and thus the conservation of morphological development over the *Oryza* genus.
3. Using the information produced in (1), produce a metabolic atlas describing the early stages of rice leaf development.
4. Using a population of transgenic rice lines, investigate the relationship between individual structural traits and photosynthetic performance.

1.6 Objectives

1. Produce a high-quality image dataset of early development in rice leaves, focusing on the P3 to P4 transition.
2. Develop a method to quantify the developmental age of leaf primordia.
3. Collect metabolomics data on staged leaf primordia.
4. Identify marker metabolomic compounds at various stages of rice leaf development.
5. Investigate the relationship between leaf anatomy and photosynthetic function in a genetically engineered rice population

1.7 Hypotheses

1. Early leaf development in rice entails a conserved pattern of morphological changes.
2. The pattern of morphological change in (1) is reflected by changing patterns of metabolite expression.
3. The transition to photosynthetic competence in leaf development in rice is reflected by a conserved pattern of metabolite accumulation related to primary plant metabolism.
4. Structure-function relationships predicted *in silico* by a novel 3D photosynthetic model would be reflected *in planta* in our study population.

Chapter 2 – A Structural Index of Early Leaf Development

2.1 Introduction

2.1.1 Rice Leaf Development

Rice is a key crop that provides over 20% of the world's calories (GRiSP, 2013). All the carbon in the grain is sourced by photosynthesis, thus there is significant interest in improving this trait to increase yield. Projects looking to alter photosynthesis, including the C_4 Rice Project, will require changes to leaf anatomy, biochemistry and physiology (von Caemmerer et al., 2012). However, there are still areas where surprisingly little is known about rice leaf development, particularly *de novo* leaf development.

Leaves are initiated as a protrusion from the shoot apical meristem (SAM), which is encased by the older leaves. Co-ordinated processes of cell division, elongation and differentiation occur as the leaf develops, eventually producing a mature leaf with a blade and sheath encasing the younger leaves. The age of a leaf can be described both by the order in which it has formed (L1, L2, L3…), with leaf 5 being considered the first adult leaf[1] (Asai et al., 2002; Itoh et al., 2005; Tanaka et al., 2011), and by the plastochron number, (Pi) (P1, P2, P3…), which describes the developmental age of the leaf. (Although confusingly, "plastochron" is also defined as the time period between the initiation of each new leaf from the SAM). In this thesis, "plastochron" will be used as a unit of developmental time. Thus, it is possible to compare equivalent developmental stages of different individual plants.

Leaf initiation is a well-characterised process. Studies of leaf initiation in grasses have been performed in maize through microdissection of primordia before P1 (Ohtsu et al., 2007), with genetic analysis demonstrating conservation of the CLAVATA3-WUSCHEL negative feedback loop which is well understood from studies of Arabidopsis development. This loop (Figure 2.1) maintains a pool of stem cells (the SAM) (Somssich and Je, 2016) from which all new leaves form. WUSCHEL (WUS) is a transcription factor responsible for regulating genes that promote SAM identity (Laux et al., 1996). CLAVATA3 (CLV3) is a small secreted

[1] In this project, leaf five is considered the fifth 'true' leaf – the coleoptile is not numbered as a leaf. Thus 'leaf five' as described in this thesis would be considered as 'leaf six' by some. As leaf five was selected as the first representative adult leaf, it is important that those who consider the coleoptile to be leaf one and wish to replicate this work ensure experiments are performed on an adult leaf.

peptide (Clark et al., 1995), which promotes the differentiation of stem cells in the SAM by downregulation of WUS. The two act antagonistically, to ensure that a pool of pluripotent stem cells remains while organs develop from the meristem.

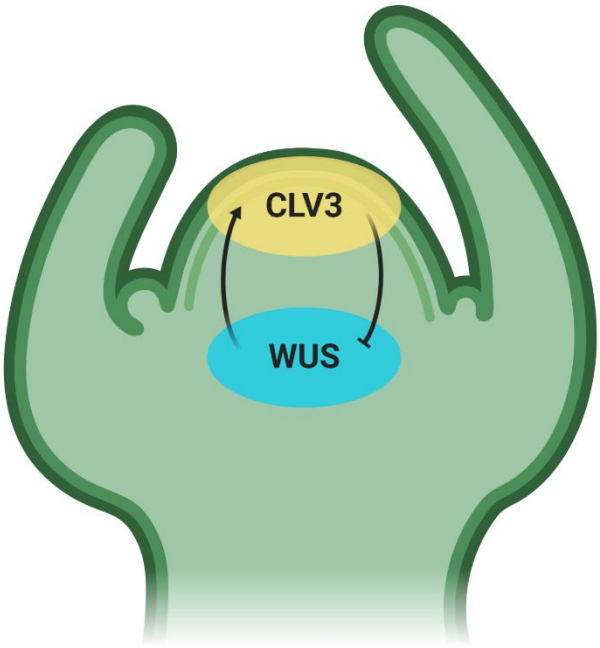


Figure 2.1 – The CLAVATA3-WUSCHEL negative feedback loop

While there are many other molecular factors at play, a simplified model of SAM proliferation is that expression of CLV3 downregulates *WUS*, whose product promotes stem cell proliferation. Pictured here is the feedback loop in *Arabidopsis*, although the same molecular system is conserved throughout the higher plants, including rice. Figure created with biorender.com

Leaves are then able to initiate from this pool of stem cells, and a key regulator of leaf initiation is auxin. Studies of plants grown in the presence of auxin transport inhibitors, in addition to studies of auxin biosynthesis mutants have shown that in dicots, cotyledons develop in areas of high auxin, while there is low auxin in the boundary space in between (Barton, 2010). The growing leaf acting as an auxin sink causes the next leaf primordium to be initiated as far away as possible, resulting in a spiral phyllotaxis in plants such as *Arabidopsis thaliana* (Jönsson et al., 2006). Other patterns of leaf initiation, such as the distichous phyllotaxy displayed in maize and rice, can also be explained by auxin fluxes, although in this case due to asymmetrical auxin distribution (Martinez et al., 2016). Much of this work has been done using computational modelling, however markers for studying auxin in rice have recently been developed (Yang et al., 2017), which will allow for complementation with biological data.

Post initiation, leaves must differentiate from stem cells to a mature, functional leaf. In rice, leaf development comprises seven broad stages, termed P0 (the group of cells in the SAM which are fated to become the next leaf) to P6 (a fully mature leaf), as summarised in (Itoh et al., 2005).

Table 2.1 – Staging leaf development in rice.
Staging system proposed by the Itoh lab. Table adapted from Itoh *et al*., 2005. Epidermal cell fate has been highlighted, as this is the focus of this chapter.

Stage		Events
Name	**Main Characteristic**	
P0	Designation of leaf cells	Leaf founder cells are designated from the SAM.
P1	Formation of primordium	The primordium begins to protrude from the SAM. Leaf margins elongate.
P2	Hood-shaped primordium	Characteristic hood shape. The two leaf margins overlap. Vascular bundle differentiates.
P3	Ligule primordium	Formation of the leaf-sheath boundary. Differentiation of sclerenchymatous cells. Initiation of epidermal-specific cells.
P4	Leaf blade elongation	Rapid elongation of the leaf blade. Differentiation of epidermal-specific cells.
P5	Leaf sheath elongation	Elongation of the leaf sheath. The leaf blade emerges from the sheath of the previous leaf. Lacunae form. Maturation of leaf epidermal cells.
P6	Mature leaf	Leaf blade bends at the lamina joint.

During the P4 developmental stage, the leaf blade elongates to more than ten times its original length, resulting in a range of morphologically different leaves being encompassed in this category. This stage has been further sub-divided into stages P4-1 to P4-12, which are determined by the leaf blade length in cm (Kusumi et al., 2010). It should be noted, however, that as these refer to absolute measurements, it is possible that a 1cm-long leaf 3 primordium and a 1cm-long leaf 10 primordium, for example, would be at different developmental stages. Thus, sub-dividing the P4 stage by length is unhelpful in a developmental context, and having more structural or biochemical information would help precisely define the developmental age.

The developmental staging above provides a good basal framework to describe rice leaf development. However, it lacks detail within specific stages. This is particularly evident during the early stages (P1-P4) when virtually all of the main patterns evident in the mature leaf are laid down. For example, data in the literature indicate that major elements of vascular differentiation occur during the P2 stage (Scarpella et al., 2000) and that the establishment of photosynthetic competence occurs in the P3 to P4 transition (van Campen et al., 2016). A deeper understanding of the timing and trajectory of the patterning of these early stages of development would be beneficial to help in the understanding of rice leaf development. This may be of particular relevance in the context of continued efforts to improve photosynthesis in rice since these early stages are likely to be key targets for the engineering of rice leaves.

2.1.2 Patterning Along Developing Grass Leaves

Studies examining leaf development in monocots frequently take advantage of the basipetal nature of monocot leaf development (Nelson and Langdale, 1989). In these leaves, cell proliferation is seen at the base of the leaf, followed by cell differentiation then elongation. This leads to the gradient along the leaf axis, with the oldest cells at the tip of the leaf and cells at the base of the leaf being in a proliferative state (either undergoing cell division or recently exited cell division). This provides a relatively easy way to study morphological and photosynthetic changes linked to differentiation along a single blade as a leaf matures. There are several advantages to this approach: larger leaves are easier to manipulate, the technique is not necessarily destructive, and many epidermal cells can be observed under a standard light microscope, reducing the need for specialist dissection skills or advanced and expensive imaging facilities. The trade-off for this, however, is that immature cells differentiating within a leaf blade are responding to different signals to those in *de novo* development. Cell fate determination has already occurred once the cells have organised into files, as position is the primary determinant of cell fate in plants (Schiefelbein, 1994). Further, the cells in a maturing adult leaf can respond to photosynthetic signals produced from elsewhere within the same leaf. Thus, it is impossible to know by looking along a maturing adult blade when cell fate was determined – information that will be vital if we are to alter the anatomy of *de novo* rice leaf development.

2.1.3 *De Novo* Leaf Development in Rice and Maize

As many studies in rice also utilise the developmental gradient along a maturing blade (Li et al., 2010), comparatively more is known about *de novo* early leaf development in the related monocot *Zea mays* (maize, Wang *et al.*, 2013) than in rice. In fact, to date there have been no in-depth studies of *de novo* early leaf development in rice. Interestingly, maize has both C_4 (foliar) and C_3 (husk) leaves, making it an interesting developmental model for the establishment of Kranz anatomy compared to 'standard' C_3 development.

The FP1 and FP2 developmental stages in maize (equivalent to P1/P2 in rice as shown in Figure 2.2B) comprise a collection of undifferentiated cells, where the mid-vein is not yet visible. By the FP4 stage (equivalent to P3 in rice as shown in Figure 2.2D), the midvein and lateral veins can be distinguished, and at FP5 (equivalent to P4 as shown in Figure 2.2E) Kranz anatomy is completely visible. This maturation of photosynthetic structures by FP5 is reflected by the work previously published by the Fleming lab, demonstrating that rice becomes photosynthetically competent during the P3/P4 transition (van Campen et al., 2016a). Interestingly, in FP3/FP4 leaves there are 4-6 mesophyll cells between every vein, while in FP5 this has reduced to the two seen in Kranz anatomy (Wang et al., 2013), further demonstrating the importance of focusing genetic changes aiming to induce Kranz anatomy in rice no later than the P3/P4 transition.

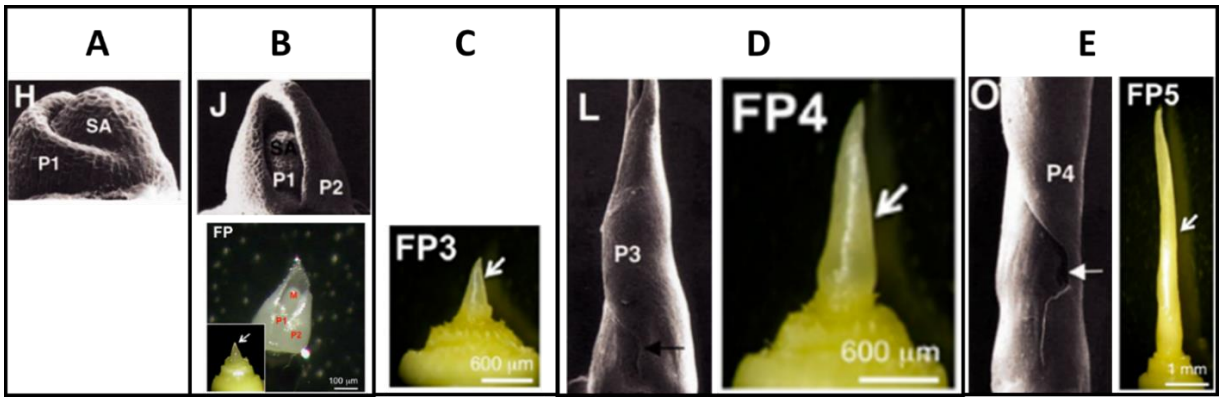


Figure 2.2 – Equivalent developmental stages in rice and maize

A – P1 and SAM stages in rice

B – P1 and P2 stages in rice (upper), FP stage in maize (lower)

C – FP3 stage in maize

D – P3 stage in rice (left), FP4 stage in maize (right)

E – P4 stage in rice (left), FP5 stage in maize (right)

All rice images from (Itoh et al., 2005), maize images from (Wang et al., 2013). FP leaves represent foliar leaves, which do exhibit Kranz anatomy.

In addition to the patterning of the internal tissue (described above), there are numerous complex patterning processes that occur within the epidermis of the rice leaf. As with all plants, the leaf epidermis in rice, like human skin, forms a boundary between the environment and the plant. In addition to fulfilling several important functions for the plant (including protection and regulation of gas and water exchange) (Javelle et al., 2011), it also provides an excellent tool for studying development and differentiation.
One reason for this is that the development of the internal and external structures of the leaf are intrinsically linked. This is because the leaf must co-ordinate, for example, the placement of air spaces between mesophyll cells with the placement of stomata on the epidermis to allow gas exchange for photosynthesis and transpiration. It is still unknown how this is achieved, and the mechanism is sure to be complex (Tsukaya, 2005). We do know that this development is plastic and responsive to environmental signals. For example, plants grown under elevated CO_2 show reduced stomatal density (number of stomata per unit area) and index (proportion of stomata compared to pavement cells per unit area) (Teng et al., 2006). It has been suggested that epidermal cell growth can regulate the division and expansion of other leaf cells, including in the mesophyll (Savaldi-Goldstein et al., 2007). Thus, as internal and external development are so intertwined, using the epidermis to study development can provide interesting insight into the development of the plant. Moreover, since it is easier to view the epidermis than the internal leaf structure, understanding the relationship of epidermal differentiation with events internal to the leaf can provide a useful tool to infer changes in differentiation at a whole leaf level.

Broadly, epidermal cells in the rice leaf can be split into pavement cells and more specialised cells. These specialised cells include guard cells and their associated subsidiary cells, and leaf hairs - better known as trichomes. Trichomes can be further characterised based on various features. While the terminology appears to differ in the literature, rice has repeatedly been described to display two types of “stinging hair” trichome (Maes and Goossens, 2010), referred to in this thesis as ‘macrohairs’ (larger hairs seen only in the silica ladders (X-shaped structures regularly spaced between two walls, as described in Yamanaka *et al*., 2009)) and ‘prickle hairs’ (smaller, pointy trichomes which are more unevenly distributed over the epidermal surface, sometimes referred to in the literature as “microhairs”).

These different epidermal cell types present on rice leaves have been documented for some time (Chaffey, 1983). More recently, the distribution – and to some degree development – of these cells have been more closely studied, including work focusing on vasculature (van Campen, 2016, Thesis), stomata (Luo et al., 2012), and silica bodies (Yamanaka et al., 2009).

In one example, Luo and colleagues documented the maturation of various epidermal features along rice leaf blades (summarised in Figure 2.3) (Luo et al., 2012). Where and when these epidermal features arise during the *de novo* development of rice leaf primordia is yet to be reported. It is also unknown to what extent these two patterns correlate, nor how external differentiation correlates with the internal changes to the leaf.

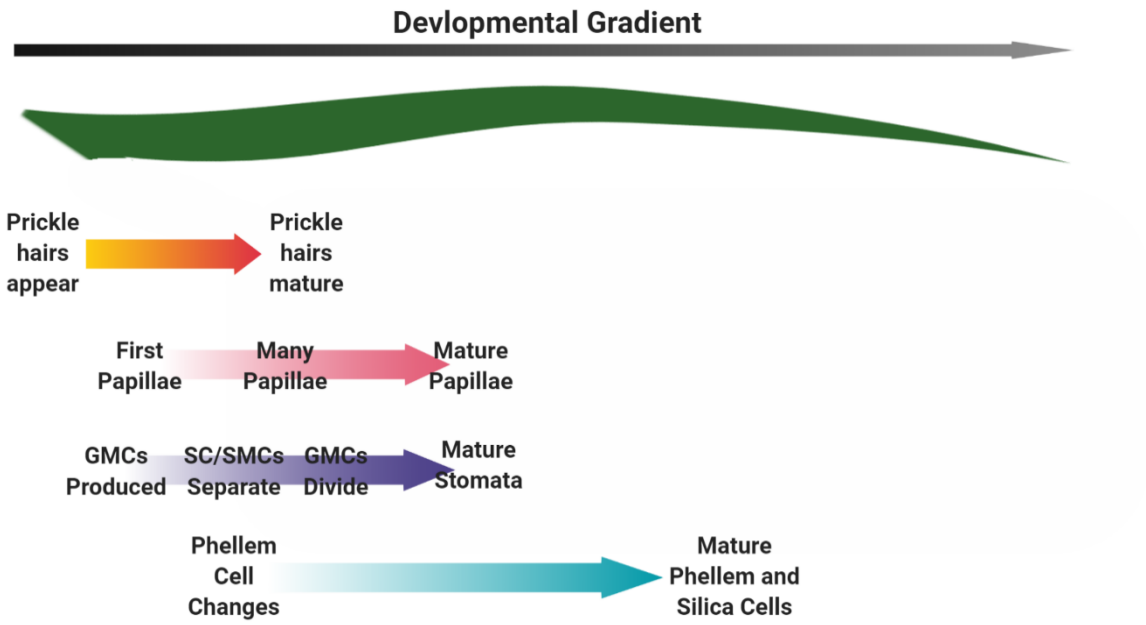


Figure 2.3 – Development of epidermal features along an immature rice leaf blade
Schematic representation of the work published by Luo and colleagues (Luo *et al*, 2012). Points are represented to scale along an "immature leaf blade" (assumed to be a P4-12 leaf, 12cm in length).
GMC – Guard Mother Cell; SC – Subsidiary Cell; SMC – Subsidiary Mother Cell
In this paper, "phellem cells" refers to the epidermal cells in between each silica cell in the silica ladder (comprising bodies of silica between two walls running the length of the leaf).
Figure created with biorender.com.

2.1.4 Aims

As discussed, although there is literature that documents spatial development of structures along a maturing leaf blade, to date there are no in-depth studies examining temporal leaf development in rice. In this chapter I aim to fill this knowledge gap by documenting early rice leaf development using staged, dissected primordia, examining the patterning of leaf structures both qualitatively and quantitatively. This structural atlas of early rice leaf development will provide the anchor for analysis of metabolic events during early rice leaf development described later in this thesis. Previous work in our laboratory has focussed on the P3/P4 transition and highlighted this as a key stage in the establishment of

photosynthesis. Thus, I intend to investigate this transition in depth to clarify the epidermal changes in this time frame.

Thus, the aims of this chapter are:

1. To generate a high-quality image set documenting early rice leaf development in IR64, focusing on the P3/P4 transition
2. To create a quantifiable index to allow for the developmental staging of young rice primordia
3. To compare early leaf development in IR64 to other varieties of domestic rice, as well as wild *Oryza* species

2.2 Methods and Materials

2.2.1 Plant Growth and Material

2.2.1.1 Varieties Used

Domestic Asian varieties used were *Oryza sativa* ssp. *indica* var. IR64 and *Oryza sativa* ssp. *japonica* var. Nipponbare, both of which were available in-house at Sheffield.
Wild varieties *Oryza latifolia*, *Oryza meridionalis, Oryza brachyantha*, and *Oryza barthii* were provided by IRRI or CIAT.

2.2.1.2 Germination

All seeds were germinated submerged in water on filter paper in deep Petri dishes. Seeds were incubated in a SANYO growth cabinet, on a 12h 26°C / 12h 24°C light/dark cycle, PAR 2000µmol^{-2}s^{-1} for several days (until the emergence of leaf 2).

2.2.1.3 Hydroponic Growth

The hydroponic system was maintained in a Conviron BDR16 growth chamber. Conditions were maintained at 28°C and 60% humidity on a 12h/12h day/night cycle, with the light set at 700µmol^{-2}s^{-1}, (around 560µmol^{-2}s^{-1} at seedling level). CO_2 was maintained at ambient, which averaged at around 480ppm. The system consisted of a 6.5L opaque plastic container, filled with 5L of hydroponic growth media (Table 2.2). The level of media was maintained at 5L with water and replenished every 1-2 weeks. Seedlings were held in microfuge tubes with the bottoms removed, placed into polystyrene racks (shown in Figure 2.4).

Table 2.2 – Compounds used for hydroponic growth of rice seedlings.

Media adapted from (Murchie et al., 2005) by previous lab members (van Campen, Thesis, 2016).

Chemical	Working Concentration (mM)
NH_4NO_3	1.4mM
$NaH_2PO_4 \cdot 2H_2O$	0.6mM
K_2SO_4	0.5mM
$MgSO_4$	0.8mM
$MnCl_2 \cdot 4H_2O$	0.009mM
$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$	0.001mM
H_3BO_3	0.037mM
$CuSO_4 \cdot 5H_2O$	0.003mM
$ZnSO_4 \cdot 7H_2O$	0.00075mM
Fe-EDTA	0.07mM

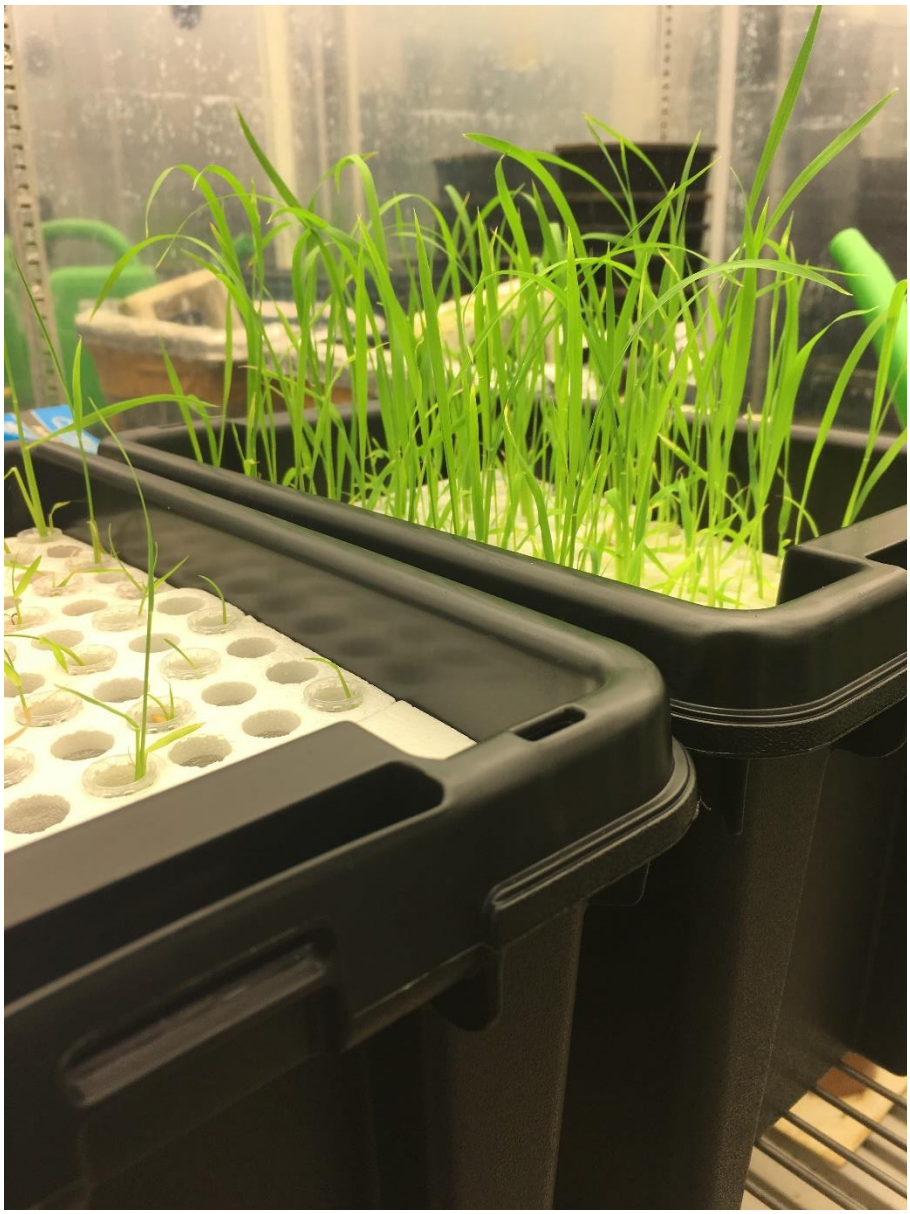


Figure 2.4 – Setup for hydroponic growth

2.2.2 Scanning Electron Microscopy

2.2.2.1 Sample Collection

Leaf 5 primordia were staged non-invasively using the plastochron index (Table 2.3). Primordia were dissected carefully with 25G and 30G hypodermic needles using a Leica MS5 dissection microscope and stored in water until imaging on the same day. Mature leaves were removed from the plant immediately prior to imaging.

Table 2.3 – Plastochron index for non-invasive staging of leaf 5

Leaf 3 length is used as a proxy for leaf 5 developmental stage. Data produced by Suppathra Narawatthana and Julia van Campen.

Leaf 5 Plastochron Stage	Leaf 3 Length (mm)
P1	4-20
P2	25-70
P3	75-110
P4	120-140
P5	150+

2.2.2.2 Imaging

Images were taken using a Hitachi TM3030 Plus Benchtop Scanning Electron Microscope. Images were captured in 15kV standard mode using secondary electron (SE) detection. Samples were blotted to remove excess water and then affixed to stubs using double-sided plastic tabs. All samples were imaged at -20°C using a cooling stage. Samples were scanned in 'fast' mode to find areas of interest and imaged in 'slow' mode. Care was taken to spend minimal time observing the area that would be photographed to prevent unnecessary degradation of the sample.

2.2.2.3 Image Analysis

All measurements were made using ImageJ Version 1.52a (Schindelin et al., 2015). Graphs were created using ggplot2 in R (R Core Team, 2019; RStudio Team, 2016; Wickham, 2016) (Versions: R – 3.6.2; RStudio – 1.1.453; ggplot2 – 3.2.1)).

2.3 Results

2.3.1 Imaging the Epidermis of Rice Leaf Primordia

Oryza sativa ssp. *indica* var. IR64 was selected for initial experiments. Unless otherwise stated, all images were taken of leaf five, as this is the first adult leaf produced. To ensure that epidermal features would be easily identifiable using this SEM imaging technique, images were taken of mature leaf blades and examined for the features previously identified in the literature (Figure 2.5A). It was decided to focus on stomata (Figure 2.5B) and two distinct types of trichome – macrohairs (Figure 2.5D) and prickle hairs (Figure 2.5E), in addition to looking at the general patterning of silica in the epidermis and the organisation of epidermal cells. Silica ladders (Figure 2.5C), while visible in mature leaf blades, develop fairly late, thus are not a suitable marker for staging in early development.

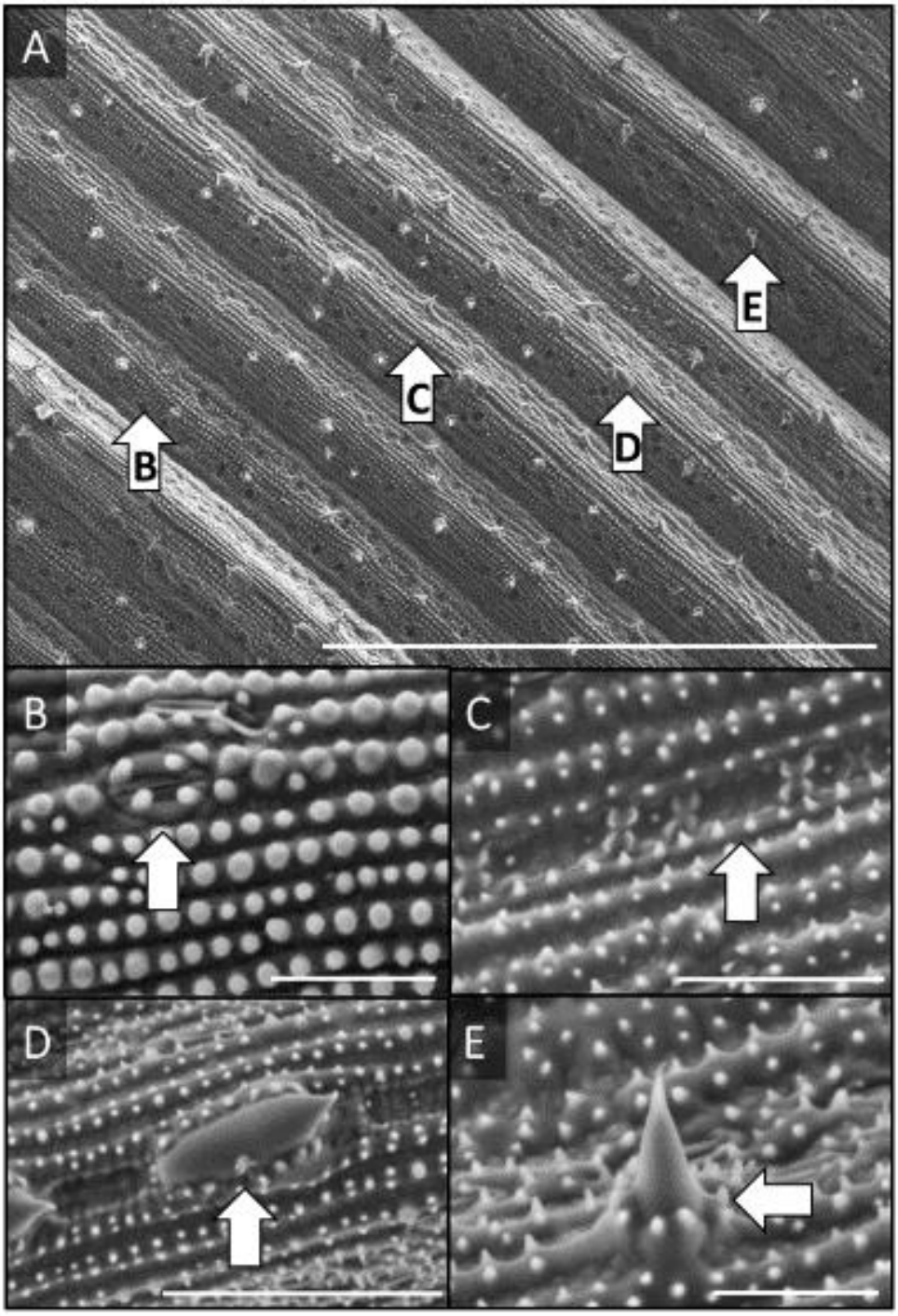


Figure 2.5 – Surface structures of mature leaf blades

Representative images of rice mature leaf blades.

A – Adaxial surface of a mature leaf blade

B – Stomatal complex

C – Silica ladder

D – Macrohair

E – Prickle hair

Scale bars: A – 1mm; B and E - 30µm; C - 50µm; D - 100µm.

Both P2 (Figure 2.6A) and P3 (Figure 2.6C) primordia have undifferentiated epidermal cells. Images of primordia at early stages of development were also taken to document the transition from the undifferentiated, differently organised cells characteristic of a P2 primordium (Figure 2.6A, B), to the undifferentiated but organised cells that in P3 can be seen starting to form the characteristic epidermal cell files (Figure 2.6D).

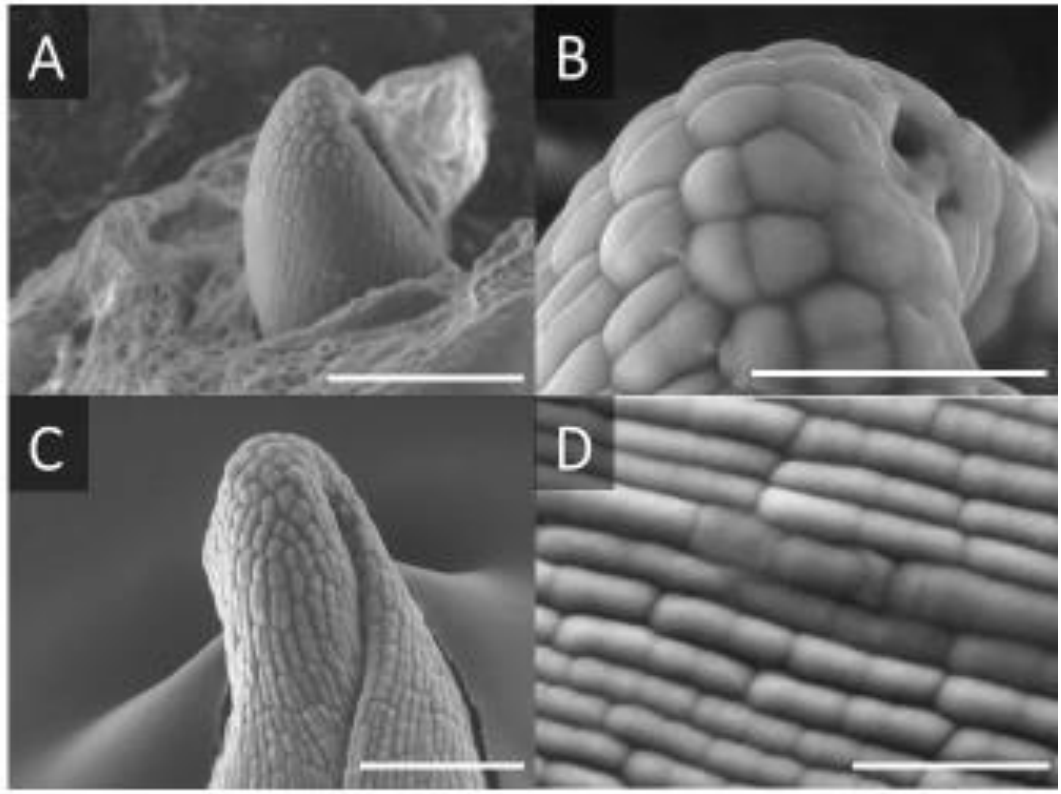


Figure 2.6 – Epidermal cell organisation in P2 and P3 primordia

Representative images of a P2 primordium, where cells are largely unorganised, and a P3 primordium, where cells begin to take on the characteristic file structure seen in older leaves.

A – A complete P2 primordium

B - Unorganised cells at the tip of the P2 primordium

C – A P3 primordium

D – Organised cells on the P3 primordium

Scale bars are 100µm (A and C) and 30µm (B and D).

As the P3 to P4 transition stage has been shown to be the stage at which photosynthetic competence is established, I was particularly interested in documenting the anatomical changes that occur during this time. As rice leaves develop basipetally, I decided to image growing primordia tips, as this is where we would first see the emergence of 'new' epidermal surface features as the plant ages (Figure 2.7). For IR64, the first feature we see differentiate from a P3 primordium (Figure 2.7A) is an immature trichome (Figure 2.7B), which will go on to form a macrohair. Immature stomata then begin to appear (Figure 2.7C),

followed by smaller immature trichomes which will go on to form prickle hairs. By P4-1, all epidermal features appear largely mature at the tip of the leaf, and the silica patterning is evident on all cells (Figure 2.7D).

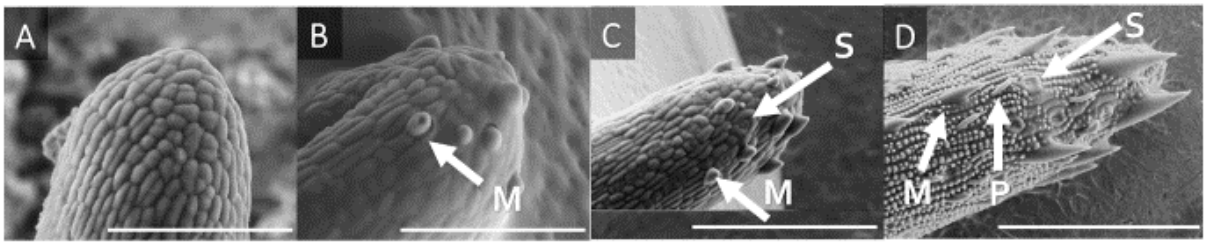


Figure 2.7 – Developing P3 and young P4 primordia

Representative images of primordia as they age (youngest left, oldest right). Features are labelled as they develop (M – macrohair, S – stomata, P-prickle hair).

Scale bars are 100µm (A and B) and 200µm (C and D).

2.3.2 A Tip Development Index

Having established a reliable method to visualise and distinguish different epidermal features on the rice leaf surface, I proceeded to develop a quantifiable index to allow staging of leaves based on these epidermal features. This was to enable a correlation of primordium length with the features present on the epidermis of the leaf. For each primordium used in this index, two images were needed: the full primordium in order to measure the length, and a high-resolution image of the primordium tip to distinguish which features had developed (example images shown in Figure 2.8).

Length was selected to act as a proxy for age for several reasons. Firstly, while hydroponic growth encourages the most homogenous growth possible, the plants still do not grow at identical rates, thus age in terms of absolute time does not necessarily equate to developmental stage. Secondly, the P3/P4 transition is very rapid, making it hard to accurately date the plants in real time using the plastochron index described in Table 2.3. Thirdly, as primordium length increases with cell divisions and elongations, the total primordium length is to some degree indicative of the cell division activity in the leaf.

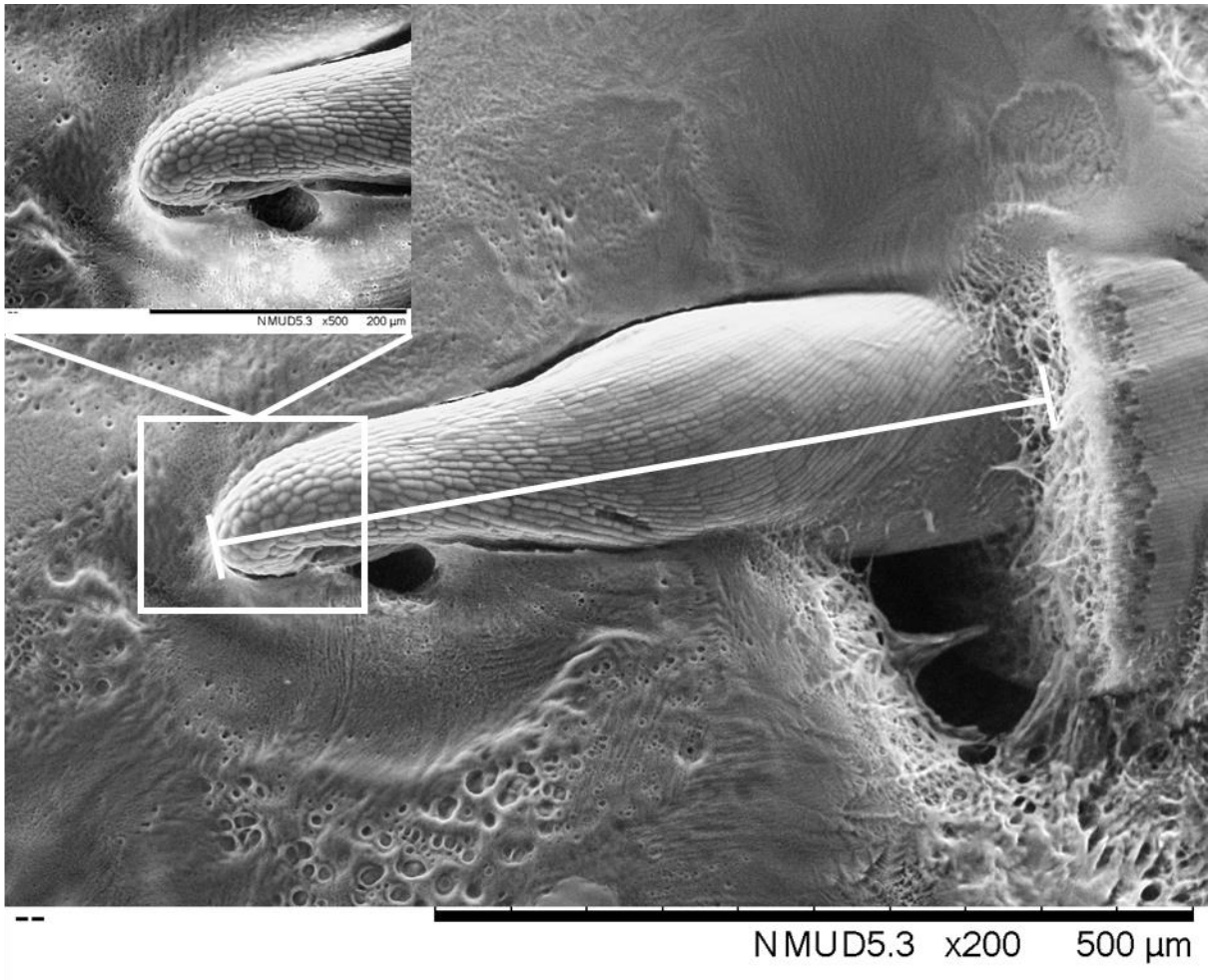


Figure 2.8 – Measurements needed for a developmental index

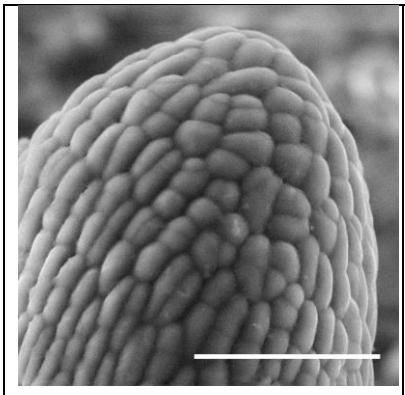
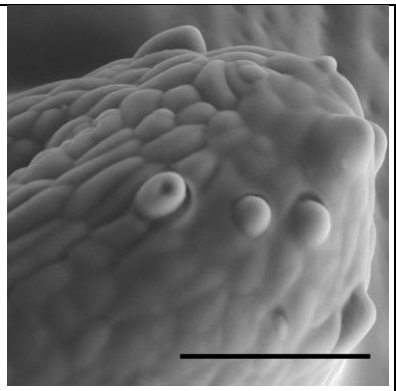
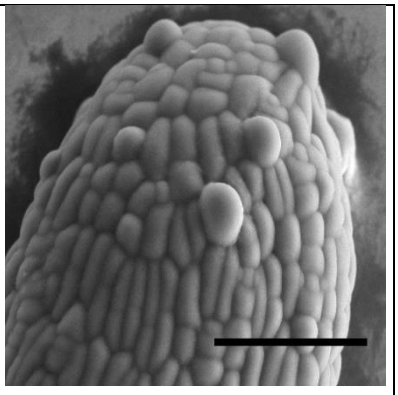
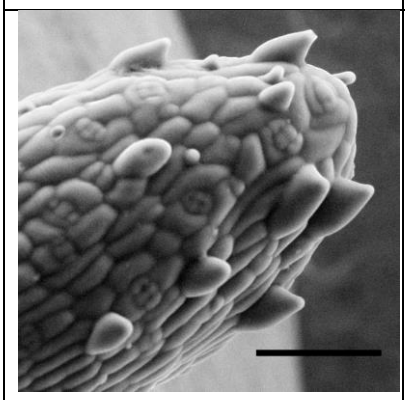
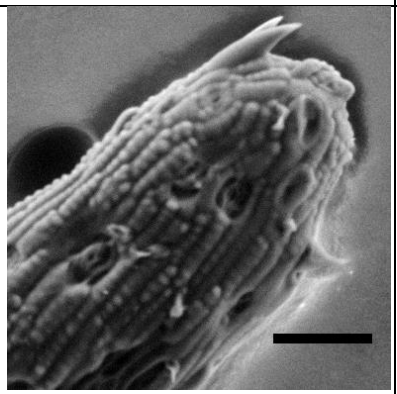
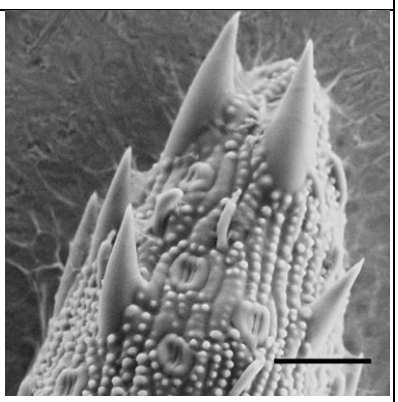
Measurements indicated show total primordia length (shown with a line) and a higher resolution image of the tip (shown with a box).

By looking at the order that epidermal features emerge on the developing primordium, it was possible to break the P3/P4 transition down into six distinct stages (termed I to VI), as shown in Table 2.4, each distinguished by the presence or absence of the various epidermal features described earlier, in addition to the overall organisation of the epidermal cells. To distinctly categorise primordia, I used the appearance of each epidermal cell type as a cut-off. It is already known that at the P3 ("Grade I") stage, no differentiated specialised cells such as guard cells or trichomes can be distinguished. The first cell type to differentiate is a kind of trichome, which in time goes on to form macrohairs, so primordia exhibiting only these were termed Grade II, Primordia were classed as Grade III if immature stomata with visible subsidiary cells were discernible in addition to these macrohair precursors. The appearance of a second type of trichome would result in a Grade IV classification. Trichomes that go on to form prickle hairs are easily distinguishable from those which appear at Grade II, as they are much smaller, round in shape, and seemingly randomly

distributed over the epidermal surface. Finally, the emergence of the characteristic silica patterning on the pavement cells resulted in a Grade V classification. At P4-1 (Grade VI), all epidermal features are mature, and the pavement cells have the distinct silica patterning seen on mature rice leaves.

Table 2.4 – Grades used for creating a developmental index

All images are of *Oryza sativa* var. IR64. Scale bars are 50µm.

		
Grade I ('P3') - Some cells unorganised, cell files beginning to form - No epidermal features visible	**Grade II** - Cells mostly in files - First trichomes appear (immature macrohairs)	**Grade III** - Macrohairs begin to mature - Immature stomata appear
		
Grade IV - Macrohairs continue to mature - Stomata begin to mature - Second trichomes appear (immature prickle hairs)	**Grade V** - First silica bumps appear on cells - Macrohairs largely mature - Stomata largely mature - Prickle hairs begin to mature	**Grade VI ('Young P4')** - Silica bumps fully visible on all epidermal cells - All epidermal features are mature, complete with normal patterning of silica

By combining this information on primordia length and tip grade, it was possible to create an index to quantify this *de novo* leaf development (Figure 2.9). For each designated developmental stage, the range of lengths observed is shown, and in each case there is very little overlap, with stages I, II, IV and VI differing significantly. While stage III is only significantly smaller than stage VI, this category is the smallest in this dataset (n=3), and the Bonferroni-adjusted p-value defines significance strictly by $p<0.005$. While non-significant, the grade III average does appear different from grades I ($p = 0.02$) and IV ($p = 0.02$), while visually overlapping with II ($p = 0.13$).

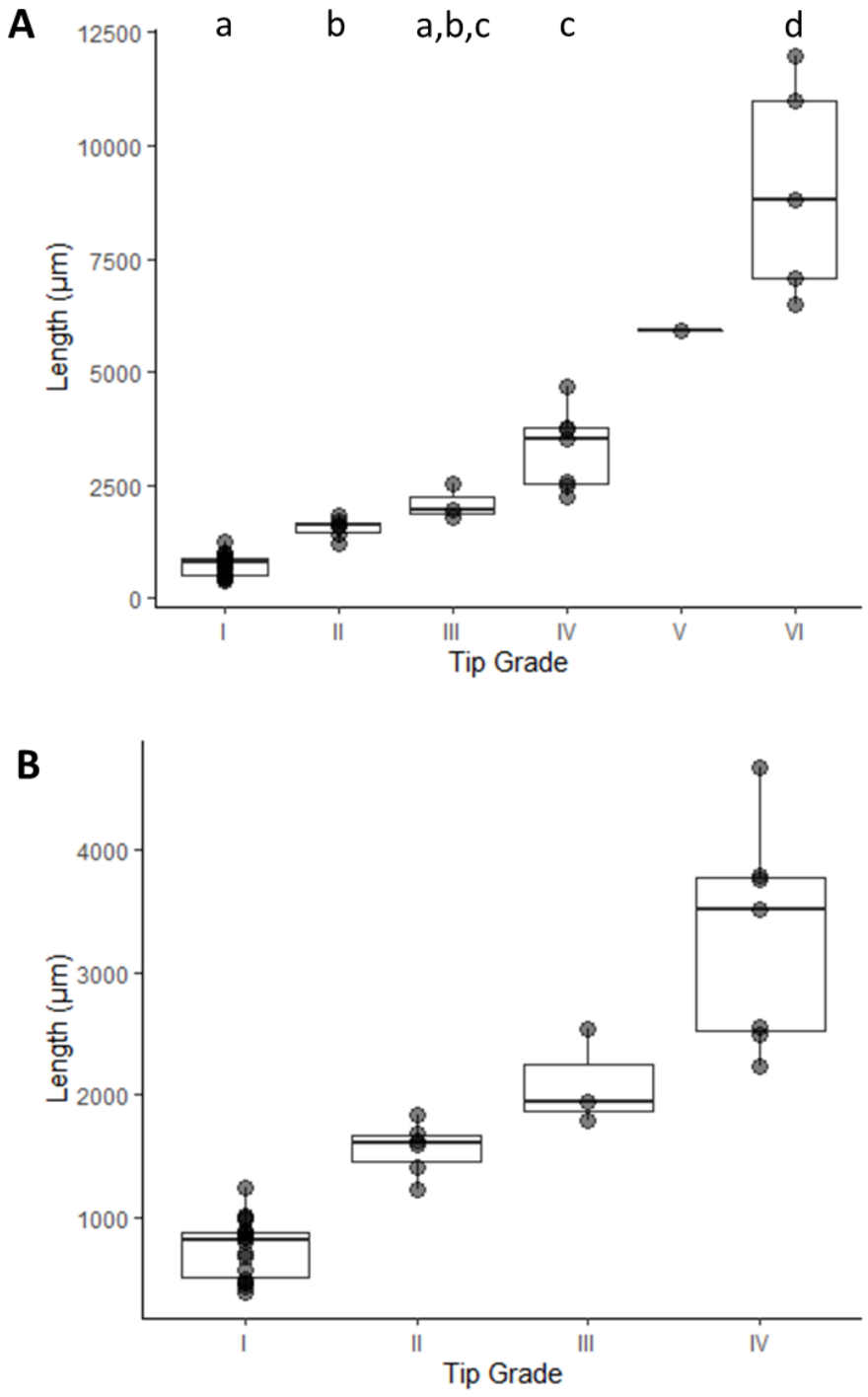


Figure 2.9 – A tip development index for IR64

A – Data shown for all primordia between P3 and P4-1. Significant differences shown by letters, as determined by t-test after correction for multiple comparisons ($p < 0.005$).

B – Data shown only for grades I to IV, where development is more rapid.

Each circle represents an individual dissected primordium. Boxplots show the first quartile, mean and third quartile. Total n = 40 (I to VI), 34 (I to IV).

Having established a framework describing the patterning and appearance of epidermal features during early development of IR64 rice, I proceeded to investigate to what extent this framework held true for other cultivated rice. I therefore proceeded to perform a similar investigation using the *japonica* domestic line Nipponbare. The two major geographic subspecies, *indica* and *japonica* are morphologically and genetically distinct (Khush, 1997), such that their hybrids are usually sterile. Both races are widely cultivated, and both IR64 and Nipponbare are commonly used in rice research, therefore a comparison of the two sub-species would enable an insight into the conservation (or otherwise) of early leaf development in these two core rice lines. Using the same approach as for IR64, primordia length was correlated with the epidermal features to generate a similar index for Nipponbare (Figure 2.10A). A smaller quantity of primordia were imaged for Nipponbare than IR64, thus grade II is omitted from this figure as no primordia imaged matched this description. Compared to IR64, there is more overlap between stages, as shown by overlap between the III, IV and V primordia, however this dataset does contain fewer primordia overall. When comparing the size of IR64 and Nipponbare primordia (Figure 2.10B), it does appear that Nipponbare primordia area slightly smaller, but this is not statistically significant.

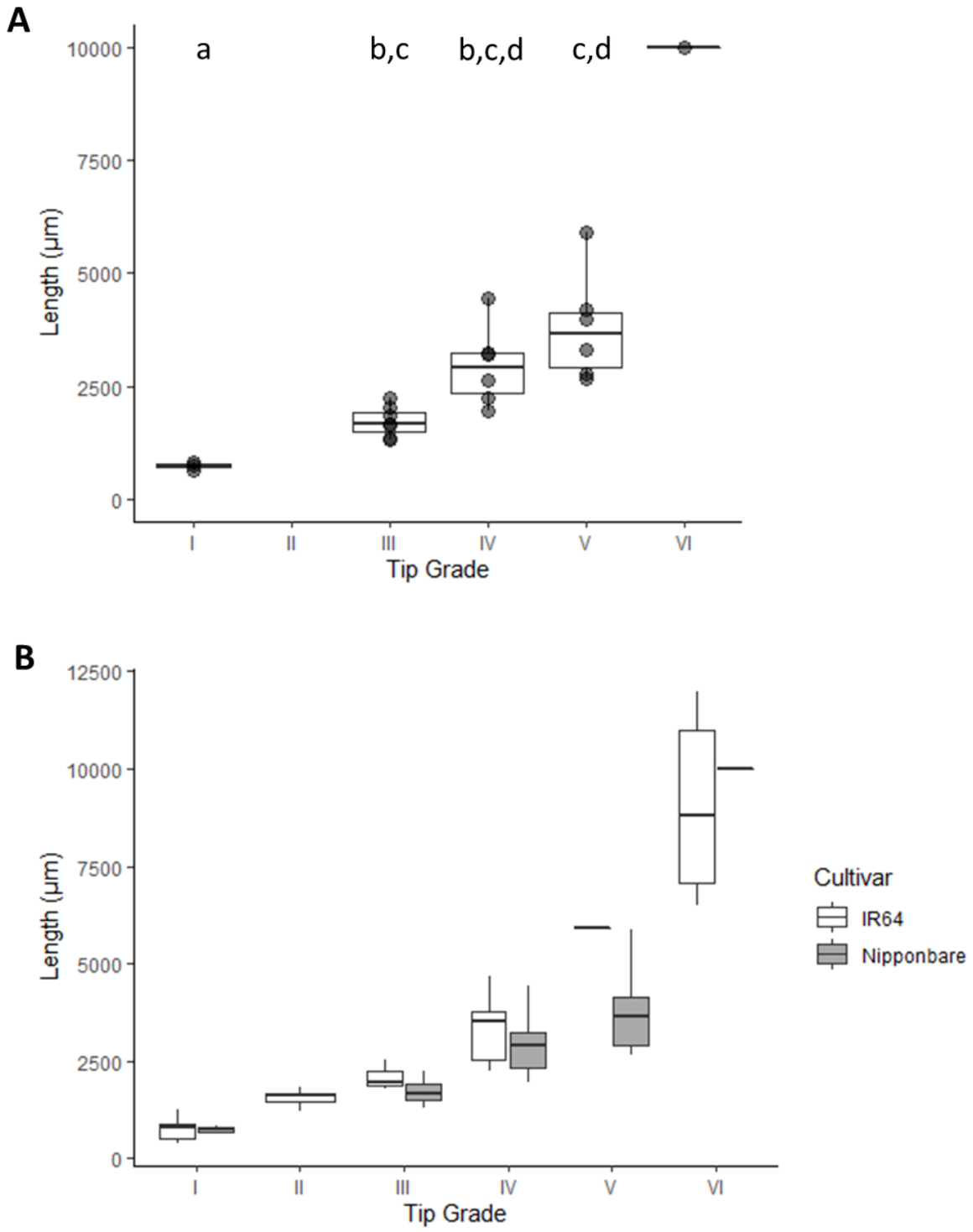


Figure 2.10 – A tip developmental index for Nipponbare

A – data shown for all Nipponbare primordia. Grade II omitted as no plants were dissected matching this grade. Significant differences shown by letters, as determined by t-test after correction for multiple comparisons ($p < 0.008$). Each circle represents an individual dissected primordium. Boxplots show the first quartile, mean and third quartile. Total n = 23.

B – comparison of IR64 and Nipponbare primordia. Boxplots show first quartile, mean and third quartile. For groups with one datapoint (IR64 V, Nipponbare VI), horizontal lines replace boxplots. Nipponbare grade II is omitted as no primordia were dissected matching this grade. Data as for Figure 2.9A and Figure 2.10A.

2.3.3 Comparison to Wild Lines

There is evidence for rice first being domesticated around 12,000 years ago (Breeding and Hall, 2007), however to this day various wild species remain important genetic resources for current domestic rice breeding programs. I was interested to see if the developmental pattern revealed in modern domesticated rice reflected an ancient conserved pattern or whether during domestication a shift in the pattern of early leaf development has occurred. To do this, I obtained a series of wild rice lines and performed a similar analysis of epidermal patterning in different stages of leaf primordia.

With this in mind, I compared the development of leaf primordia in two species of *Oryza sativa* (IR64 and Nipponbare), to *Oryza meridionalis* and *Oryza latifolia* (Figure 2.11). *Oryza latifolia* is a wild relative of rice found in South America that is known to be very stress tolerant. Compared to *O. sativa*, which has a diploid AA genome, *O. latifolia* has a tetraploid CCDD genome (Ge et al., 1999), thus representing at least 6.3 million years of evolutionary divergence and a hybridisation event between the two. *Oryza meridionalis* is a heat-tolerant wild rice first identified in Northern Australia, that shares the AA genome with *Oryza sativa*, and is therefore a promising candidate in some breeding programs.

Despite notable differences in the appearance of the mature plants, both *Oryza meridionalis* and *Oryza latifolia* appear to reflect the characteristic temporal development described fully in this chapter shown in domestic rice (as shown for IR64 and Nipponbare), with a completely undifferentiated primordium (I) followed by initiation of macrohair-type trichomes (II), initiation of stomata (III), prickle-hair type trichomes (IV), then silica patterning only appearing later in leaf development when other epidermal organs have initiated and begun to mature. Areas where primordia appear damaged are largely due to deterioration of the soft tissue caused by the electron beam during imaging. Where images are omitted (*O. meridionalis* grade I and *O. latifolia* grade V), this is because no primordium of that age was observed during sample collection.

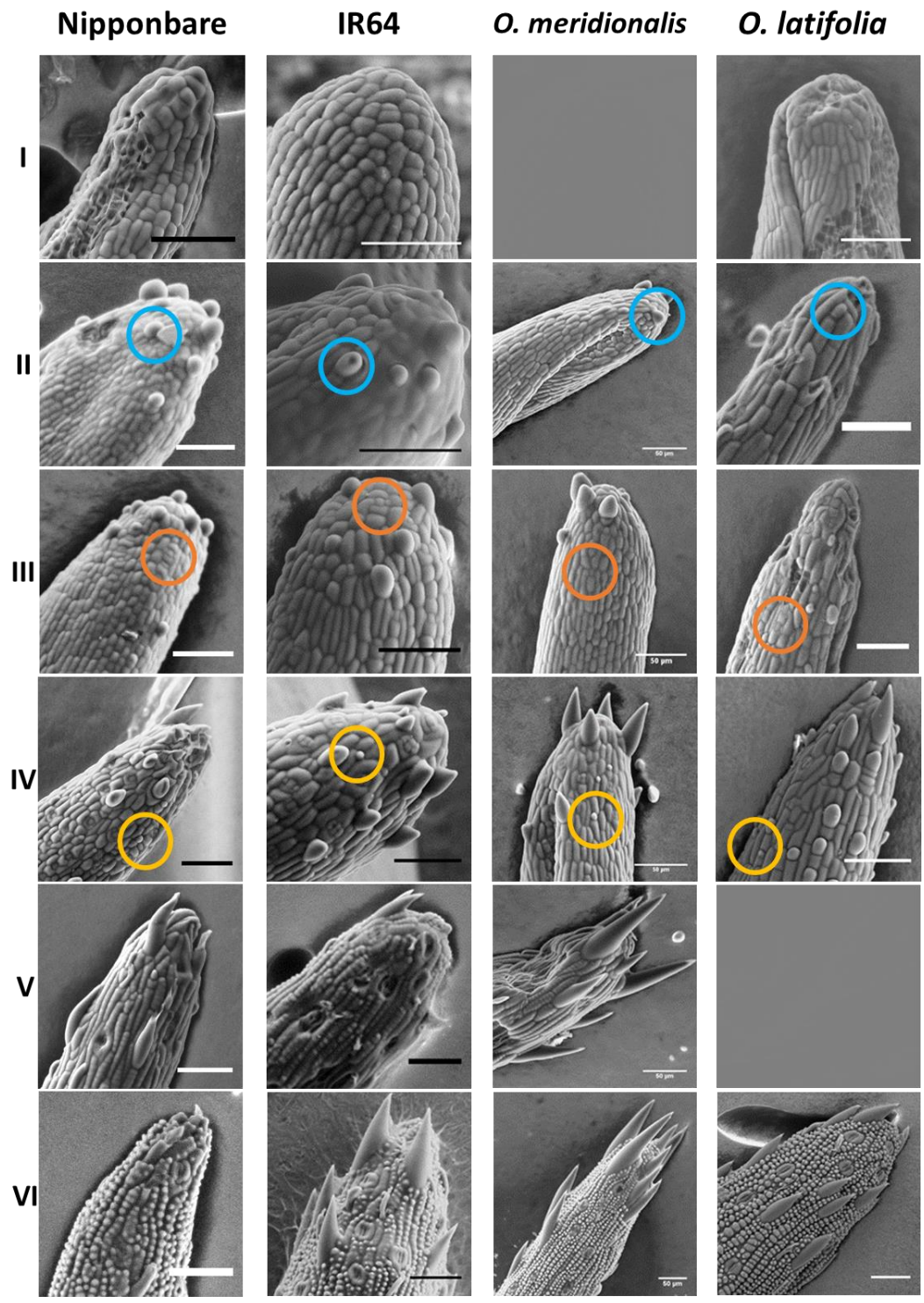


Figure 2.11 – Leaf primordia epidermal patterning in domestic and wild rice

Shown are example images for grades I-VI for four varieties of rice: *Oryza sativa* ssp. *japonica* var. Nipponbare, *Oryza sativa* ssp. *indica* var. IR64, *Oryza meridionalis* and *Oryza latifolia.* James Pitman contributed some images of *Oryza latifolia* primordia.

For panels II, III and IV, emerging features of macrohairs (blue circles), immature stomata (orange circles) and prickle hairs (yellow circles) have been highlighted.

All scale bars 50µm. Where images are omitted, a primordium of that grade was not dissected from the available specimens.

While similar primordia dissections of all available wild lines were not performed, images of the abaxial and adaxial surfaces of *Oryza barthii* (AA) and *Oryza brachyantha* (FF) revealed similar mature epidermal patterns to wild-type rice (shown in Figure 2.12), as shown by representative images of the abaxial and adaxial surfaces, with the only notable difference being the absence of macrohairs on the abaxial surface of *Oryza barthii* and the apparent absence of macrohairs in *Oryza latifolia*. However, absence of these features from the images does not mean that they are completely absent in the plants. Macrohairs have been previously documented as present on the adaxial surface of *Oryza latifolia* (Sánchez et al., 2003), and can be seen above in Figure 2.12 on developing *Oryza latifolia* primordia. Thus, while we do not know if the temporal developmental trajectory in these additional wild lines reflects the conserved pattern seen in *Oryza sativa*, *Oryza meridionalis* and *Oryza latifolia*, it does appear that the final mature leaf epidermis is very similar.

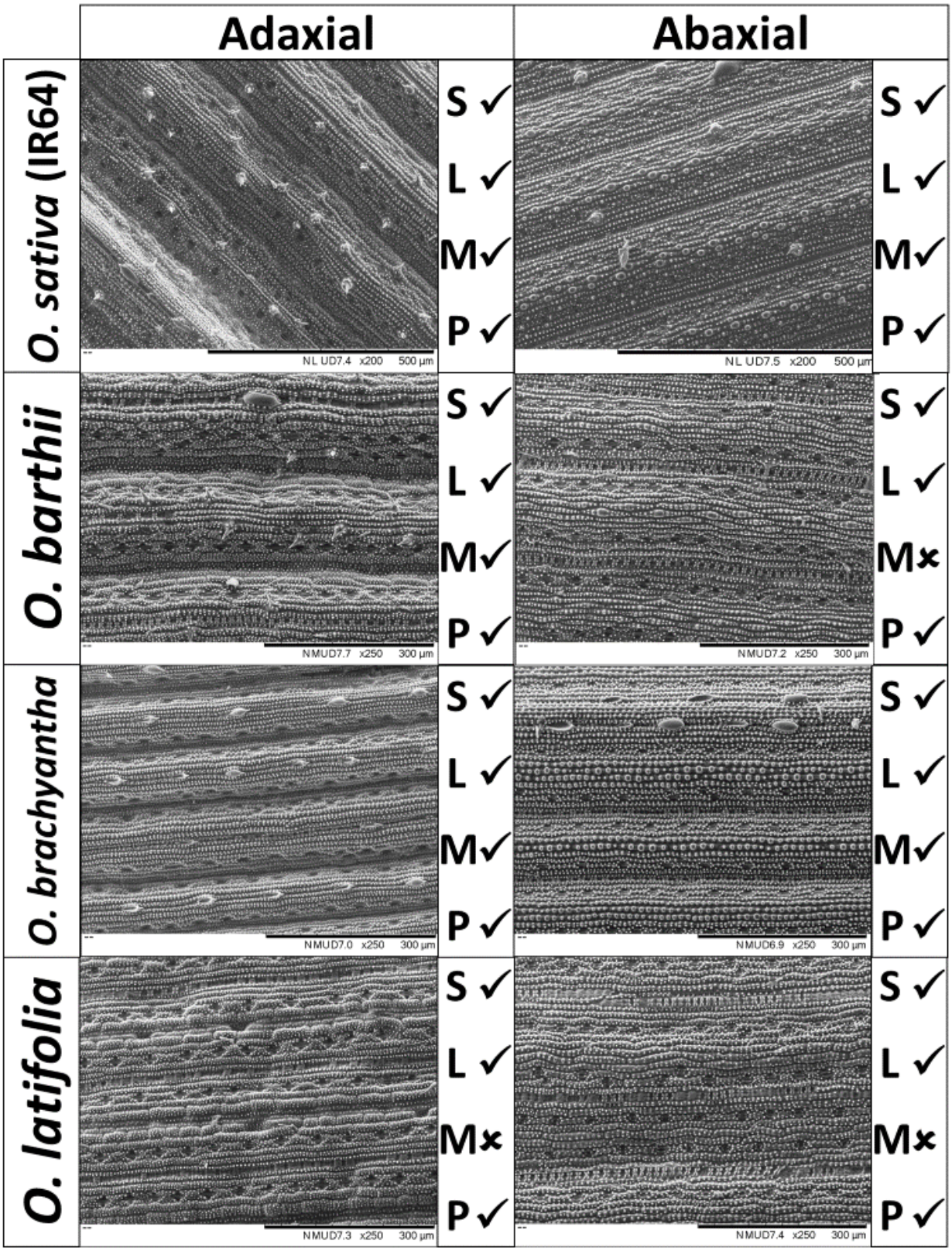


Figure 2.12 – Mature epidermal patterns in domestic and wild rice

Representative images for the abaxial and adaxial surfaces are shown for each line. Key indicates if certain epidermal features can be seen on this leaf surface. S – stomata; L – silica ladders; M – macrohairs; P – prickle hairs.

Images for *O. barthii, O. Brachyantha* and *O. latifolia* taken by Hannah Jones.

2.4 Discussion

In this chapter, I set out to both document and quantify the developmental transition between P3 and P4-1 leaves. The overall anatomy of a developing leaf has been previously documented for Nipponbare (a *japonica* variety; Itoh *et al.*, 2005; Kusumi *et al.*, 2010), with a more in-depth look at the overall early development of IR64 (an *indica* variety) performed previously in our lab (van Campen, 2016, thesis). This work also identified the P3 to P4-1 transition as a key developmental stage where photosynthetic competence is established (van Campen et al., 2016), hence the narrow focus in this chapter.

To date, most studies of rice leaf development have exploited the developmental gradient present along a maturing leaf blade from undifferentiated to mature cells, or affected development non-invasively while the plant is still growing. While this is both practical and provides much information, we cannot assume that *de novo* development of a leaf follows the same trajectory as a maturing leaf, as discussed earlier. Here I have produced a comprehensive dataset documenting *de novo* leaf epidermal development in rice. This high-quality image dataset documenting early rice leaf development is a novel addition to the field. While SEM has been used to obtain clear images of each discrete developmental stage across the whole lifetime of a rice leaf (P1-P6) (Itoh et al., 2005), there are no currently published studies examining the transitions between these stages in rice, nor any with a focus on the early stages described here.

When comparing data produced in this chapter on primordium differentiation with studies that consider a developmental gradient along an established P4 leaf, we can see that the sequence of epidermal cell differentiation differs slightly (Figure 2.13). In particular, in formed leaves it appears that prickle hairs mature before stomata, whereas in developing primordia the stomata emerge prior to the prickle hairs (although they do mature concurrently with macrohairs, another form of trichome). Further, papillae are seen much earlier in the developmental gradient along a mature blade than in a developing primordium relative to other epidermal features. There are, however, some important caveats in this comparison. For example, in the work performed by Luo and colleagues it is unclear whether the data refer to leaf 3, 5, 7 or the flag leaf. As it is known there are anatomical differences between juvenile leaves (leaves 1-3) and adult leaves (leaf 5 and older) in rice (Sylvester et al., 2001), it is entirely plausible that the developmental trajectory of a leaf depends on the point in development it is initiated. In this chapter, all samples imaged were leaf 5 primordia. Leaf 5 is the first adult leaf in rice (Xie et al., 2012), and shows a reliable, typical growth pattern (van Campen, thesis, 2016). Secondly, the method used by Luo *et al* involved producing epidermal impressions to obtain images of the leaf surface, which may not allow for an

entirely accurate replication of the epidermal surface. In contrast, the SEM approach taken in this chapter involved minimal tissue preparation, while maintaining the higher resolution offered by this microscope, to provide a reasonably direct view of the epidermal topography. Nevertheless, the overarching epidermal patterns described in the existing literature are also seen when examining established leaf blades using the SEM technique described in this chapter.

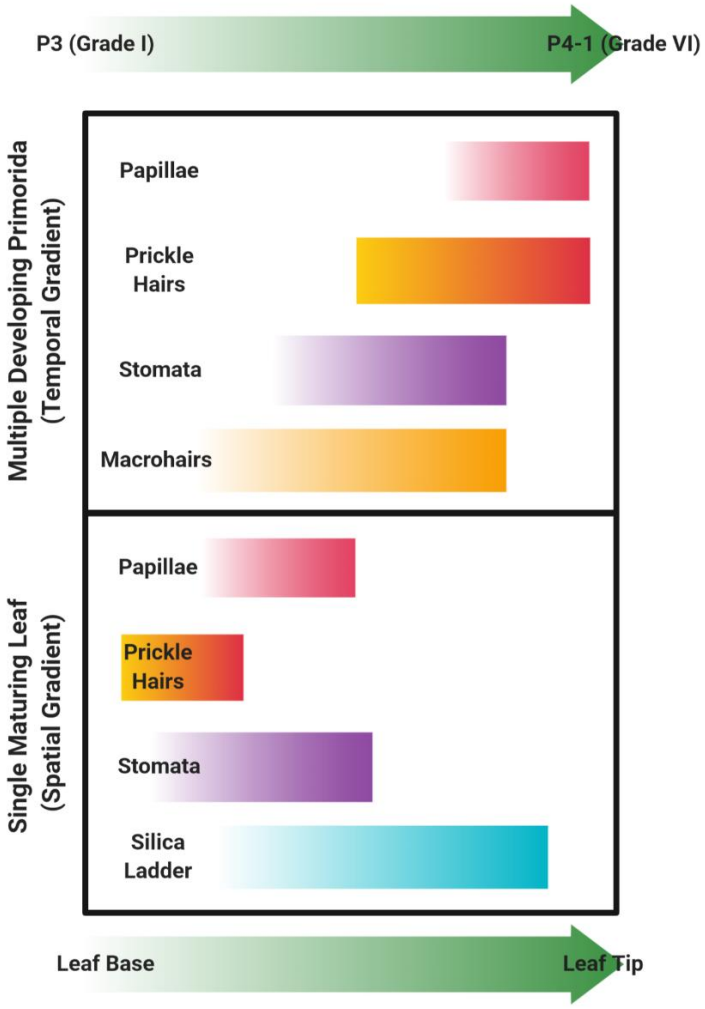


Figure 2.13 – Comparison between development along a P5 blade compared to *de novo* primordium formation

"Multiple developing primordia" shows differentiation of epidermal structures in time on increasingly mature primordia from youngest (top) to oldest (bottom).

"Single maturing leaf" shows differentiation of epidermal structures along a developing leaf blade from leaf base (top) to leaf tip (bottom), based on Luo *et al.*, 2012.

Figure made with biorender.com.

It was already known that the cells at the P3 stage are pluripotent, but lose this trait by P4 as the cells differentiate (Itoh et al., 2005). The results reported in this chapter establish a more

detailed temporal analysis of epidermal differentiation in domestic rice, which seems to reflect the pattern observed during the early leaf development of a variety of wild rice species. This suggests a highly conserved pattern of early leaf differentiation. Engineering of rice leaf structure has emerged as a key aim for global initiatives aimed at improving photosynthesis. In this context, having a reliable atlas of what “normal” early leaf development looks like will be a useful tool to allow assessment of any altered development resulting from manipulations designed to achieve this aim. In a broader context, understanding the mechanism regulating the rate of differentiation during the flux of cells from a dividing to a non-dividing state remains a major challenge in biology. Creating a standard atlas of the changes that occur along the surface of the differentiating rice primordium with time provides a scaffold against which genetic regulators of this process might be identified and temporally mapped.

It is also interesting to consider whether the development of leaf ultrastructures is evolutionarily conserved. When examining wild rice species, Ethel Sanchéz from the University of Costa Rica has documented the ultrastructure of mature leaves of *Oryza grandiglumis* (Sánchez et al., 2006) and *Oryza glumaepatula* (Sánchez and Espinoza, 2005), the latter of which was previously considered to be the Latin American race of *Oryza rufipogon*. While both lines show some differences to *Oryza sativa*, these are mainly in silica patterning, suggesting the leaves ultimately have similar cell-level structures despite very different gross morphologies.

The apparent conservation of epidermal patterning in *de novo* developing leaves between *Oryza* species of different genomic groups implies this differentiation pattern has been conserved throughout the evolution of the genus. Given more time, and more wild rice lines, it would be interesting to see if this apparent conservation of epidermal patterning continues across other species, particularly for the other cultivated rice species *Oryza glaberrima*, which is predominantly grown in Africa. If it is conserved, then the information gleaned here could aid in guiding alterations to leaf development in this species too. Interestingly, large trichomes also emerge as the first visible epidermal structure in SEM studies of leaf development in a variety of dicot species (Hagemann and Gleissberg, 1996; Poethig, 1997; Sylvester et al., 1996), reflecting the similarities between dicot and monocot development. As it becomes clearer that there are a large number of conserved processes in leaf development between monocots and dicots (Nelissen et al., 2016), it follows that the *de novo* patterning of these highly related rice species would be similar, even if final leaf shape and morphology differ greatly between them.

In addition to providing novel insight and a method to stage both normal and mutant

developing leaves, the data produced here will be a useful addition to other molecular studies of early leaf development in producing a developmental atlas.

Chapter 3:
Metabolite Profiling of Early Development of Leaf Primordia

3.1 Introduction

3.1.1 Why Use a Metabolomics Approach?

As discussed throughout this thesis, there are a variety of ways to measure rice leaf development. The formation of a functional leaf also requires the co-ordination of numerous biochemical and physiological processes by the plant, and no single technique can give a complete picture. Physiological measurements, such as chlorophyll fluorescence data (van Campen et al., 2016) are informative of the biological processes being carried out by the leaf and how these processes may change over time. Histological approaches can map development visually, either focusing on individual cell types (Luo et al., 2012) or overall patterning (this thesis, Chapter 2). There are also several approaches to investigating developmental processes at the molecular level, through examining transcripts, proteins or metabolites produced during development. RNAseq experiments to study gene expression (van Campen et al., 2016) have proven highly informative, however mRNA is not the final product of gene expression. Such transcriptomics data can be complemented by proteomics data, to show which proteins are being produced, and metabolomics data, to directly assess changes to metabolism (for a brief overview of how “omics” technologies are integrated, see Horgan and Kenny, 2011). As the metabolome represents the final, most downstream result of gene expression and protein activity, I was interested to see how the events of leaf development relate to metabolic state.

3.1.2 Introduction to Metabolomics

Metabolites are either intermediates or products of metabolism. They are chemically diverse small molecules, often 1200Da or less in size. Many classical biochemical techniques will measure one metabolite at a time, however modern metabolomics approaches allow for surveying a larger panel, or even the study of the global metabolite profile (or metabolome) of a given biological system. As metabolism rapidly changes with gene expression and interactions with the environment, metabolomics can be used to get a snapshot of what is

happening in a biological system at a specific point in time. There are several different applications for this technology with varying broadness and specificity, from untargeted metabolomic fingerprinting (analysing the pattern of metabolites in a given sample without identifying them), to metabolic profiling (looking specifically at a pre-determined set of compounds) to targeted analysis (where the procedure can be optimised to study a compound of interest) (Rochfort, 2005).

Metabolites are interesting not just as a reflection of what has already occurred in the system, but are directly responsible for the function and survival of the cell on an instantaneous basis. They also have the potential to affect future gene expression. Sugars, for example, are a core to primary plant metabolism, engrained as they are in photosynthesis and respiration. Not only is growth and development highly impacted by carbohydrate availability, but sugars themselves can also act as signalling molecules, directly affecting gene expression (Eveland and Jackson, 2012). Sucrose is utilised as a long-distance signal (Chiou and Bush, 1998), glucose is known to differentially regulate over 1000 genes in Arabidopsis (Price et al., 2004) and functioning trehalose metabolism has been shown to be vital for embryo maturation in Arabidopsis (Eastmond et al., 2002).

Often these studies focussing on only one metabolite of interest will use tailored biochemistry techniques to measure the metabolite directly, or use classical genetics approaches to examine the gene expression changes surrounding its presence. However, focusing on gene expression alone ignores the advantages of directly studying metabolites to gain an instant snapshot of cell function, and measuring only one metabolite still means that vital information can be missed.

To get a truly global metabolite profile of the system of interest, either mass spectrometry (MS) or nuclear magnetic resonance spectroscopy (NMR) techniques can be used (Alonso et al., 2015). NMR is used to observe atomic nuclei in a strong magnetic field. Not all nuclei will resonate using this technique, but several of biological relevance do, including ^{13}C and ^{15}N. The technique is used for a range of applications, including protein structure determination. MS is reliant not on the presence of specific nuclei, but on ionisation of the metabolites. Thus, the limitations for mass spectrometry experiments are the volatility of the compounds, their respective ionisation energies, and the ability to identify them.

3.1.3 Introduction to Mass Spectrometry

Mass spectrometry is a technology which allows for the study of the relative abundance of ions within a given sample by measuring their mass: charge (m/z) ratios. Multiple different kinds of mass spectrometer are available, but all require an ion source (to ionise the

sample), a mass analyser (to separate the ions by m/z) and an ion detector, as shown in Figure 3.1.

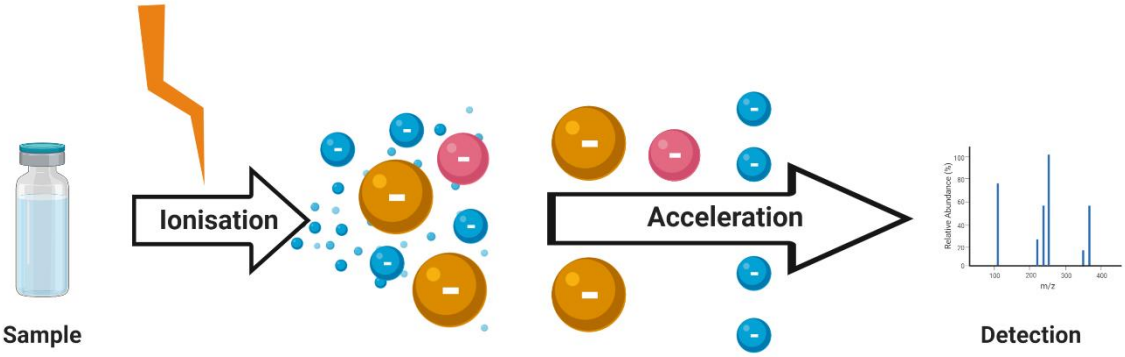


Figure 3.1 – The basic steps of mass spectrometry

An ion source is used to ionise the samples. These ions are then separated by mass to charge ratio (m/z) in a mass analyser, and detected, resulting in a spectrum showing count data for each ion.

Figure made with biorender.com.

A commonly used ion source is Electrospray Ionisation (ESI), which uses high voltages to create a charged aerosol from a liquid sample. ESI is deemed a “soft” ionisation technique that generally results in unfragmented ions, sometimes with multiple charges, especially on larger molecules (Banerjee and Mazumdar, 2012). To focus on a more specific mass range for ionisation, ESI can also be coupled to other technologies, for example high performance liquid chromatography (HPLC), for increased fractionation prior to ionisation. This ion source must then be paired with a mass analyser in order to separate ions by m/z. In the equipment used to collect the data described in this thesis, ESI is coupled to a time-of-flight (TOF) mass spectrometer, which accelerates the ions towards the detector. As flight time is proportional to energy, distance travelled and m/z, the mass of the ions can be easily calculated (Agilent Technologies, 2011): for a given energy (E) and distance (d), the mass (m) of the ion is directly proportionate to the flight time (t) squared (see Equation 3.1). As the energy and flight path (distance) are kept constant within a mass spectrometer, the time taken allows us to calculate the ion mass.

Equation 3.1

$$m = \left(\frac{2E}{d^2}\right) t^2$$

3.1.4 Applications of Metabolomics to Leaf Development

Metabolomics is a useful tool in plant science for studying various aspects of development and crop improvement. In species other than rice, various approaches have been taken to study leaf development, and its interaction with environmental factors.

Pick and colleagues have performed in-depth analysis of the maize leaf developmental gradient using a GC-MS (gas chromatography-mass spectrometry) metabolomics approach on leaf blades divided in to ten sections (Pick et al., 2011). In this work, 118 metabolites were grouped in to five distinct clusters which displayed different patterns along the spatial developmental gradient. This work comprehensively documented the normal progression of C_4 photosynthetic development along the maize leaf developmental gradient, combining transcriptomic and metabolomic data. However, even the youngest leaf sections included in this study have already differentiated into a characteristic Kranz anatomy, thus represent more mature tissue than the young primordia used in the experiments presented in this chapter.

Work performed in other species have taken alternative approaches to performing metabolomics on leaves of different ages. A study of secondary metabolism in yacón (a South American medicinal plant) sampled leaves from different points during whole plant development (split into five stages from “leaf development” to “flowering”), and suggested that leaf developmental stage is a regulator of secondary metabolism (Padilla-González et al., 2019). In an approach more similar to that taken in this chapter, work in date palms has demonstrated that “still developing” young leaves are potentially more resilient in the face of abiotic stress challenges than mature leaves (Du et al., 2019).

With respect to rice, due to the global and small farm-focussed nature of rice farming, it is predicted that more than forty thousand cultivars exist globally (The Rice Association, website), with over 130,000 accessions held in the International Rice Genebank (IRRI, website). Thus, there is sizeable chemical and metabolic diversity in rice, as reviewed in Kusano et al., 2015. Consequently, practical applications of metabolomics techniques in recent years have led to an exploration of factors affecting grain quality (Mumm et al., 2016), grain chalkiness (Lin et al., 2017a) and responses to both biotic (Parker et al., 2009) and abiotic (Lawas et al., 2019) stress. Metabolic trait locus mapping, which examines the association between genetic variation and changes in metabolite levels, has also been a useful tool for plant breeding (Okazaki and Saito, 2016; Wei et al., 2018).

However, while a popular tool for breeding improvements, comparatively less literature describes the use of metabolomics as a tool to study rice leaf development. Work by Wang

et al., 2014, which exploited the developmental gradient along individual maturing maize and rice leaves, presents data for the changing levels of 28 metabolites (identified through database searches) across 11 rice leaf sections, and directly compares this to a complementary maize metabolite profile. These data provides an interesting perspective of core differences between C_3 and C_4 metabolism between these two species with similar developmental profiles, and showcases how a subset of metabolites changes in relation to spatial regulation of development. However, as indicated by a growing body of work (reviewed in Nelissen et al., 2016), monocot leaf development is influenced by both spatial and temporal organisation. This chapter aims to add to the information about the temporal trajectory of rice leaf development.

As the leaf develops, cells shift from a pluripotent, undifferentiated state to their mature, differentiated form. This shift in cell identity is accompanied by changes in metabolism. Thus, not only does the metabolome reflect the current state of the cell, but evidence from studying animal embryos has revealed that metabolism can directly impact whether a stem cell proliferates or differentiates (reviewed in Shyh-Chang et al., 2013). In this chapter, I describe the work I have undertaken to document the global metabolome changes over the course of *de novo* leaf development in the rice variety IR64.

3.1.5 Aims

1. To use a global metabolomics approach to produce metabolomic fingerprints for P3, P4-1 and P5 staged primordia
2. To identify potential molecular markers unique to these stages
3. To establish whether P3/P4-1 transition intermediate primordia (as described in Chapter 2) can be distinguished by metabolomic fingerprinting
4. To identify metabolic pathways of potential interest for further experimentation

3.2 Methods and Materials

All metabolomics work was performed in the Faculty of Science biOMICS facility. Use of the mass spectrometer was performed under the supervision of Dr Heather Walker.

3.2.1 Plant Growth and Material

Oryza sativa ssp. *indica* var. IR64 were grown in hydroponics as detailed in Chapter 2. Primordia were staged while still attached to the plant by measuring under a dissecting microscope during dissection. Great care was taken to ensure that only leaf primordia tissue was sampled. Primordia were flash frozen instantly in liquid nitrogen after removal from the plant.

3.2.2 Metabolite Extractions

Staged, dissected primordia that had been flash frozen in liquid nitrogen, were freeze-dried for around six hours prior to extraction. Metabolites were extracted using the methanol-chloroform method (Walker, thesis, 2013). For all extractions, LC-MS grade solvents and distilled, deionised water were used. Samples were kept on ice and all centrifugation steps were refrigerated to 4°C.

Solvents used are documented in Table 3.1, with values shown for 1mg of tissue. Small primordia where an accurate mass could not be determined were extracted as 1mg of tissue.

Table 3.1 – Methanol-chloroform extraction solvents

Values shown are accurate for a 1X extraction.

Solvent Name	Composition	Ratio (v, v:v, v:v:v)	µL solvent per mg tissue
A	MeOH:$CHCl_3$:H_2O	2.5:1:1	10
B	MeOH:$CHCl_3$	1:1	5
C	H_2O	1	3.5
D	$CHCl_3$	1	2

For 1mg tissue, 10µL pre-chilled solvent A was added to the sample on ice, and the samples ground directly in the microfuge tubes with a micropestle to minimise sample loss. Samples were then vortexed for 10 seconds and left on ice for 5 minutes; then centrifuged at 14,000rpm for 2 minutes. Supernatant A was removed to a prechilled storage tube. The sample pellet was then re-extracted with 5µL solvent B. Samples were then vortexed and left on ice for 10 minutes; then centrifuged at 14,000rpm for 2 minutes at 4°C. Supernatant B was removed and added to supernatant A.

To separate the aqueous and chloroform phases, 3.5µL H_2O and 2µL $CHCl_3$ were added to

the supernatant mix; the sample was then centrifuged at 14,000rpm for 15 minutes at 4°C, resulting in two distinct phases. The aqueous and chloroform phases were then moved to separate tubes, and centrifuged again at 14,000rpm for 2 minutes at 4°C. Finally, the supernatants were removed to clean tubes for storage. Samples were stored at -80°C, and thawed on ice prior to use.

When extracting individual primordia, small 1mg samples were extracted directly at a 1 in 10 dilution, while larger samples were diluted post-extraction.

3.2.3 ESI-TOF-MS

Metabolite extracts were analysed using a Waters G2-Si mass spectrometer. For each sample, 3 repeats of 10µL aqueous phase sample were directly injected using a Waters Acquity UPLC. Acetonitrile was run before and between samples. Spectral data were collected in both positive and negative ionisation modes, with a spectral window of 50 to 1200Da at a scan rate of 1 scan per second. A leucine enkaphalin lock mass was run alongside the samples, but not corrected for, in order to check for drifts in mass measurement and accuracy over time.

3.2.4 Data Analysis

For each biological replicate, data were simplified by combining the three technical replicates, and binning masses to 0.2 Da using an in-house macro described in Overy et al., 2005. Multivariate analyses were performed using SIMCA (15.0.2) (Eriksson *et al*., 2006). For all multivariate analysis, % total ion count data were used. This normalisation step converts expression data for each bin as its percentage of the total ion count for that sample.

There are important considerations when interpreting the mass spectrometry data, particularly when assigning putative identities to ions of interest. Firstly, exact masses of compounds are calculated values, usually based on the most common isotope for that compound. Accurate mass refers to the measured quantity of that ion, and reflects the isotope constitution of the sample in addition to any errors during measurement. The discrepancy between calculated exact mass and observed accurate mass is known as mass measurement accuracy. The accuracy of the measurement tends to decrease for low intensity samples, as noisier spectra lead to larger errors in assignment. Higher purity samples, or peaks with a high ion intensity, are more easily to measure accurately, but there may still be slight differences in assigned m/z values. For example, in Supplemental Figure 3.1, we see that the citrate peak in the citrate standard is assigned at 191.02, while the citrate peak in the P5 sample is assigned as 191.05 (Figure S3.1A, B), despite the fact that both are confirmed as citrate by MS/MS in the same experiment.

Another consideration is whether the machine is assigning the correct values to these peaks. One way to do this is to run a lock mass – a known compound standard of a known mass – alongside the samples, to which experimental data can be compared or normalised. For these fingerprinting experiments, the lock mass compound run was leucine enkaphalin, which is a standard reference material used with Waters mass spectrometers. In negative ionisation mode, leucine enkaphalin m/z should read as 554.2693. However, in the experiment described in 3.3.2, it reads ~0.1Da higher than it should – for example reading 554.3668 in Figure S3.1F. While this shift does not change the overall spread of metabolites or the differences identified by multivariate analysis, it must be considered when assigning putative identities to m/z bins of interest.

In this chapter, masses were rounded (or "binned") to the nearest 0.2Da. Due to the mass shift error explained above, this can mean that we look for putative expression of certain metabolites in the "wrong" bin. For example, citrate appears to be highly expressed in the mature leaf tissue, and a peak would be expected at around 191.02 – 191.05 (as shown in Figure S3.1B). Instead, a strong peak is seen at 191.13 (as shown in Figure S3.1C). Adjusting for the 0.1Da shift that is known from the lock mass value, the actual m/z of this peak should read at around 191.03 – what we would expect from citrate. As expression values are rounded to the nearest 0.2Da, this shifts citrate expression from bin 191 to bin 191.2. This will not alter the assignment for all compounds, however. The strong peak in mature leaf tissue identified as malate by tandem mass spectrometry (MS/MS) reads as 133.02 (Figure S3.1D). However, in the previous fingerprinting experiment, a strong peak is observed at 133.08. In this case, the m/z increase does not change the assigned bin, as 133.08 still rounds down to 133, the bin of interest for malate.

3.2.5 MS/MS

M/z ratios of interest were putatively identified using PubChem (Kim et al., 2019), with searches narrowed down by monoisotopic exact mass, and exclusion of compounds based on their lack of involvement in plant biochemistry. Expected fragmentation patterns were available online using the METLIN database (Guijas et al., 2018). Where standards were run, these were at a concentration of $1 mgmL^{-1}$ in 1:1 v:v methanol:water. Samples were run as for 3.2.3, but by selecting an ion (to the nearest Da) to fragment. Samples were run for one minute, with 1 scan/second MS and MS/MS scans collected. Where possible, real standards were run alongside leaf samples so fragmentation patterns from the leaf tissue sample could be matched computationally and to the known standard metabolite.

3.3 Results

3.3.1 Untargeted Analysis of P3, P4-1 and P5 Primordia

To investigate whether rice leaf primordia of different developmental stages could be distinguished through metabolomics, I performed test metabolite extractions to ensure the extraction method chosen was appropriate for rice, and to optimise signal levels in the mass spectrometer by diluting the sample where necessary (data not shown).

Having concluded that sufficient signal was produced, I designed an experiment to look at the profiles of P3, P4-1 and P5 primordia in IR64 to complement the earlier work our lab had produced examining transcriptomic changes during rice leaf development (van Campen et al., 2016). Three biological replicates were tested for each stage, with the P3 samples consisting of 8-10 pooled leaf primordia, the P4-1 stage comprising 2-3 pooled leaf primordia, and the P5 stage consisting of approximately 1cm of leaf blade. Each individual sample was then run in triplicate for both the positive and negative phase. It is important to note for these experiments the P4-1 primordium stage was judged by eye, so there is likely some variation in precise developmental age (see Chapter 2).

The resulting data was then analysed using an unsupervised principal component analysis (PCA), separately for both the positive (Figure 3.2A) and negative phase (Figure 3.2B). For both phases, primordia of each of the three developmental ages were clearly clustered together, with this clustering being particularly strong for the P5 staged tissue.

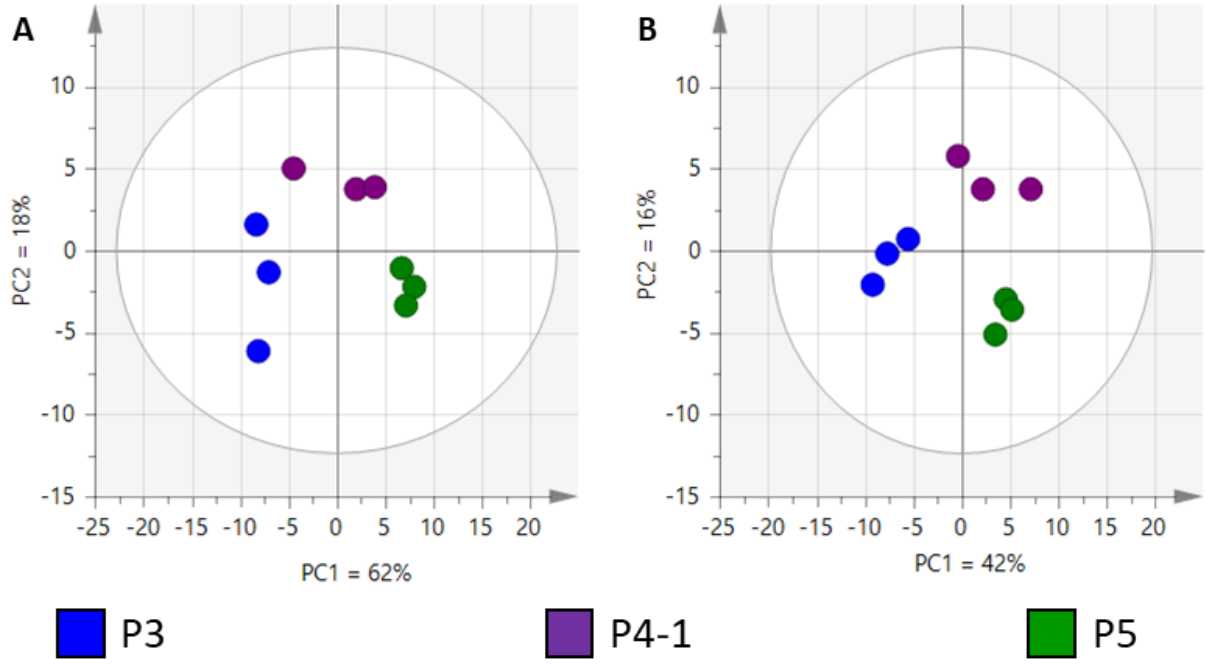


Figure 3.2 – Multivariate analyses of staged leaf primordia

Aqueous extracts from staged leaf primordia (P3 (8-10 pooled primordia) (blue), P4-1 (2-3 pooled primordia) (purple), and P5 (green)) were analysed in positive (A) and negative (B) mode. Masses were binned to 0.2Da. Data are shown as PCA score plots where the ellipse = 95% confidence (Hotelling's T2) and axes show loadings t[1] and t[2] and refer to principal components 1 and 2 for the model (for positive phase A, R2X[1] = 0.618, R2X[2] = 0.182; for negative phase B, R2X[1] = 0.417, R2X[2] = 0.164). Each point represents an average of 3 technical replicates of 1 biological sample. n=3 for each developmental stage.

As good separation of developmental stage was apparent in the PCA plots (both in positive and negative modes), I examined each stage in a pairwise fashion to see which masses were underpinning the majority of the separation. I was particularly interested in masses allowing the P3 primordia to be distinguished from P4-1 primordia, as this transition is the focus of this thesis. While it is not possible to confidently assign a metabolite to a given bin without performing further work, such as MS/MS, it is possible to make preliminary identifications, based both on the IDs assigned in available databases and some *a priori* knowledge of the samples analysed (leaf tissue). This analysis suggested that sucrose (monoisotopic mass 342.116) was playing a major role in the positive phase in distinguishing P3 and P4-1 primordia in the PCA. The 3 bins most significantly affecting R2X[1] (which describes the fraction of X variables which the model explains, and has a value of 0.616) are 381, 382 and 383 – likely potassium adducts of monoisotopic sucrose and sucrose with one or more C_{13}. This interpretation raised the issue that high levels of sucrose can suppress the detection of other metabolites which may be expressed at lower values (Annesley, 2003).

This issue was not apparent in the negative mode analysis, thus, I decided to perform subsequent experiments in negative mode only.

3.3.2 Untargeted Analysis of Staged and Measured Primordia

To get a better idea of how the metabolite profile of a leaf primordium shifts over developmental time, I performed a larger experiment with individual dissected primordia. This involved individually measuring and instantly freezing each primordium to accurately stage it using the Tip Development Index developed in Chapter 2. While this meant it was not possible to accurately weigh each sample (an individual primordium is too light to accurately weigh), each small primordium at the P4-1 stage and below was extracted as if it had a weight of 1mg – the lowest weight at which I could still accurately measure extraction buffers and obtain enough sample to run in triplicate, while the P5 primordia could still be weighed accurately. This approach allowed me to more accurately designate P4-1 primordia than in my previous experiment, with only those primordia measuring 1cm in length used. Furthermore, this dataset encompasses primordia fitting all stages I to VI as designated in Chapter 2, allowing the assessment of the potential to distinguish between these stages using metabolomic fingerprinting in addition to SEM.

While all samples were analysed in the same experiment on the mass spectrometer, I initially performed multivariate analysis using only the data for the P3, P4-1 and P5 primordia separately, as I now knew each stage should distinctly cluster. After both PCA (Figure 3.3A) and orthogonal partial least squares discriminant analysis (OPLS-DA) (Figure 3.3B) three distinct groups were visible, with the biological replicates for each group this time clustered closer together than shown in Figure 3.2– likely due to the more accurate designation of the primordia stage during sample dissection. When performing multivariate analysis using all samples together, samples intermediate between either P3 and P4-1, or P4-1 and P5, appear to have footprints intermediate between these stages (Figure 3.3C, D).

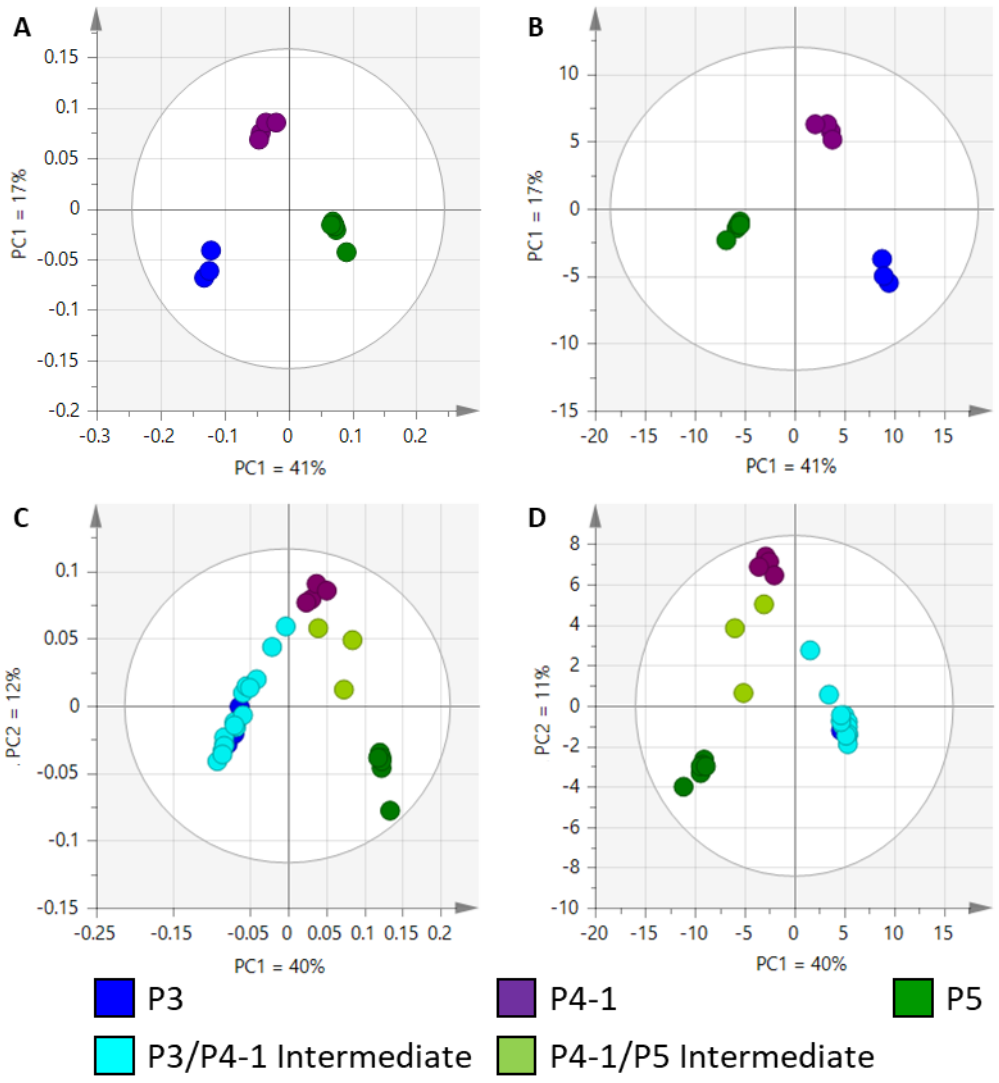


Figure 3.3 – Multivariate analyses of staged leaf primordia

Aqueous extracts from individual leaf primordia (staged P3, P4-1, P5, and intermediate samples of a known length) were analysed in negative mode only. Masses were binned to 0.2Da.

A – PCA Score plot for staged P3 (dark blue), P4-1 (purple) and P5 (dark green) primordia only, where ellipse: 95% confidence (Hotelling's T2) and axes show loadings t[1] and t[2] and refer to principal components 1 and 2 for the model (R2X[1] = 0.41, R2X[2] = 0.172). Each point represents an average of 3 technical replicates of 1 biological sample. n=3 for P3, n=4 for P4-1, n=7 for P5.

B – OPLS-DA Score plot for data shown in A. R2X[1] = 0.412, R2X[2] = 0.173.

C - PCA score plot for staged P3 (dark blue), P3/P4-1 intermediate (light blue), P4-1 (purple), P4-1/P5 intermediate (light green) and P5 (dark green) and intermediate primordia, where ellipse: 95% confidence (Hotelling's T2) and axes show loadings scaled t[1] and t[2], and refer to principal components 1 and 2 for the model (R2X[1] = 0.399, R2X[2] = 0.12). Each point represents an average of 3 technical replicates of 1 biological sample. For P3, n=3; P3/P4-1 intermediate, n=16; P4-1, n=4; P4-1/P5 intermediate, n=3; P5, n=7.

D – OPLS-DA Score plot for data shown in C. R2X[1] = 0.396, R2X[2] = 0.111.

3.3.3 Specific Metabolic Changes Over Early Leaf Development

3.3.3.1 Putative Identification of m/z Bins

The data obtained in the experiments described above gives an overview of the global metabolome changes in developing leaves. It is impossible to assign definitive identities to m/z values collected in these experiments. Discussion of putative identities below is based on combining spectral information with database searches and knowledge of the biological system in question: the developing leaf. With this in mind, I was still interested to speculate on which pathways may be altered over the course of development, in order to focus future experiments which can exploit this novel metabolome of the developing leaf, as well as the existing transcriptome produced by Julia van Campen in a previous PhD project.

The data presented in 3.3.2 was largely analysed through PCA, whereas to investigate further marker metabolites of interest I largely used OPLS-DA models. Both PCA and OPLS-DA analyses are used to examine complex, multivariate datasets, but provide different information. PCA models allow for the definition of class boundaries, and show which samples group together. OPLS-DA models use this class information, and ask which variables are causing this separation between the known groups.
For the initial multivariate analysis, it was important to examine the data in an unbiased fashion, by providing no class information in the model. The good clustering observed through PCA in the first incidence made it appropriate to use a more supervised form of analysis to delve into the specific variables contributing to class differences. That is to say, that PCA tells us the samples are different, and defines class borders. OPLS-DA allows us to look at why the samples differ by identifying the variables most strongly associated with each class.

Figure 3.4 below shows the score and loadings plots for both the pooled (A, B) and individual (C, D) data. In both datasets, the masses most strongly associated with the P3 cluster are m/z 195.8 and 197.8. For the pooled data, m/z 377 and 379 are strongly associated with the P4-1 cluster, while for the individual data, m/z 377.2 and 379.2 emerge. The P5 cluster is associated with m/z 191 in the pooled data and m/z 191.2 in the individual data. Pairwise comparisons also reveal masses of interest for maturing tissue are m/z 133 and m/z 115, which are associated with P4-1 in P3/P4-1 comparisons, and with P5 in P4-1/P5 comparisons.

The discrepancy of 0.2 Da between some bins of interest between the two experiments can be explained by minor changes in instrument calibration over time (section 3.2.4). As this analysis bins to 0.2Da, a small shift of 0.1Da in sensitivity over different experiments can

result in some ions being assigned to different bins. When examining specific ions, this can be considered by directly examining the spectra, taking into account the lock mass value, explained fully in Figure S3.1.

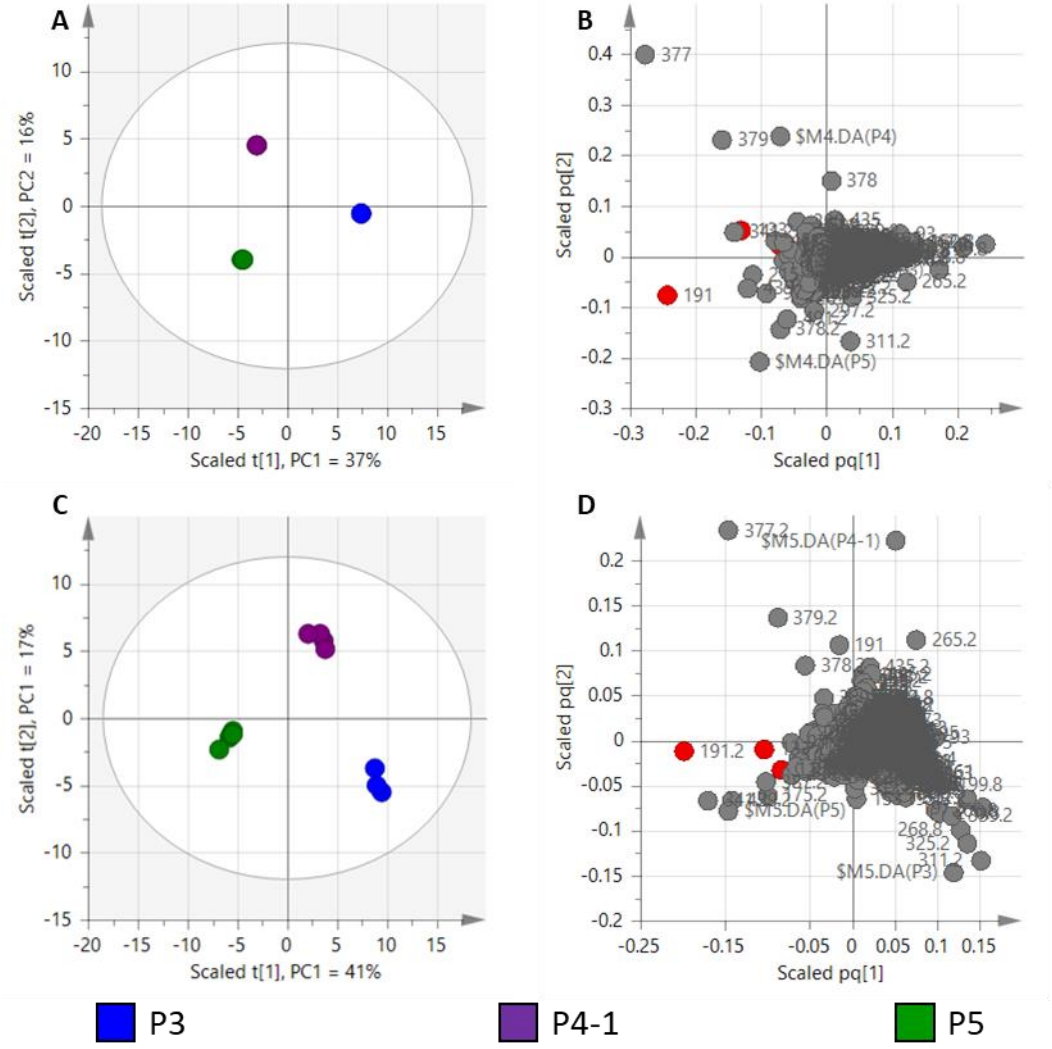


Figure 3.4 – Score and loadings plots for staged leaf primordia

A – OPLS score plot for pooled primordia data shown in 3.3.1

B – OPLS loadings plot for pooled primordia data shown in 3.3.1

C – OPLS score plot for individual primordia data shown in 3.3.2

D – OPLS loadings plot for individual primordia data show in 3.3.2

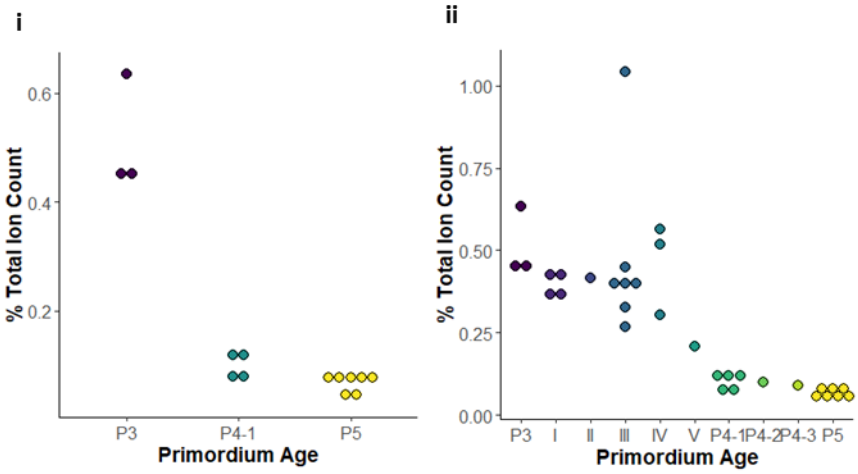
For score plots, ellipse: 95% confidence (Hotelling's T2), and points represent P3 (blue), P4-1 (purple) and P5 (green) samples. Axes represent components 1 and 2 of the OPLS model.

For loading plots, each point refers to an m/z bin and is labelled accordingly. Axes represent predictive and orthogonal components of the OPLS model. Compounds of interest are highlighted in red.

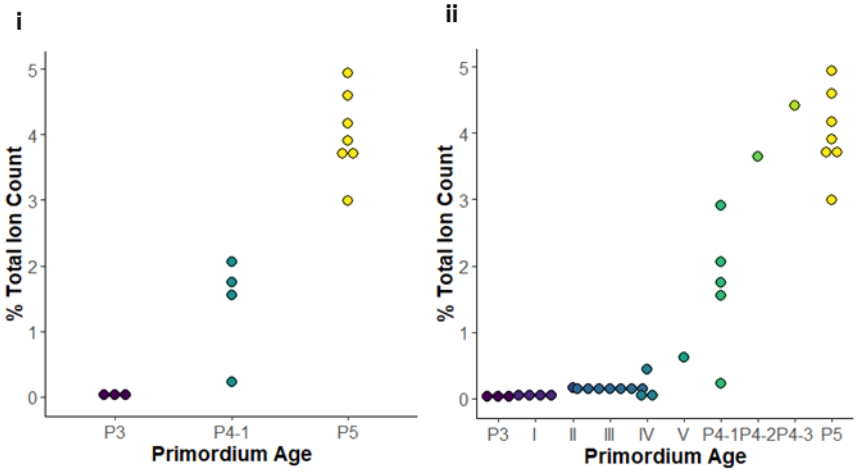
In both pooled and individual experiments, ions at m/z 191(.2), m/z 133 and m/z 115 appear as influential markers of the maturing leaf, highlighted in red on the loadings plots (Figure 3.4B, D). Through a database search using PubChem (Kim et al., 2019), these markers can be putatively identified as citrate (exact mass 192.027), malate (exact mass 134.022) and fumarate (exact mass 116.011). These three compounds have an obvious association as part of the tricarboxylic acid (TCA) cycle. To see if other m/z bins associated with TCA metabolites were also implicated in leaf maturation, I examined the data for m/z 145 (associated with α-ketoglutarate, exact mass 146.022), m/z 131 (associated with oxaloacetate, exact mass 132.006) and m/z 117 (associated with succinate, exact mass 118.027). While not a definitive identification, I was interested to see if there were interesting fluctuations over the course of development that would warrant further experimentation.

Bins m/z 191.2 (citrate), m/z 133 (malate) and m/z 115 (fumarate) all constitute a relatively high proportion of the total ion count at their highest level, and all show an increase in relative abundance between P4-1 and P5 (Figure 3.5B, C, D). In contrast, the relative abundance of m/z 145 (α-ketoglutarate) drops over the course of development from P3 to P5 (Figure 3.5A), and the relative abundance of m/z 131 (oxaloacetate) drops between P4-1 and P5, although in both cases these markers constitute a lower total percentage of ion counts (Figure 3.5E). The proportion of m/z 117 detected in the sample does not change significantly over the course of development (Figure 3.5F). It was also possible to visualise this trend over intermediate stages, and each graph in Figure 3.5 has a part ii, showing data for all primordia including those not falling into a P3, P4-1 or P5 classification.

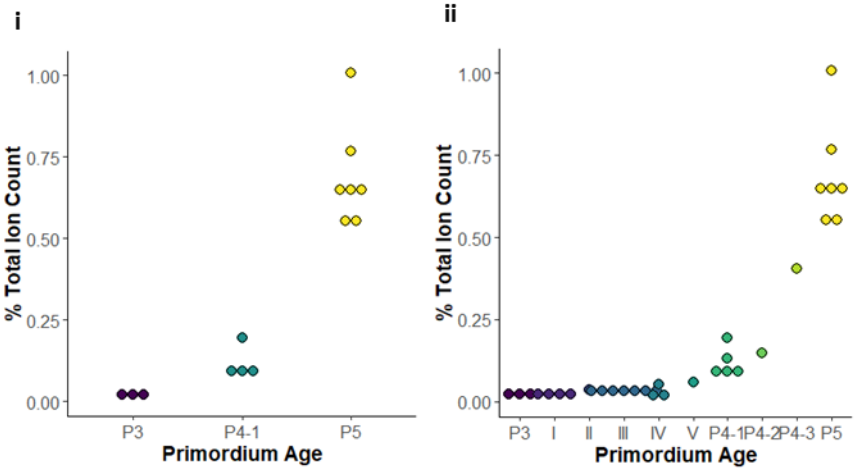
A – m/z 145 – putative α-ketoglutarate



B – m/z 191.2 – putative citrate



C – m/z 115 – putative fumarate



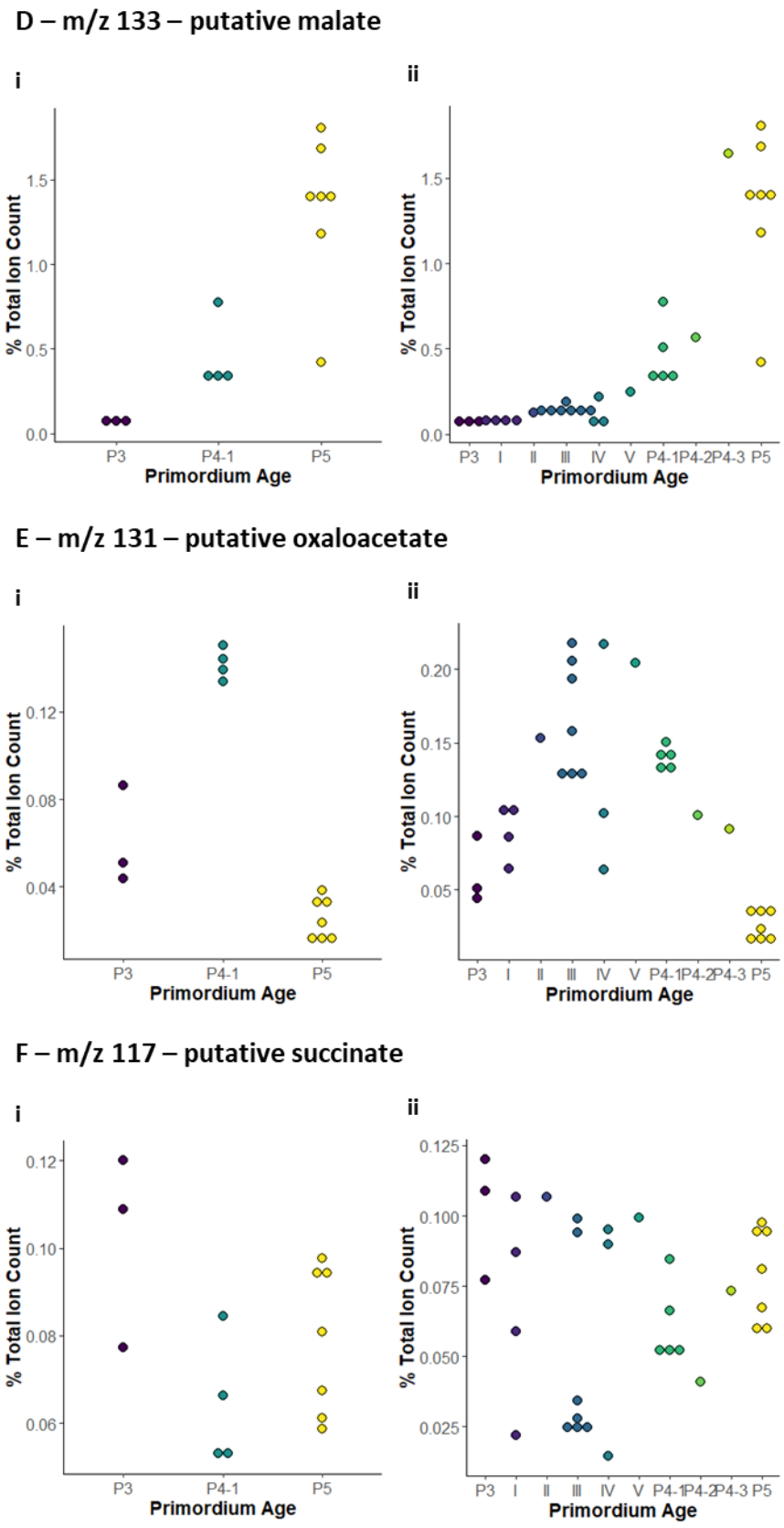


Figure 3.5 – Expression of putative TCA cycle metabolites

Data shown are from the individual primordium dataset described in 3.3.2. Each point represents an individual primordium. For each metabolite, data is shown for staged P3, P4-1 and P5 primordia only (i), as well as for all

primordia including developmental intermediates (ii). Different age groups are shown across the x axis and indicated by colour, with dark purple showing the youngest leaves and yellow the oldest.
A – Bin 145, associated with α-ketoglutarate
B – Bin 191.2, associated with citrate, isocitrate
C – Bin 115, associated with fumarate
D – Bin 133, associated with malate
E – Bin 131, associated with oxaloacetate
F – Bin 117, associated with succinate

The data shown in Figure 3.5 can also be presented in a diagrammatic form of the TCA cycle whereby the data are assigned to specific TCA metabolites for P3, P4-1 and P5 stages, with the relative level of metabolite being indicated by a grayscale (Figure 3.6A). This method of presentation can be extended to the intermediate stages between P3 and P4-1 (Figure 3.6B).

From this we can see that the m/z bin assigned to α-ketoglutarate remains fairly constant throughout much of the P3 to P4-1 transition (P3 to V), albeit at lower levels than in P3. In contrast, putative oxaloacetate levels appear to peak during the P3 to P4-1 transition at intermediate stage III. These data shows a potentially interesting flux in a core area of central metabolism over leaf development, and merits further work to confirm the identities of these ions.

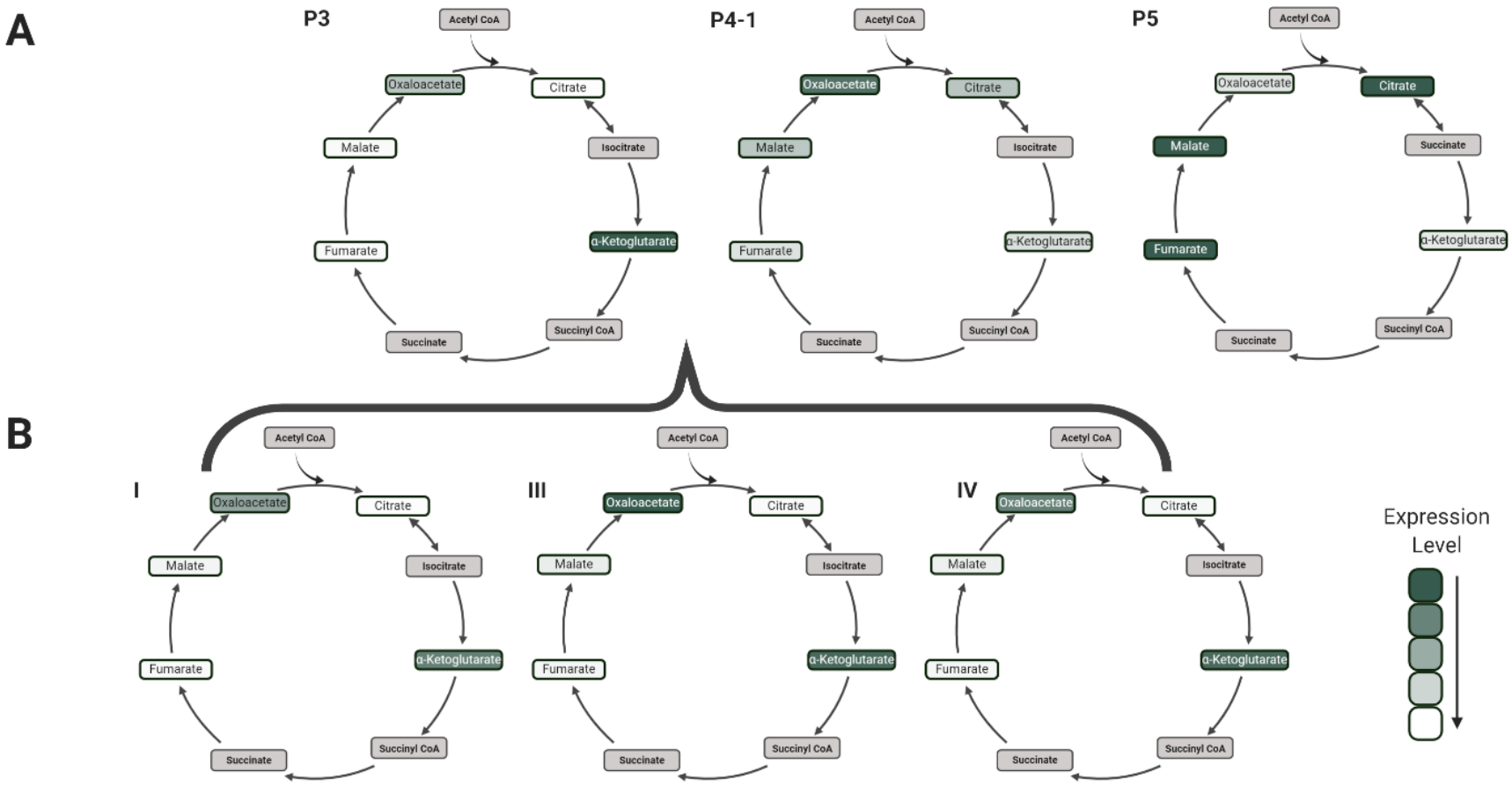


Figure 3.6 – Visualising putative TCA cycle metabolite flux

Each metabolite level has been normalised to its highest total % ion count value as 1, represented by dark green. Lower expression values are represented by a lighter shade. Grey boxes cover acetyl coA, succinyl coA, isocitrate and succinate, which we either do not detect, cannot be distinguished or do not change over the course of development. Cycles for staged P3, P4-1 and P5 primordia are shown in A, and for staged P3/P4-1 intermediates in B. Data for intermediate stages where n=1 has been omitted. For P3 n=3, P4-1 n=4, P5 n=7, I n=4, III n=7, IV n=3.

Figure made with biorender.com.

As we are particularly interested in when and where photosynthetic function is established, a key metabolic pathway of interest is the Calvin-Benson cycle. RNAseq data showed that Calvin-Benson cycle enzyme transcripts, including Rubisco, increased from P3 to P4-1 to P5 as the leaves matured (van Campen, thesis, 2016). However, this was not reflected in the accumulation of metabolites in m/z bins of interest (described below in Table 3.2). No bin associated with a Calvin-Benson metabolite was represented as containing a high total % ion count for any developmental stage in the dataset, and amongst those that were detected, there was no consistent pattern as the leaf matured. Taken overall, there was little metabolite data suggesting a convincing rise in Calvin-Benson related metabolites as the leaves progressed from P3 through to P5. Several possible explanations for this observation are discussed in 3.4.1, however the key conclusion is not that these metabolites are not present, but that they are not detected using this technique.

Table 3.2 – Exact masses of Calvin-Benson cycle metabolites

Data from PubChem (Kim et al., 2019).

Compound	Exact Mass (PubChem)	Expected Bin $[M-H]^-$
3-phosphoglycerate	185.993	185
bisphosphoglycerate	265.969	265
glyceraldehyde-3-phosphate	169.998	169
ribulose-5-phosphate	230.019	229
ribulose-1,5-bisphosphate	309.985	309

3.3.3.2 MS/MS of TCA metabolites

As indicated above, although it is possible to assign putative identities to some m/z bins, these identities need to be validated by other techniques. MS/MS, also known as tandem mass spectrometry, further fragments ions at an m/z of interest. Each compound will fragment in a unique and predictable manner, and this characteristic pattern can be compared to the fragmentation pattern of standard solution, either run alongside the samples or obtained from an online database. The initial aim of the following section was to perform a series of MS/MS experiments to validate the identity of some of the metabolites described in the previous section that showed significant changes during early rice leaf development. Unfortunately, due to the COVID-19 crisis it was not possible to carry out plans to repeat the analysis shown below, or to perform a larger scale experiment which would have tested more compounds and used MS/MS conditions tailored to individual metabolites. Here I

report on initial experiments that were performed to validate data, providing an oversight of the approach that would have been taken.

One m/z bin displaying a large change in level over the course of leaf development, and making up around 5% of the total ion count in mature P5 leaves, was m/z 191(.2) – putatively identified as citrate. Performing MS/MS on a citrate standard showed that the precursor ion 191.048 had an obvious fragment at 111.01, and two smaller fragments at 85.03 and 87.01 m/z, as highlighted in Figure 3.7B, reflecting the information available on the METLIN database (Figure 3.7A). When comparing to MS/MS of m/z 191 in P3, P4-1 and P5 leaf extracts (Figure 3.7C, D and E respectively), we can see that P5 leaves have both the 111.01 and 87.01 fragment, while the P4-1 shows only the 111.01, but at a lower level. It is not possible to visualise any fragmentation in the P3 sample, which is from a much smaller primordium. However, the sample fragmentation pattern reflecting both the citrate standard fragmentation and the information available online (Guijas et al., 2018) allows me to conclude that this 191.048 precursor ion can be reasonably identified as citrate in the P4-1 and P5 samples.

It is also possible that some fragmentation has been masked by the unusual high signal at m/z 160.84 and 162.84. This fragment also appears at the same unusually high level in the MS scans of the same samples. Further, these ions appear at around ten times higher levels in the P3 and P4-1 samples, which were not diluted prior to fragmentation, unlike the P5 sample which was diluted one in ten. In addition, these metabolites were not seen at these high expression levels in any previous MS samples. Taken together, this leads me to conclude that they represent a contaminant in the samples introduced during extraction.

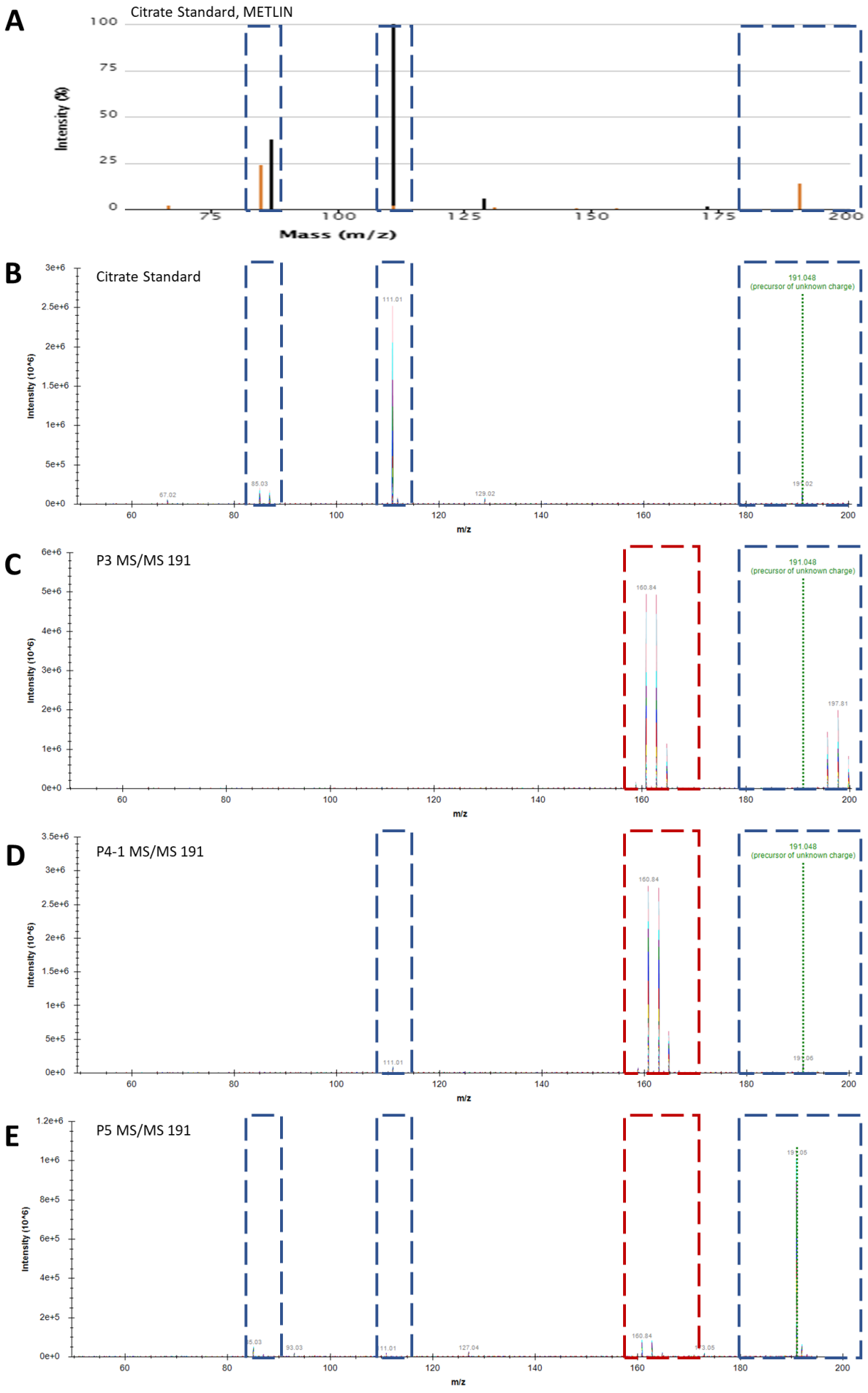

A
Citrate Standard, METLIN
Intensity (%)
Mass (m/z)
B
Citrate Standard
Intensity (10^6)
m/z
191.048
(precursor of unknown charge)
C
P3 MS/MS 191
D
P4-1 MS/MS 191
E
P5 MS/MS 191

Figure 3.7 – MS/MS of m/z 191

Each spectrum represents a sum of 60 scans. A – experimental data for citrate fragmentation at -10V, obtained from the METLIN database; B – citrate standard 1µgµL^{-1}; C – P3 primordium; D – P4-1 primordium; E – P5 primordium. Blue boxes represent precursor ions or fragments of interest, red boxes represent other or contaminating peaks of interest.

Another putative TCA metabolite identified from my metabolomics data is malate, which is found in m/z bin 133. From experimental data available on METLIN, we see that in negative ionisation mode the 133.01 precursor ion produces large fragments at 115 and 71.01 (Figure 3.8A). Whilst once again ion levels in the P3 stage (Figure 3.8B) are too low to deduce any meaningful fragmentation patterns, both the precursor ion and two main fragment ions can be seen in both the P4-1 and P5 samples (Figure 3.8C and D respectively). Although in both cases the spectra are noisier than for citrate, possibly due to the lower precursor ion levels, the fragment peaks are clearer. This information leads me to conclude that the precursor ion 133.031 can be reasonably identified as malate in the P4-1 and P5 samples.

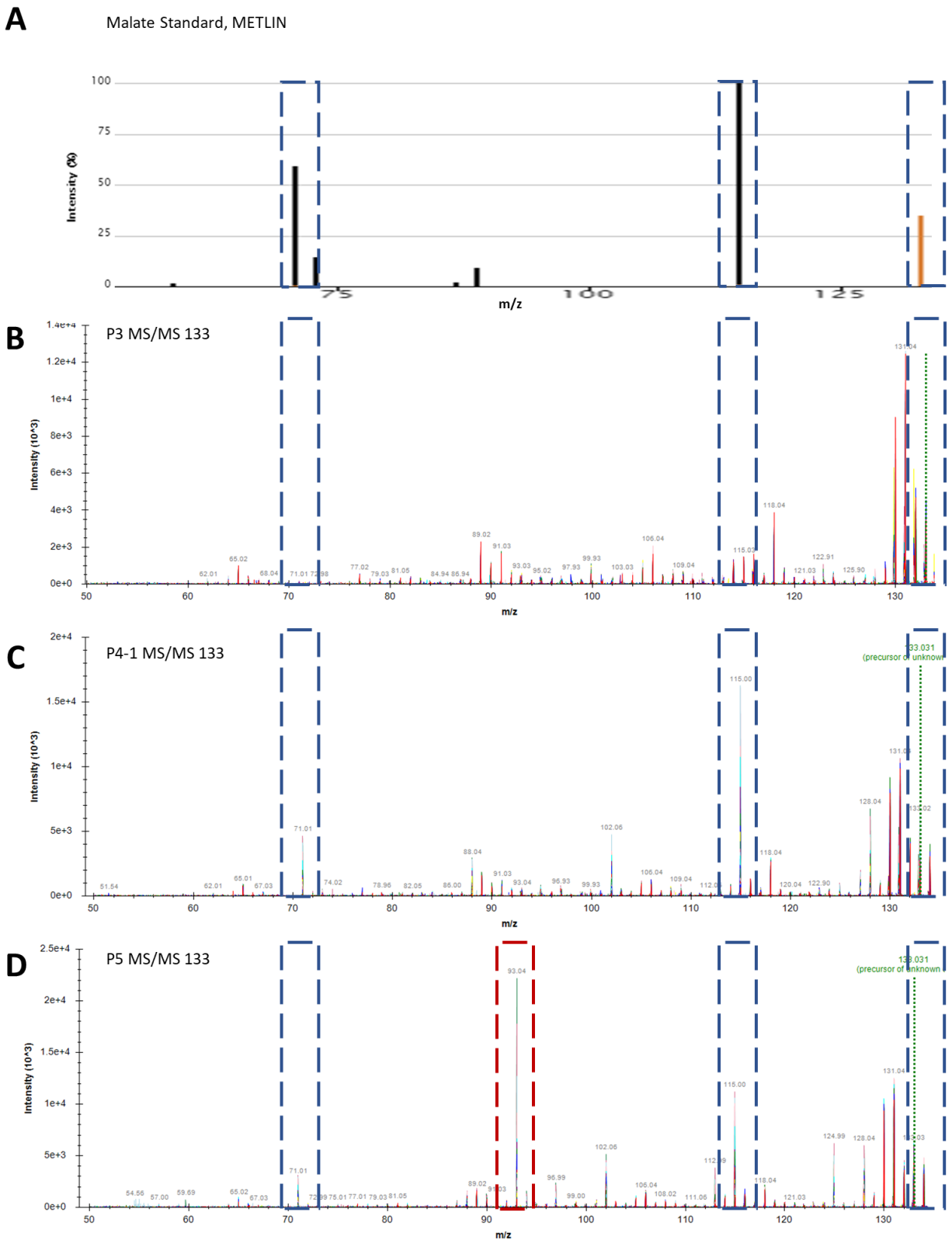


Figure 3.8 – MS/MS of m/z 133

Each spectrum represents a sum of 60 scans. A – experimental data for malate fragmentation at -10V, obtained from the METLIN database, B – P3 primordium, C – P4-1 primordium, D – P5 primordium. Blue boxes represent precursor ions or fragments of interest, red boxes represent other or contaminating peaks of interest.

In both cases this presumed identification is concentration dependent, as it is impossible to glean conclusive evidence from the small P3 samples.

3.3.4 Metabolomic and Transcriptomic Changes Over Leaf Development

3.3.4.1 Clustering of Metabolomics and RNAseq Data

While the analysis method described above used bins data to 0.2Da and so an individual bin may contain more than one metabolite, I was interested to look at the global shift in metabolite expression over the course of development. Mirroring the transcriptome clustering performed by Julia van Campen on similarly staged primordia (van Campen, 2016, thesis), all metabolite bins where differences in expression were detected were clustered into one of the 11 groups described in Figure 3.9, providing a general framework for the description of metabolic shifts during the transition P3 through P4-1 to P5.

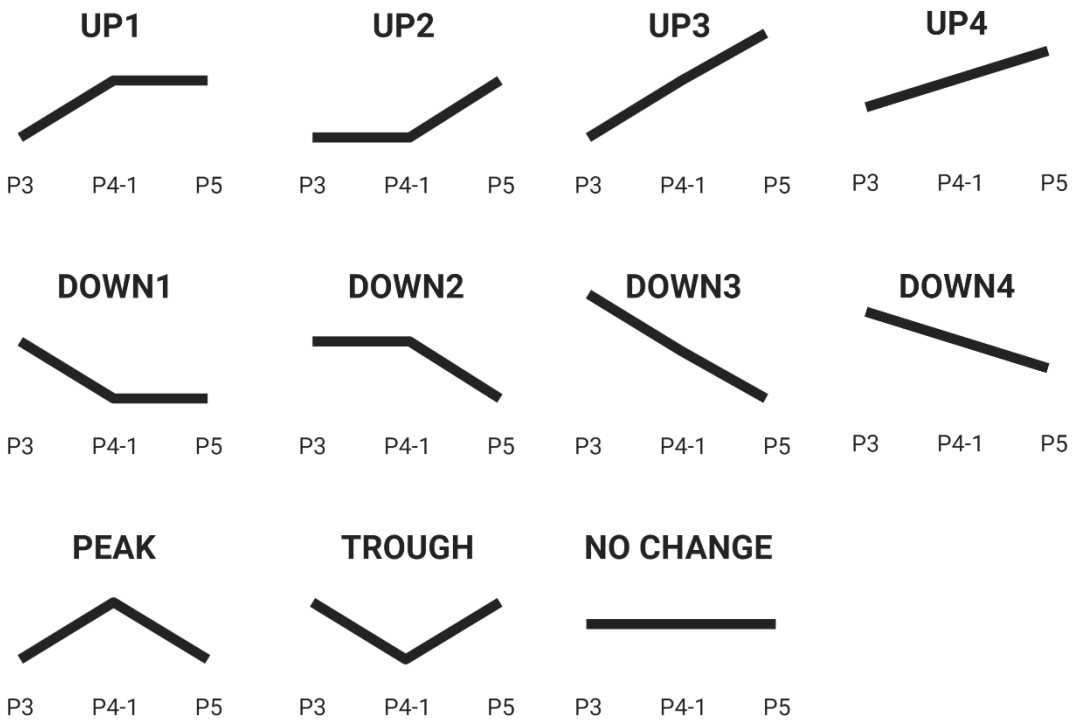


Figure 3.9 – Expression clusters for metabolomics analysis

For all comparisons in metabolite data, a significant difference is determined as a Benjamini-Hochberg corrected $p<0.05$, with a false discovery rate of 10%.

UP1 – significantly upregulated between P3 and P4-1, no change between P4-1 and P5

UP2 – no change between P3 and P4-1, significantly upregulated between P4-1 and P5

UP3 – significantly upregulated between P3 and P4-1, and between P4-1 and P5

UP4 – no change between P3 and P4-1 or P4-1 and P5, significant upregulation between P3 and P5

DOWN1 – significantly downregulated between P3 and P4-1, no change between P4-1 and P5

DOWN2 – no change between P3 and P4-1, significantly downregulated between P4-1 and P5

DOWN3 – significantly downregulated between P3 and P4-1, and between P4-1 and P5

DOWN4 – no change between P3 and P4-1 or P4-1 and P5, significant downregulation between P3 and P5

PEAK – significantly upregulated between P3 and P4-1, significantly downregulated between P4-1 and P5

TROUGH – significantly downregulated between P3 and P4-1, significantly upregulated between P4-1 and P5

NO CHANGE – no change between P3, P4-1 and P5

When examining the metabolomics data, only 31.8% of the metabolite bins fitted to one of the clusters describing some pattern of change over development, whereas over 56% of the published transcriptome data set did so. However, we cannot infer from this that a smaller proportion of metabolites than transcripts are changing over the course of leaf development for several reasons. Firstly, an individual bin may contain more than one metabolite. Secondly, metabolomic fingerprinting experiments are designed to measure as many ions as possible. However, these compromise conditions may still mean that some ions are missed.

Finally, mass profiles are frequently data-rich of small metabolites (below around m/z 350), as this range encompasses most primary metabolites. This may lead to changes in larger metabolites, which make up a much lower proportion of total signal, being missed, especially in small samples potentially due to ion suppression effects, where the overrepresentation of small metabolites masks detection of other, larger metabolites.

Of the transcripts and m/z bins which did change in expression over leaf development, the overall trend between the two datasets was fairly similar, with 48.4% of metabolite bins and 51.7% of transcripts upregulated, and 42.2% of metabolite bins and 40.8% of transcripts down-regulated.

However, of the data which showed changes in expression over the course of development, the proportion of each clustering pattern was very different between the two datasets. Comparatively more transcripts than metabolites change significantly (up or down) between P3 and P4-1, as shown by much higher levels of the UP1, UP3, DOWN1 and DOWN3 clusters in the transcriptomics dataset (Figure 3.10). Conversely, comparatively more metabolites than transcripts change significantly (up or down) between P4-1 and P5, shown by higher levels of the UP2 and DOWN2 clusters in the metabolomics dataset. There was also a higher proportion of metabolites than transcripts changing gradually (up or down) between P3 and P5, evidenced by higher representation in the UP4 and DOWN4 clusters.

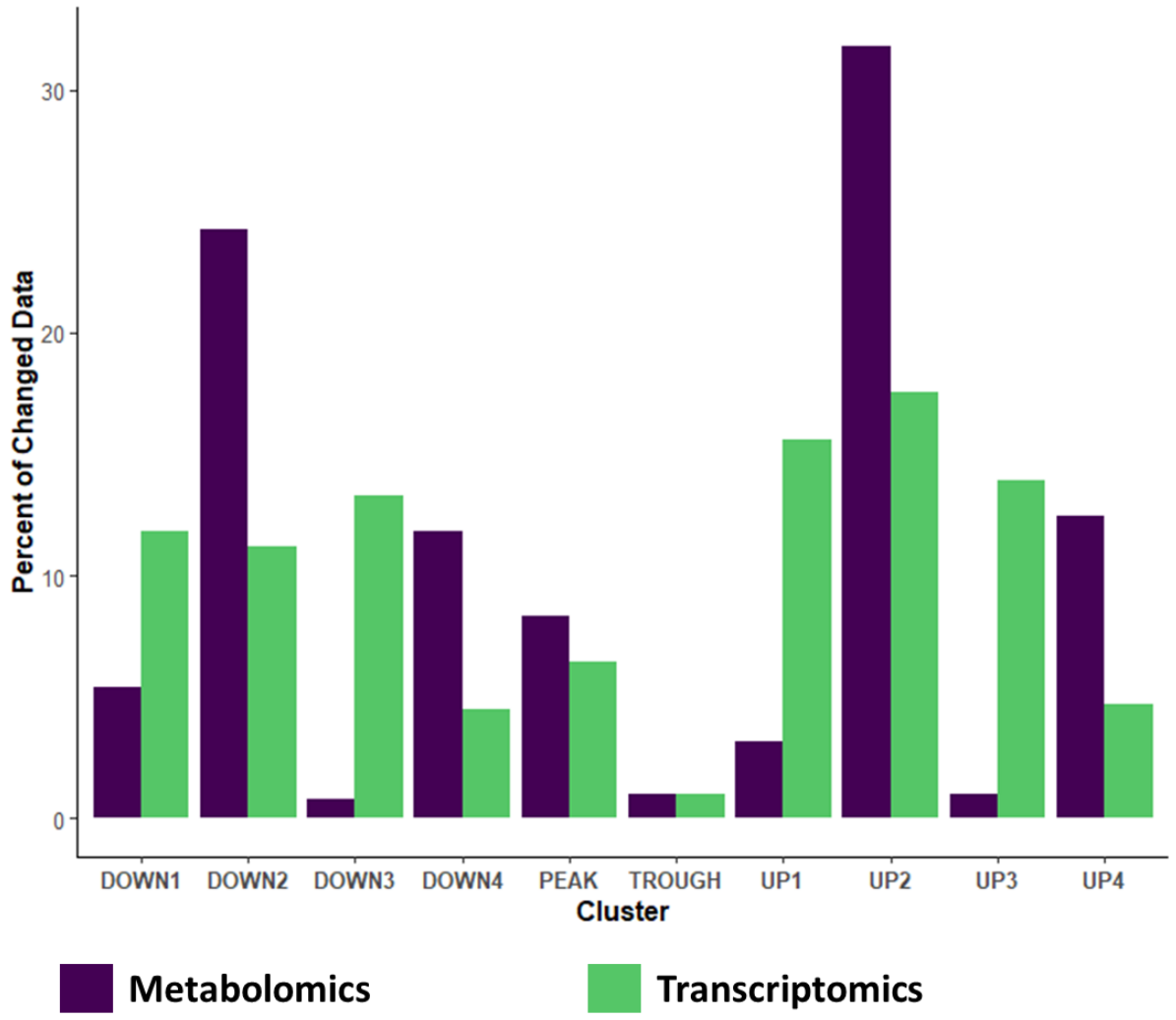


Figure 3.10 – Global metabolite changes over leaf development

Metabolomics data (purple) is from the individual primordia dataset shown in Figure 3.3. Transcriptomics cluster data (green) values from van Campen, thesis, 2016. In both cases, data shown is percentage of transcripts/metabolite bins where expression is changing over time – bins and transcripts with no change or no expression are omitted.

3.3.4.2 Specific Changes in Metabolomics and Transcriptomics Data

As discussed earlier, one of the purposes of this metabolomics fingerprinting dataset was to complement existing transcriptomics data. The metabolomics data presented in this chapter gives a snapshot of the state of the tissue in the moment of sampling. In contrast, transcriptomics data may reflect longer-term effects of signalling than metabolomics data, and is itself not necessarily indicative of corresponding protein changes, or indeed in changes to metabolite levels.

When considering fluxes through a given metabolic pathway, it is important to remember that they do not operate in isolation, rather that many pathways are interconnected. Pictured in Figure 3.11 is a diagram showing some of the other pathways connected to the TCA cycle. Notable related pathways include links to nucleotide biosynthesis through α-ketoglutarate and oxaloacetate, a route into porphyrin biosynthesis through succinyl coA, and the inhibition of glycolysis by citrate.

Thus, there are several possible biological explanations for changes in metabolite expression over time. For example, citrate level increases from P3 to P5, while α-ketoglutarate levels display the opposite pattern, decreasing from P3 to P5. One explanation for this could be the decreased expression of the enzymes catalysing this process (aconitase and isocitrate dehydrogenase), so that less citrate is converted to α-ketoglutarate. However, α-ketoglutarate also feeds directly into glutamine biosynthesis (and indirectly to the biosynthesis of other amino acids and of purines). Gene ontology analysis of the transcriptomics data (van Campen, thesis, 2016) shows that transcripts annotated as involved in transcription and DNA synthesis are enriched early in leaf development, so less α-ketoglutarate will need to feed into this pathway over time. The biological context of a given metabolite must be given careful consideration when designing future experiments, so that alternative pathways are not ignored.

Similarly, we cannot assume the biological outcome of a change in transcript levels without further experimentation. For example, citrate synthase transcript levels are shown to drop between P4-1 and P5, while the proportion of citrate in the samples as measured by metabolomics vastly increases during this transition.

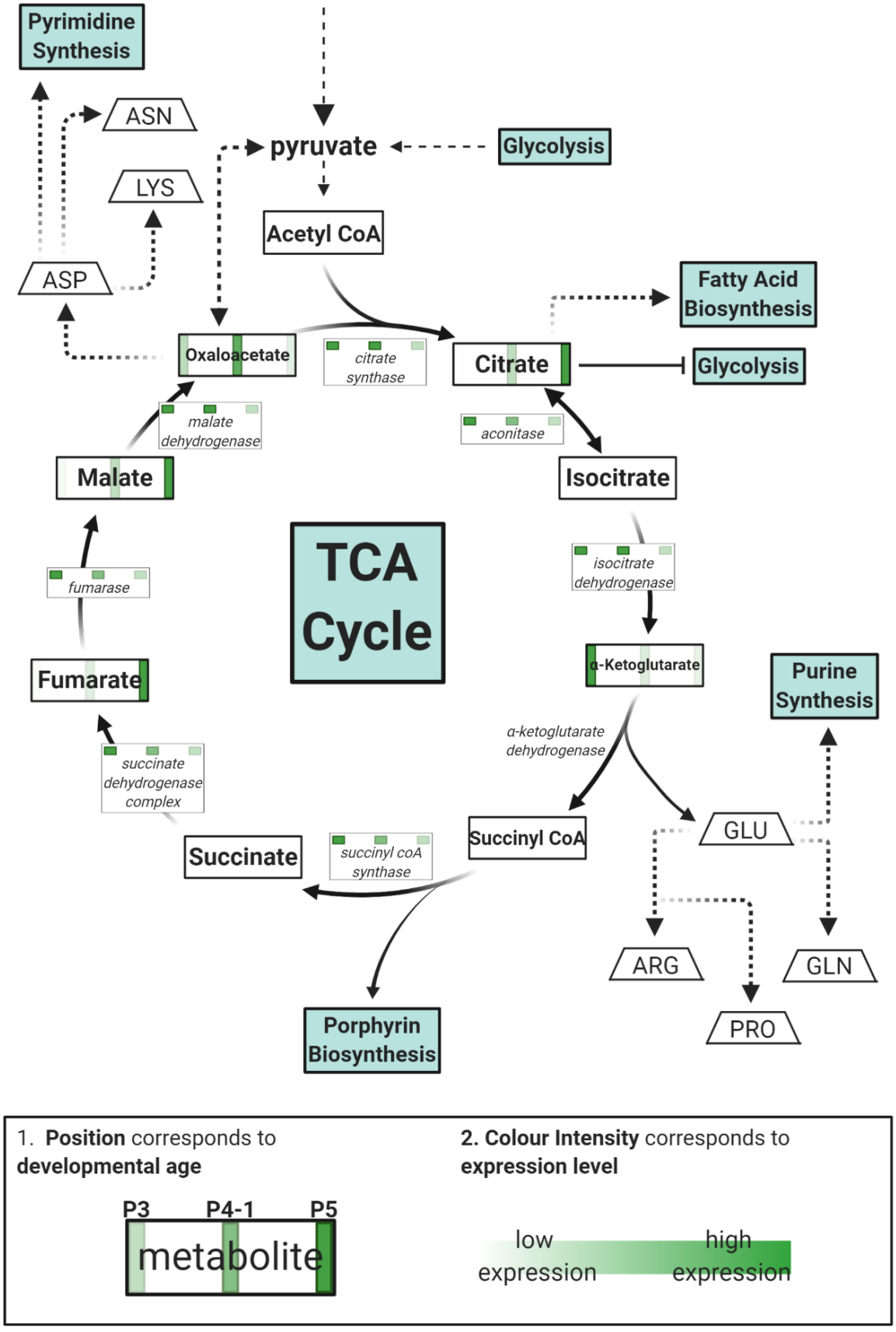


Figure 3.11 – Diagrammatic summary of TCA cycle metabolism

Simplified diagram to show some of the metabolic pathways that interact with the TCA cycle.

Blue rectangles represent pathways. Trapeziums represent amino acids. For the TCA cycle, a **bold font** shows metabolites, and *italics* shows transcripts. For both metabolites and transcripts, relative expression for P3 (left), P4-1 (centre) and P5 (right) is shown through colour intensity, with bright green showing the highest expression and pale green the lowest.
Metabolite data is based on putative identification described in Figure 3.3.4. Transcript data is from van Campen, thesis, 2016. Specific transcripts are described fully in Appendix 3.1. In both cases, expression level is normalised for each transcript or metabolite.

3.4 Discussion

3.4.1 Overview of Data Presented

In this chapter I present data setting out metabolomic fingerprints for the developing rice leaf. As discussed in Chapter 2, the majority of studies examining monocot leaf development take advantage of the basipetal developmental gradient of maturing leaves. In this thesis, I take the approach of examining whole leaf primordia which have developed *de novo* to different stages. This novel dataset complements the existing transcriptome for *de novo* developing leaves (van Campen et al., 2016), and provides further clarity to changes during the P3/P4-1 transition.

In contrast to other metabolomics studies of rice leaf development (Pick et al., 2011; Wang et al., 2014), the experiments presented in this chapter profile very small leaf primordia samples. These young primordia represent younger, less developed leaf tissue than reported in previous studies. Further, as entire primordia of different ages were sampled, in comparison to dissected sections of the same leaf, we can conclude that metabolites detected in these samples were either produced by the developing leaf or imported from older leaves, and not as a result of mobilisation from more mature cells in the same leaf. This perspective allows for an alternate view of leaf development: how is the leaf formed *de novo* from pluripotent cells?

As no published work exists to date profiling the metabolome of rice leaves as they develop *de novo*, I used an untargeted approach to maximise information gathered in this first instance. This was especially key due to the small sample sizes used in this work. The GC-MS approach adopted by Pick *et al* was unsuitable for these experiments due to prohibitively large minimum sample size requirements. Furthermore, using untargeted, unsupervised analyses (such as principal component analyses) to explore the data allowed for the discovery of patterns in the data without bias. PCA loadings plots were a useful tool to identify candidate marker m/z bins unique to a given developmental stage (Overy et al., 2005). Supervised analyses, such as OPLS models, were also helpful to confirm these candidate markers while inputting more information about the samples used. This method is

fast, sensitive, and has produced sizeable, global datasets using the very small amounts of sample. The experiments were also designed to complement existing transcriptomics data, to allow for better integration of available information about *de novo* rice leaf development. I hope that this unique resource will be useful for further investigations in this area.

3.4.2 Experimental Limitations

It is important to consider the conclusions drawn from this dataset in the context of the technical and experimental limitations. One such consideration is that the volumes of solvent used during sample extraction are dependent on the weight of the sample. When collected, samples were flash frozen in pre-weighed tubes, so the sample weight could be calculated after freeze drying. For P5 leaf blades, and the pooled primordia samples described in 3.3.1, accurate sample masses were obtained, and solvent quantities adjusted accordingly. However, many of the leaf primordia used in the individual extractions described in 3.3.2 had a negligible, undetectable weight, making this impossible. In these cases, samples were extracted as if they had a total weight of 1mg, as any lower levels of solvents would not have yielded sufficient eluate for mass spectrometry. For these samples, I attempted to collect similarly-sized samples at the time of dissection to minimise weight variation, but must acknowledge that concentration-dependent overrepresentation of ions may be a possibility in some samples, or that significant changes of ions expressed at very low levels may have been missed. The percent total ion count normalisation step described in 3.2.4 helps overcome this issue.

Extracting individual primordia also has several advantages over using pooled samples, however. Dissecting individual primordia is laborious, increasing the time taken to collect each sample. It would have been impossible to perform an experiment covering the same breadth of developmental stages using pooled samples, especially regarding covering developmental transitions. By using individual primordia, I could be more confident that each sample was not contaminated by older leaf tissue, as pooling would require only one primordium to have been transferred with older tissue to change the profile of the entire sample.

A lack of detection of a given compound with this untargeted approach does not reflect a lack of presence of said ion in the sample. For example, m/z bins in which we would expect Calvin-Benson metabolites to fall (described in 3.3.3) contain very small proportions of the total ion count and do not appear to change as the leaf matures. Common sense, however, dictates that photosynthetically active tissue such as P4-1 and P5 leaves do in fact accumulate Calvin-Benson metabolites. Detection will largely be determined by the volatility and ionisation energy required by the metabolites in the sample, and the untargeted

technique described here is used to observe the maximum range of compounds. Furthermore, phosphate is a very stable group which is easily lost from compounds during ionisation, meaning Calvin cycle metabolites may be observed in this screen, just not at their expected masses. Targeted metabolomics of the Calvin-Benson cycle requires comparatively more pre-processing, such as linking to LC-MS or isotope labelling (Arrivault et al., 2009) than this global approach. In summation, the lack of detection in this screen does not mean that further work with targeted techniques will not elucidate this information.

A further key limitation of the work presented here is the assignation of putative identities to ions in the dataset. In untargeted metabolomics experiments such as those described in this chapter, it is impossible to assign a definitive identity to any ion. It is possible to make educated guesses about likely candidates, through combining m/z information with knowledge about the biological properties of the sample. However, to definitively assign an identity to a given ion, further experiments such as tandem mass spectrometry are necessary. Thus, further discussion of fluxes through specific pathways in this chapter are a best approximation based on the information available, and would require future follow-up experiments.

As mentioned in 3.3.3.2, some follow-up experiments were planned to conclude this chapter, but were unfortunately curtailed by the COVID-19 crisis. While this early end to experimentation was disappointing, the preliminary experiments allowed me to work out key considerations for moving forward and further exploring this dataset. I do believe this is an interesting dataset, and would be a good starting point if this work were to continue in the future.

3.4.3 Potential Applications of Metabolomic Fingerprinting of Leaf Primordia

As shown for both pooled and individual leaf primordia samples, the metabolic fingerprints of P3, P4-1 and P5 primordia are easily distinguished using this mass spectrometry technique, and each developmental stage has unique potential markers associated with it, satisfying two of the core aims of this chapter. The similar clustering, and identification of the same marker metabolites for each developmental stage between the pooled and individual primordia samples is a promising indication that extracting from individual primordia provides sufficient and useful information to obtain a global profile of metabolism. This will allow for the possibility of larger scale studies, and requires fewer plants to be grown for sample collection, which is particularly important given the collection of leaf primordia is a destructive technique.

In the context of C_4 rice, this approach could be particularly useful to assess global metabolic changes in, for example, a wide variety of transgenic plants. Work comparing the metabolome of C_3 rice and C_4 maize has revealed very different metabolite profiles between the two (Wang et al., 2014), and there are known differences in Calvin-Benson cycle profiles between C_3 and C_4 plants (Schreier and Hibberd, 2019). A logical extension is that C_3 rice and C_4 rice should also exhibit differing metabolic profiles as a result of these different modes of photosynthesis. An initial global metabolite screen could identify transformants whose metabolism has been altered for further, targeted testing. This could be an advantageous approach to screen plants of interest compared to histology and genetic approaches. Screening using a histology approach can be very time consuming, especially as transgenic plants may not yield the desired phenotype ((Wang et al., 2017a), discussed further in Chapter 4), and genetic approaches that rely on the quantification of transcript levels do not necessarily mean there has been an alteration to protein levels, or indeed to central metabolism. Identification of C_4 photosynthetic activity by measuring gas exchange requires mature leaves, and is a time-consuming procedure. A direct, rapid analysis of early primordium metabolism may provide a more reliable indication of acquisition of C_4 traits.

3.4.4 Investigating Early Leaf Development

A third core aim of this chapter was to investigate whether P3/P4-1 intermediate stages, as assigned in Chapter 2, have significantly different global metabolic profiles. From Figure 3.3, we see that the majority of P3/P4-1 intermediates cluster around the P3 primordia. However, there are notable individuals which reside between the P3 and P4-1 clusters in the score plots. Examining this transition in detail (shown below in Figure 3.12A) shows that this central cluster comprises all the youngest transition primordia (grades I to IV). Indeed, this is reflected in pairwise comparisons between grades (only those with three or more biological replicates were compared) – there is a visible separation of P3 and grade I primordia (Figure 3.12B), and of grade IV and P4-1 (Figure 3.12E), but no separation between I and III, or between III and IV (Figure 3.12C, D). Interestingly, there is one grade IV primordium that does not cluster with the other young leaves, but instead can be seen nearer to the grade V and P4-1 primordia. This individual measured 4mm in length, while the other two in this grade measured 3mm. This stark separation suggests that the point in the P3/P4-1 transition where sizeable global metabolic remodelling occurs is around the grade IV/V point. Interestingly, this also reflects the greatest contrast in epidermal patterning changes (described in Chapter 2), as papillae begin to appear and epidermal features rapidly begin to mature at this stage. Taken together, these data suggest that there may be an important transition phase in early rice leaf development, where there is a rapid switch in cell identity, rather than only gradual differentiation. This observation is in contrast to the conclusion of

other metabolomic studies of monocot leaf development (Pick et al., 2011; Wang et al., 2014), which surmise that differentiation occurs as a result of morphogenic gradients only. However, the youngest tissues used in these studies already had much of their anatomy determined, and were already green at the base, so represent tissue much more mature than the young primordia described in this thesis. Conversely, work in *Arabidopsis thaliana* has shown that exit from cell proliferation through a cell cycle arrest front – where there is a spatial transition from cell proliferation to cell elongation – is rapid, and dependent on the differentiation of photosynthetic machinery (Andriankaja et al., 2012). While the data presented here do not suggest the existence of a cell cycle arrest front in rice, they do suggest that morphogenic gradients may not be acting alone in controlling cell differentiation.

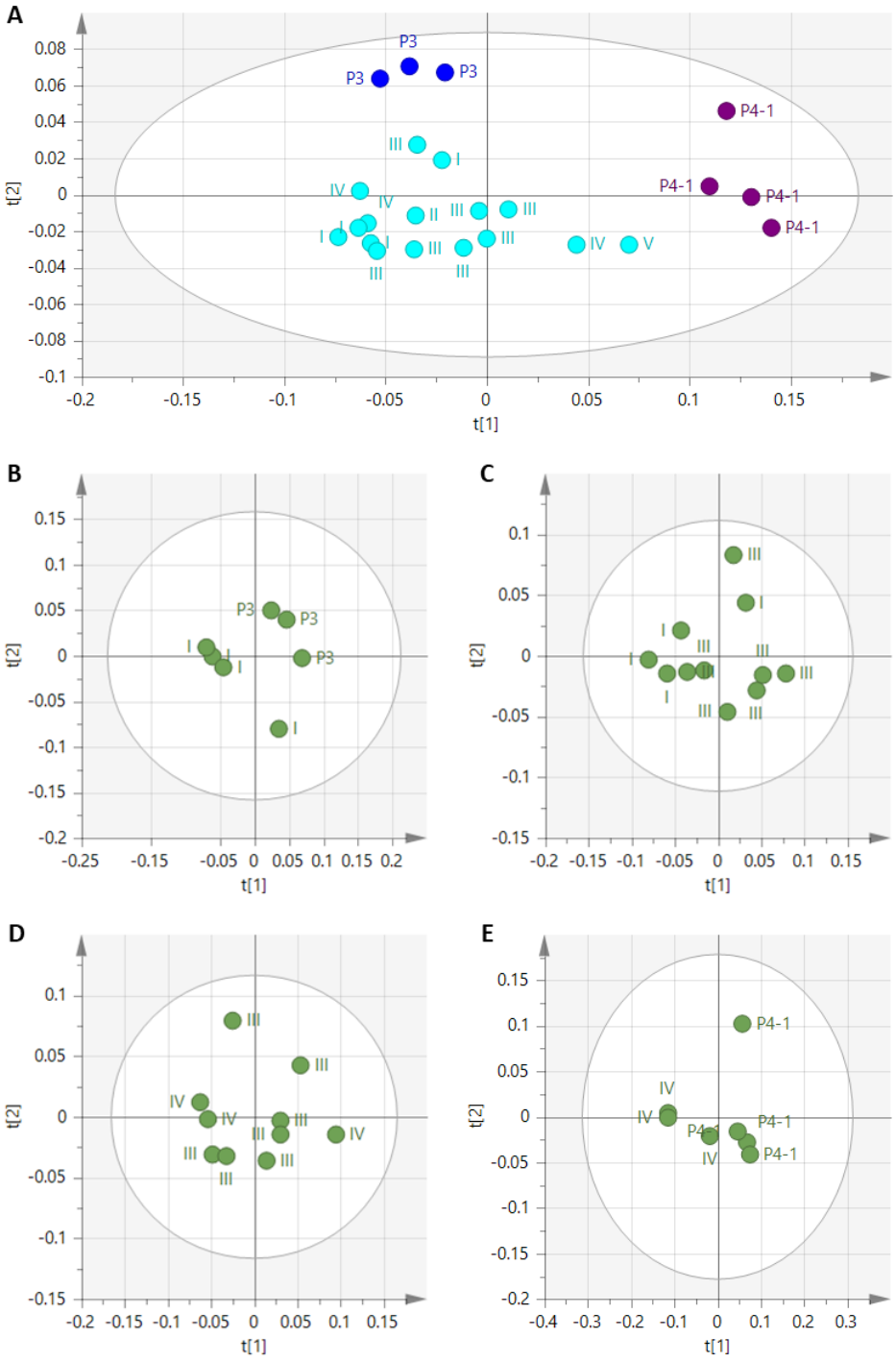


Figure 3.12 – Unsupervised analysis of the P3/P4-1 transition

All figures shown are PCA score plots. The ellipse represents the 95% confidence interval (Hotelling's T2), and the axes components 1 and 2 of the model. All points are labelled with their developmental grade, as assigned based on Chapter 2.

A – P3 (blue), P3/P4-1 intermediate (turquoise) and P4-1 (purple) individual primordia.

B – Pairwise comparison of P3 and I

C – Pairwise comparison of I and III

D – Pairwise comparison of III and IV

E – Pairwise comparison of IV and P4-1

One such example of this potential metabolic remodel can be seen when examining putative TCA data. Both (putative) citrate and malate have a significantly higher % total ion count between grade IV and P4-1, while remaining at the same level between P3 and grade IV inclusive. Further, the metabolite level of both citrate and malate in the 4mm grade IV primordium is between four and ten times higher than in the two 3mm grade IV primordia. Similarly, putative α-ketoglutarate, which drops in time between P3 and P5, is expressed at 90% of its P3 level at grade IV, and 19% at P4-1.

Very limited definitive metabolite identification was possible in the available time. While citrate and malate were convincingly (although not definitively) identified in the P4-1 and P5 samples, this would require repeating with cleaner samples. For both metabolites, strong concentration-limiting effects were observed, with ion counts at the P3 stage too low to deduce meaningful fragmentation patterns. Both ions were fairly abundant in the dataset, comprising up to 5% total ion count for citrate at the P5 stage. As a result of this, pooling and concentration of multiple leaf primordia would be necessary to detect lower abundance ions, or indeed to reliably detect any ions in the smallest primordia. By necessity, this experiment would have to be planned carefully to maximise information to be gleaned from each extract, as sample collection is a significant time investment in these experiments. Coupling MS/MS with HPLC could also ensure a better separation of parent ions, and a more effective use of limited sample. It would be particularly interesting to confirm the identities of other TCA cycle metabolites given the interesting fluxes through the cycle of their putative ions. As plants are known to utilise both cyclic and non-cyclic flux modes for a variety of biochemical processes (reviewed in Sweetlove et al., 2010), I would be particularly interested to know how this changes over the course of leaf development.

In addition to TCA cycle metabolites, it would be interesting to attempt to identify the marker metabolites for the P3 and P4-1 stages; m/z 195.8 and m/z 197.8 for P3, and m/z 377 and 379 for P4-1. In both cases, it is possible that both m/z bins represent the same species, either that is multiply charged (although ESI-TOF MS tends to create singly charged small ions) or though addition of one or more heavier isotope (although we do not see the expression ratio we would expect if this were the case). A useful starting point for this could be to perform MS/MS on the precursor ions in question, and compare the fragmentation patterns to information in a database such as METLIN.

3.4.3 Future Directions

The data presented in this chapter represent, to my knowledge, the first global look at the metabolome of the *de novo* developing rice leaf. Here, I will discuss potential future work building off of this data.

Firstly, it would be interesting to identify some of the key marker metabolites changing throughout development. I believe that TCA cycle organic acids would be a good starting point for this for several reasons. Within my dataset, several putative TCA metabolites have been identified at high % total ion counts, suggesting there may be suitably high signals for further work with minimal pooling of samples. Further, MS/MS scans of putative citrate and malate suggest that this is the correct identification for these compounds. In addition, an obvious next step is to use MS/MS to identify the marker metabolites for P3 and P4-1 primordia (section 3.3.3.1), and potentially investigate further any pathways they are involved in. A targeted screen to examine the expression of Calvin-Benson metabolites over time would also be useful, especially in the context of photosynthetic improvement.

It would also be interesting to utilise mass spectrometry to gain spatial information about metabolite distribution. While I am confident that my data reflect only the metabolites present in developing primordia, we know that the onset of chlorophyll fluorescence (van Campen et al., 2016) and epidermal differentiation (Chapter 2) occur rapidly from tip to base in the P3/P4-1 transition, and the mass spectrometry data presented in this chapter uses samples from entire primordia. For larger primordia, it may be possible to bisect during dissection, allowing for the examination of “tip” and “base” samples. However, it does appear there is already a big metabolic shift when the primordia measure just 3-4mm long, and are thus too small to dissect further by hand. In these cases, it may be possible to apply MS imaging techniques to visualise spatial distribution of the metabolites in sections of entire primordia, although this would be an intensive protocol, and likely require much optimisation.

This work also has the potential to be of use in the context of genetic improvement and the C_4 rice project. Introducing C_4 photosynthesis into rice requires the introduction of substantial changes to the metabolism and anatomy of the leaf. To assess the efficacy of this approach, a thorough baseline of “normal” metabolism will be useful. Existing comparisons between rice and maize have focussed on already mature leaves (Wang et al., 2014) but it is not yet clear if these patterns are reflected in *de novo* development. Observation of transgenic lines overexpressing potential C_4 genes has shown that in some cases, leaves appear to shift back to a C_3 state (Paul Quick, personal communication). It is possible that by examining *de novo* leaf development we can pinpoint when this is happening, and attempt to overcome it.

Chapter 4 – Exploring the Relationship Between Rice Leaf Anatomy and Photosynthesis

4.0 Collaborations

The work presented in this chapter was performed as part of a larger, ongoing project at Sheffield co-ordinated by Dr Jen Sloan (UoS). All experimental work was performed at the International Rice Research Institute (IRRI) in the Philippines. Data for the initial analysis was provided by Dr Xiaojia Yin (IRRI). Experiments were planned jointly by myself and Jen Sloan. All histology work was collected and analysed by myself. Gas exchange data was collected by Irma Canicosa (IRRI), fitted by Jen Sloan, and further analysed by myself and Jen Sloan.

4.1 Introduction

4.1.1 The Relationship Between Leaf Anatomy and Photosynthesis

As discussed in Chapter 1, one of the major limiting steps to photosynthesis is the inefficiency of Rubisco in rice and other C_3 plants. Over time, plants have evolved to counteract this, in part, by optimising several anatomical parameters to maximise internal CO_2 concentration in the stroma of the chloroplasts. For the CO_2 to reach the chloroplasts, however, it must pass through a series of boundaries: from the air, through the stomata and into the intercellular airspace (stomatal conductance, or g_s, reviewed in Farquhar and Sharkey, 1982); then through the cell wall, plasma membrane, cytoplasm, chloroplast membranes and finally the chloroplast stroma (collectively termed mesophyll conductance, or g_m, reviewed in Terashima et al., 2011) before it can reach the site of carbon fixation, as shown in Figure 4.1. In species where photosynthetic activity occurs in the bundle sheath cells, there is a further diffusion barrier between the mesophyll and bundle sheath (Brown and Byrd, 1993).

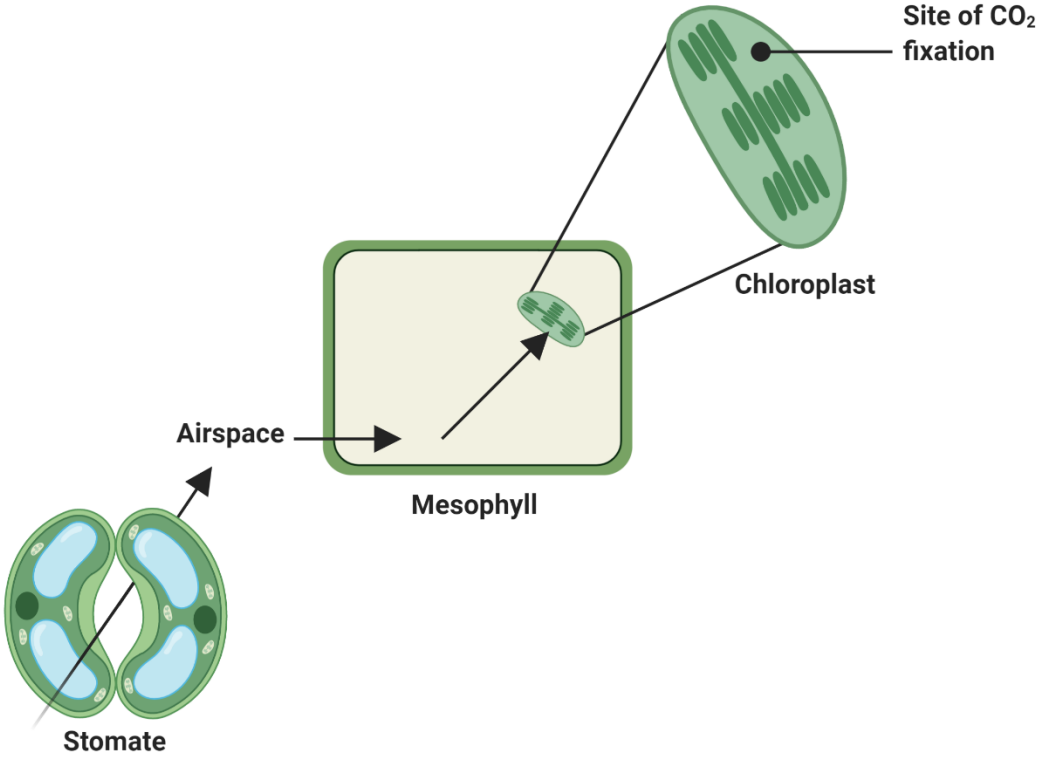


Figure 4.1 – Simplified diagram showing the CO_2 diffusion pathway in the leaf.

The path of CO_2 through the leaf is shown here. CO_2 passes through stomata to the airspace. It then enters the mesophyll cell, passing through the cell wall and cytoplasm. It can then pass into the chloroplast where it is fixed in the stroma.

Figure created with biorender.com

While the plant can transiently alter its g_s by changing the stomatal aperture in response to environmental stimuli (including light and external CO_2 (Farquhar and Sharkey, 1982)), g_m is less flexible, as it is set largely by leaf anatomy (Terashima et al., 2011). There is some evidence that g_m can partly respond flexibly to the environment (Flexas et al., 2008), and is also influenced by both aquaporins and carbonic anhydrases (Buckley and Warren, 2014). Within a functional group of plants, it has been shown that internal CO_2 concentration is inversely proportional to mesophyll cell wall thickness, and proportional to surface area of chloroplasts which are facing the intercellular airspace (Terashima et al., 2006, 2011). Currently, costly techniques such as scanning or transmission electron microscopy (SEM, TEM) or sequential light sheet microscopy are required to visualise plastid distribution. In rice, three dimensional reconstructions of mesophyll cells have shown that chloroplasts line the lobes of the cells in sheets (Oi et al., 2017; Ouk et al., 2019). This suggests that examining the shape and degree of lobing in the rice mesophyll – for example through hand sections and light microscopy – could be an acceptable proxy for direct measurement of plastid distribution, and a less time-consuming and more accessible technique. Overall, g_m is

under-investigated compared to g_s, even though it can be a sizeable limitation on carbon assimilation (Flexas et al., 2012).

4.1.2 Modelling Photosynthesis

During photosynthesis, CO_2 and H_2O are exchanged, and this exchange can be indirectly measured by an infra-red gas analysis system (IRGA) (Long and Bernacchi, 2003), as shown in Figure 4.2. The concentration of carbon dioxide and water vapour is measured (through their absorbance of infra-red light) in both an empty reference chamber, and an airtight sample chamber containing the clamped leaf. When combined with flow rate, differences in the levels of these gases can be used to calculate CO_2 assimilation rate (A) per unit of leaf area.

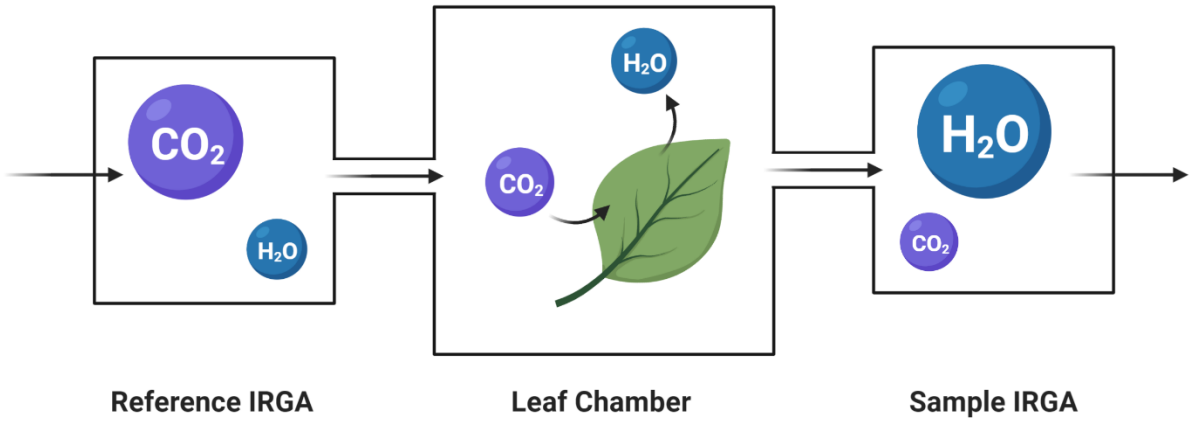


Figure 4.2 – Simplified schematic of an infra-red gas analysis system (IRGA)
Concentrations of CO_2 and water vapour are detected through their absorbance of infra-red light. Combined with flow rate, information can be compared between an empty reference chamber and sample chamber containing a leaf to give an indirect measurement of gas exchange.
Based on Wilson, 2019, thesis. Figure made with biorender.com

Several other parameters can be calculated from the IRGA measurements (Long and Bernacchi, 2003) or derived from tools based on the Farquhar-von Caemmerer-Berry (FvCB) model of photosynthesis (Farquhar et al., 1980). Some commonly used parameters are summarised in Table 4.1.

Table 4.1 – Commonly used photosynthetic parameters

Abbreviation	Parameter	Measured By	Type of Measurement
A	Leaf CO_2 uptake	Direct measurements of CO_2 flow in sample chamber relative to reference chamber	Direct measurement
C_i	Intercellular CO_2 concentration	Calculated from direct measurements of *A* and the CO_2 concentration around the leaf (C_a)	Calculated from direct measurements
$V_{c,max}$	Maximum rate of Rubisco activity	Calculated based on A/C_i response and other parameters	Calculated via. curve fitting
J_{max}	RuBP regeneration driven by electron transport	Calculated based on A/C_i response and other parameters	Calculated via. curve fitting
TPU	Triose phosphate utilisation	Calculated based on A/C_i response and other parameters	Calculated via. curve fitting

Fitting photosynthetic data using tools based on the FvCB model requires skilled user input as the output is modelled, at minimum, on two non-linear equations. Many such tools (for example, Duursma, 2015; Gu et al., 2010; Patrick et al., 2009; Sharkey et al., 2007) require either a minimum number of measurements, a number of assumptions and pre-fixed photosynthetic parameters, significant user input or a combination of these. Further, many fitting tools have been optimised around commonly used model laboratory plants such as *Arabidopsis thaliana* or *Nicotiana benthamiana*, which are often grown in 25°C – well below the temperatures at which rice is grown. As many parameters are calculated based on data from A/C_i curves, which themselves have been fitted subjectively based on the tool used, it is important to have the best data and best fit possible to accurately predict other photosynthetic parameters.

4.1.3 Testing the Link Between Rice Leaf Anatomy and Photosynthesis

An ongoing project in the Fleming lab is examining this potential link between leaf anatomy and photosynthetic performance in rice to contribute to a new 3D model of photosynthesis (Yi *et al*., in preparation). Of particular interest is how lines with different leaf anatomy may perform in increasing CO_2 concentrations. To aid this work, the group were interested in identifying lines with variations in individual structural parameters (but that were otherwise phenotypically normal) to help quantify these parameters' contributions to photosynthesis.

In 2017, 60 maize genes were transformed into *O. sativa* var Kitaake or IR64 under constitutive ubiquitin promoters as part of the international C4 Rice Project (Wang et al., 2017a). These 60 genes were a subset of the 283 potential regulators of Kranz anatomy identified by the work of Wang et al., 2013, and were chosen based on their predicted involvement in gene regulation (for example, transcription factors). Many of these transgenic rice lines were phenotypically normal, with very few exhibiting anatomical variation from the wild type. However, further analysis of the supplementary data by colleagues in Sheffield suggested that there was some variation amongst individual anatomical parameters such as bundle sheath cell size and mesophyll cell size, making these lines potentially useful tools in the context of the modelling project. In Sheffield, colleagues have tested several of these "JL" lines in a Kitaake background. Some transgenic lines appeared to have changes in a single structural parameter (for example, mesophyll cell lobing) in addition to altered rates of photosynthesis, while exhibiting no other gross phenotypic differences to the wild type. It would be interesting to explore if there were similar transgenic individuals amongst the IR64 background. These stocks are maintained at IRRI, so the opportunity to visit IRRI (the CASE partner in this PhD studentship) was exploited to investigate this unique resource. Shown below in Table 4.2 is a summary of the four lines of initial interest: JL75, JL80, JL81 and JL82.

Table 4.2 – Summary of JL genotypes of interest
Shown for each line is the maize gene (and its family) transformed into rice under a ubiquitin promoter, and the growth phenotype described in Wang et al., 2017a. All four lines are transformed in an IR64 background.

Line	Maize gene	Gene family	Growth phenotype
JL75	GRMZM5G850129	Growth Regulating Factor (GRF)	Normal
JL80	GRMZM2G126018	SQUAMOSA promoter binding protein-like (SBP)	Normal
JL81	GRMZM2G312419	R2R3 MYB	Normal
JL82	GRMZM2G478876	Serine Threonine Kinase	Normal

4.1.4 Chapter Aims

1. To screen existing data on the "JL" IR64 lines at IRRI, and select transformation events suitable for further study
2. To collect in-depth histology data on selected lines
3. To examine this new histology data alongside novel gas exchange data, and identify:
    a. Any line that has changes to both a given anatomical parameter and photosynthesis
    b. Which structural parameter changes reliably correlate with photosynthetic changes by studying all individuals as a population

4.2 Methods and Materials

4.2.1 Plant Growth and Materials

O. sativa var IR64 and the "JL Lines" (Wang et al., 2017a) were grown from seed. T2 "JL" seeds were germinated on hygromycin to select for transformants. Seeds were germinated in the lab, and transferred to the transgenic screenhouse after 4 days where they were grown in soil in root trainers for two weeks, then transferred to buckets for further growth. Plants in the screenhouse were grown under natural light conditions and in average outdoor day/night temperatures of 30ºC/25ºC ± 3ºC.

4.2.2 Histology

4.2.2.1 Sample Collection

Samples were collected within the first two hours of daylight to minimise starch accumulation in the leaves. As leaves could not be sectioned immediately, they were kept in water in darkness and sectioned the same day. Hand sectioning was performed with fresh razor blades, and the leaves cleared in lactic acid solution (~85% (v/v) lactic acid) saturated with chloral hydrate (~13g in 30mL) until clear, then transferred to 70% (v/v) ethanol for storage.

4.2.2.2 Imaging

Samples were mounted in water. Images were taken using an Olympus BX51 microscope, controlled by Olympus software. Images were analysed in ImageJ (version 1.52a) (Schindelin et al., 2015) using an in-house macro, and subsequently the data was analysed using R ((R Core Team, 2019; RStudio Team, 2016; Wickham et al, 2020; Wickham, 2016; Wickham and Henry, 2020; Garnier, 2018), (Versions: R – 3.6.2; RStudio – 1.1.452; dplyr – 0.8.4; ggplot2 – 3.2.1; tidyr – 1.0.2; viridis – 0.5.1)).

4.2.2.3 Image Analysis

For all image analysis, data from the region of three minor vein pairs were measured and analysed, and the mean of each parameter (as defined below) calculated for each biological sample.

Initial image analysis was performed on leaf four using simplified parameters from an in-house macro. These parameters were: interveinal distance (measured as a straight line between two adjacent minor veins, Figure 4.3A); average mesophyll cell length and mesophyll cell number (of the middle row of cells between two minor veins, Figure 4.3B); average bundle sheath cell diameter and number (for the cells directly contacting the vascular bundle, measured on the left-hand vein, Figure 4.3C) and internal (yellow) and external (blue) vascular bundle diameter (measured at the widest part on the left-hand vein,

Figure 4.3D).

For each trait, an ANOVA was performed to test for significant differences between genotypes. When $p<0.05$, pairwise t-tests were performed using the Bonferroni correction.

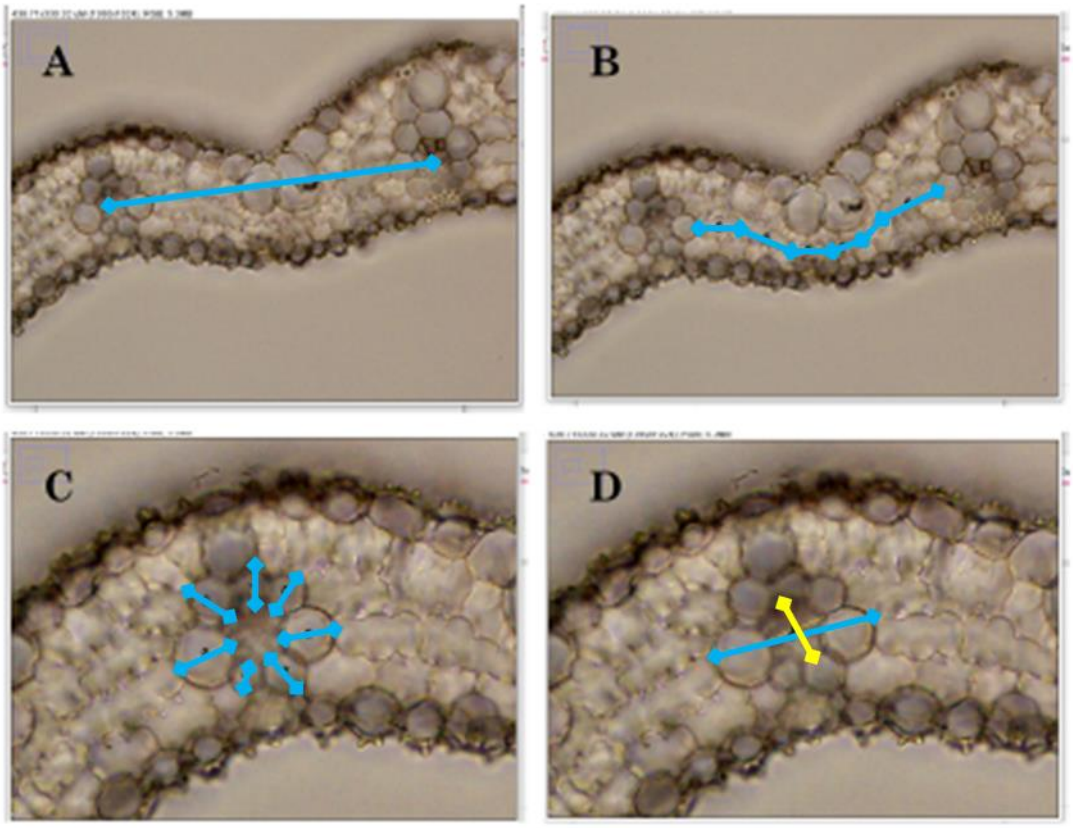
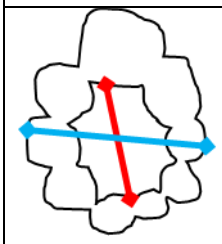
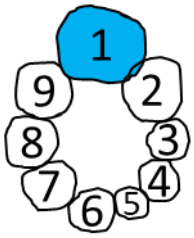


Figure 4.3 – Parameters measured for initial screening

A – Interveinal distance, B – mesophyll cell length, C – bundle sheath cell width, and D – inner and outer vascular strand diameter.

Detailed histology was performed on leaf seven using an in-house macro written in ImageJ and R. This macro collected more in-depth structural information, described fully in Table 4.3 below. As for the initial image analysis, data for a biological replicate was produced from three separate minor vein pairs.

Table 4.3 – Description of parameters measured for detailed histology measurements

Vasculature Parameters		
	Interveinal Distance	Straight line distance between the centre of two minor veins
	**Vein Area**	Total area from the outer perimeter of all bundle sheath cells
	Vascular Bundle Area	Total area from the inner perimeter of all bundle sheath cells
	Vein:Vascular Bundle Ratio	Ratio of vein area to vascular bundle area
	Vein Annular Radius	Average total bundle sheath radius, measured as: $0.5 \times \left(\frac{\max + \min \textit{ vein diameter}}{2} - \frac{\max + \min \textit{ vascular bundle diameter}}{2}\right)$
Bundle Sheath Cell Parameters		
	**Bundle Sheath Number**	Number of bundle sheath cells directly contacting the minor vein
	Bundle Sheath Area	Average individual bundle sheath cell area
	Total Bundle Sheath Cell Area	Total bundle sheath cell area around a minor vein
Mesophyll Cell Parameters		
	**Mesophyll Cell Number**	Number of mesophyll cells in the middle row between two minor veins
	Mesophyll Cell Width	Average mesophyll cell width of cells in the middle row between two minor veins
	Mesophyll Cell Length	Average mesophyll cell length of cells in the middle row between two minor veins
	Mesophyll Cell Perimeter	Average mesophyll cell perimeter of cells in the middle row between two minor veins
	Mesophyll Cell Area	Average mesophyll cell area of cells in the middle row between two minor veins
	Lobing	"lobiness" and "lobing area" describe the degree of average mesophyll cell lobing. Described fully in Figure 4.7.

4.2.3 Infra-Red Gas Exchange Analysis

4.2.3.1 Data Collection

Gas exchange measurements were taken on six individual LI-COR LI-6400XT InfraRed Gas Analyser machines with standard LED heads. All curves were performed on leaf seven. Conditions in the IRGA were set to match the environment, with TLeaf at 30°C, humidity maintained between 65% and 75% and flow rate set at 400µmols^{-1}. *A/Ci* curves were commenced after readings had stabilised.

A/*Ci* curves were performed at saturating *PPFD* 2000 µmol m^{-2} s^{-1}. For each plant, the curve was performed in two sections, with a 180s wait at each data point, with *A/Ci*1 at 400, 300, 250, 200, 150, 100, 50, 400ppm [CO_2] and *A/Ci*2 at 400, 500, 700, 800, 900, 1000, 1200, 1500ppm [CO_2]. The IRGAs were matched at every [CO_2].

4.2.3.2 Data Analysis

Data was fitted using the "*A/Ci* curve fitting utility version 2.0 (T)" (Sharkey, 2016), originally designed for tobacco.

4.3 Results

4.3.1 Screening Available Data

Interest in the IR64 "JL" lines had previously been narrowed down to 4 genotypes: JL75, JL80, JL81 and JL82. However, JL81 seed was unavailable at the time. For each genotype, several independent transformation events had been taken to the T1 stage. Thus, initial phenotypic data on every line in a genotype was screened to identify which lines exhibited the most extreme anatomical changes compared to IR64 whilst otherwise growing and developing normally. As described in the original paper, transgenic plants were compared to their null segregants to identify differences in leaf anatomy.

Data that had been produced by the Quick Lab is summarised in Figure 4.4 below. For all lines, neither leaf width (Figure 4.4A) nor leaf length (Figure 4.4B) significantly differed from the wild-type. Similarly, for vein density (Figure 4.4C) and interveinal distance (Figure 4.4D), all transformants appear to be within the normal wild-type range. There do, however, appear to be differences in bundle sheath and mesophyll cell anatomy between transformants. Several transformants have either unusually large or small bundle sheath cells (Figure 4.4E), including large cells in the JL80 lines and one JL75 line, and small cells in several JL82 lines. Mesophyll size data was not collected, however cell number does appear to vary (Figure 4.4F) with JL75 and JL80 having fewer mesophyll cells, and several JL82 transformants having more than the wild-type lines. As interveinal distance does not

significantly differ between these lines, it is possible that these differences in cell numbers are a result of alterations to mesophyll cell size compared to wild-type.

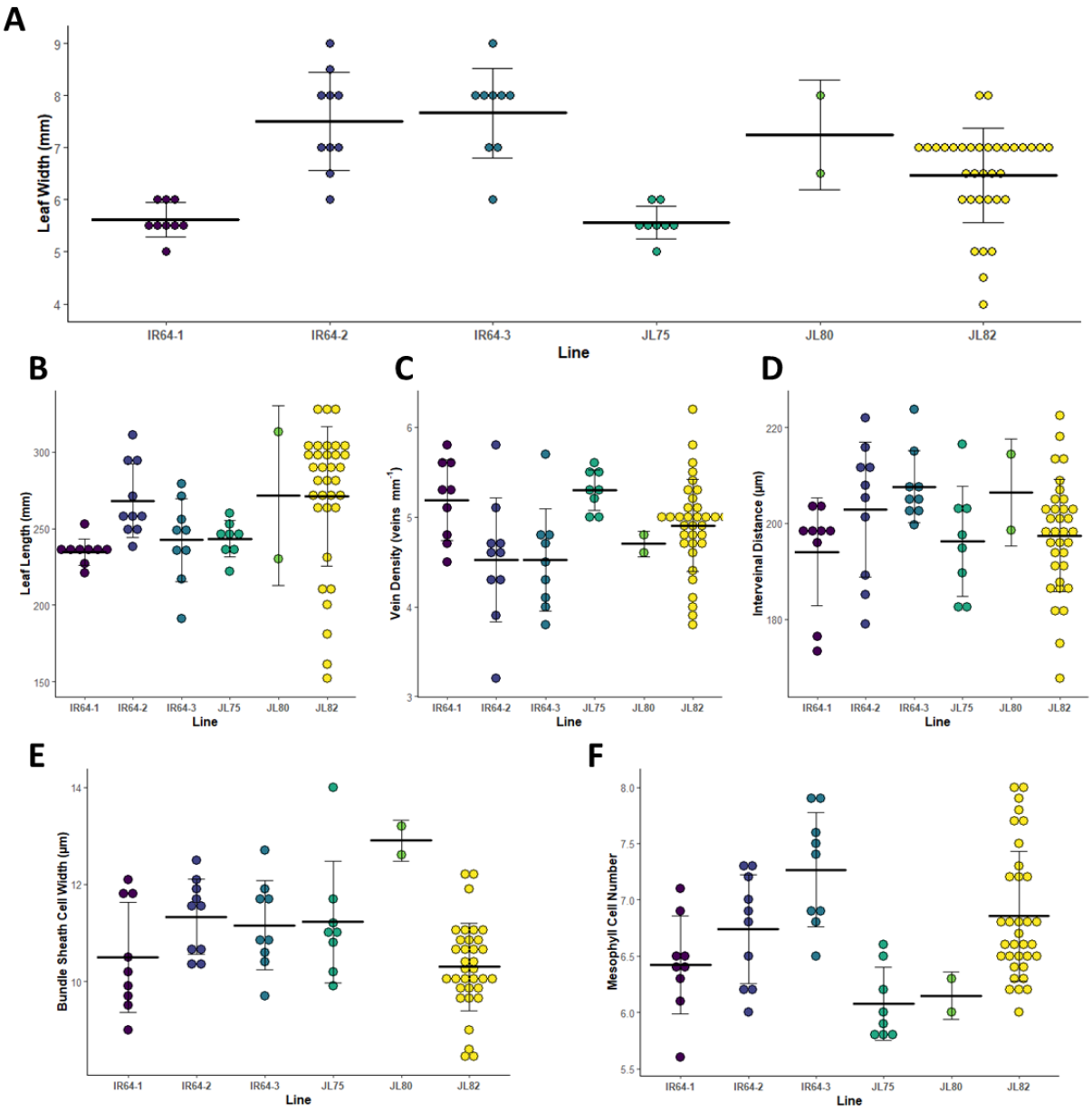


Figure 4.4 – Available anatomical data by genotype

For every graph, each point represents the average of all plants in a single transformation event. Horizontal bars show the mean value for the genotype, and error bars show standard deviation.

Each "JL" line was grown alongside an IR64 control, JL75 with IR64-1, JL80 with IR64-2 and JL82 with IR64-3.

All data is measured on leaf 6.

A – Leaf width.

B – Leaf length.

C – Interveinal distance.

D – Vein density (number of veins per mm).

E – Average bundle sheath cell width.

F – Average interveinal mesophyll cell number in the middle row.

Raw data provided by Xiojia Yin, and collected in the Quick lab at IRRI.

While on the average anatomy of the transgenic lines did not differ significantly from null transformants, the variation between the genotype suggested that individual transformation events could fit the criteria for this project, exhibiting variation in only one anatomical characteristic. Transformation events were chosen based on both phenotype and availability. Two independent JL75 lines, one JL80 line and three JL82 lines were planted. The JL80 plants were subsequently eliminated from further study due to a poor germination rate, slow-growth, and a high mortality rate of the resultant seedlings.

Below in Table 4.4, the individual lines carried forward are summarised. A full description, including construct identity is detailed in Appendix 4.1.

Table 4.4 – Summary of individual JL lines of interest
For each line, the designation in this chapter is shown, alongside the genotype and transgene expression level (as documented in Wang et al., 2017a), and the reason they were selected for consideration in this work.

Line	Genotype (Copy Number)	Transgene Expression	Phenotype of Interest
JL75A	JL75 PCR+ (3)	High	Smaller interveinal distance while maintaining mesophyll cell number – potential for smaller mesophyll cells.
JL75B	JL75 PCR+ (2)	High	Normal interveinal distance with lower mesophyll cell number – potential for larger mesophyll cells.
JL80	JL80 PCR+ (4)	N/A	Larger bundle sheath cells.
JL82A	JL82 PCR+ (1)	N/A	Smaller interveinal distance while increasing mesophyll cell number – potential for smaller mesophyll cells.
JL82B	JL82 PCR+ (1)	N/A	Smaller bundle sheath cells.
JL82C	JL82 PCR+ (1)	N/A	Smaller bundle sheath cells.

4.3.2 Preliminary Histology

For the remaining JL75 and JL82 lines, leaf four was taken from ten plants of each genotype at the five-to-six leaf stage, and hand-sectioned to perform a preliminary screen. All plants were screened for mesophyll size and number, bundle sheath cell size and number, as well as for vascular size and spacing. All ANOVA statistics and significance tables can be found in Appendix 4.2.

Previous data suggested that there was some variation in mesophyll cell number between veins. While we do not see this variation in these plants (Figure 4.5A), there are some significant differences in mesophyll cell length (Figure 4.5B). JL82A has shorter cells than JL82B ($p=0.0002$) and JL82C ($p=0.00001$). We also see that JL82C has longer cells than both JL75A ($p=0.0002$) and the wild type ($p=0.0032$). Further, bundle sheath cell number, which was not included in the original data, remains fairly consistent across all genotypes (Figure 4.5C), with the only significant difference being JL82A having fewer cells than JL82C ($p=0.0009$). Despite selecting for smaller-than-normal bundle sheath cells in the original available data, none of the lines used in this experiment showed a smaller bundle sheath cell phenotype on average (Figure 4.5D). However, JL82B has larger bundle sheath cells than both JL75B ($p=0.0005$) and JL82C ($p=0.0009$). We also examined the mesophyll to bundle sheath cell ratio (Figure 4.5E), as this differs between C_4 and C_3 plants. Amongst these lines, JL82C has a higher mesophyll:bundle sheath ratio than JL75A ($p=0.0005$), JL82A ($p=0.0008$) and the wild-type ($p=0.0029$).

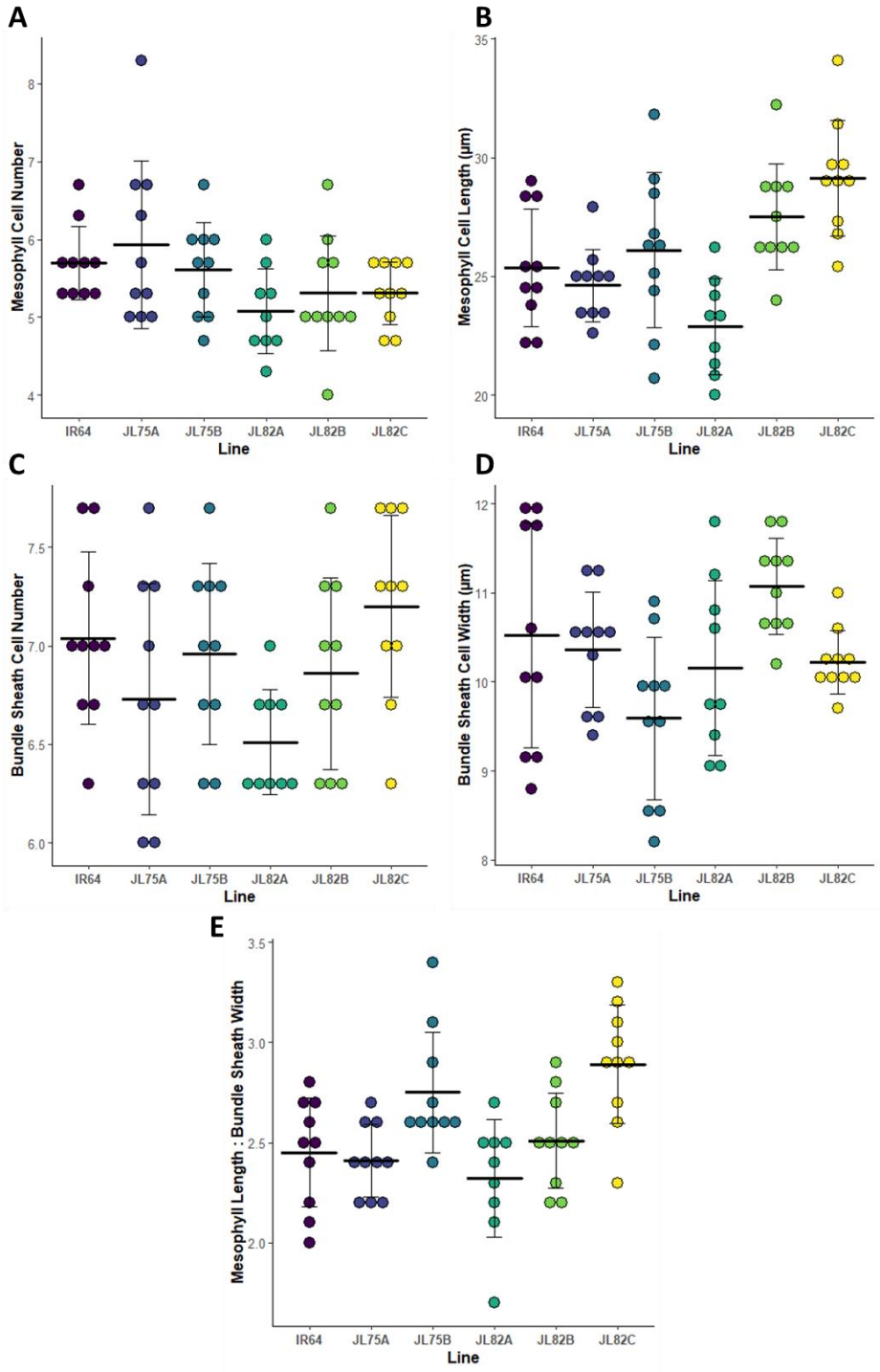


Figure 4.5 – Mesophyll and bundle sheath size and number in leaf four

For every graph, each point represents an individual plant. Horizontal bars show the mean value for the genotype, and error bars show standard deviation.

The WT line is IR64.

A – Mesophyll cell number (measured on the middle row between two veins) by genotype.

B – Mesophyll cell length by genotype.

C – Bundle sheath cell number (measured as bundle sheath cells directly contacting the vascular bundle) by genotype.

D – Bundle sheath cell length (measured as the widest diameter) by genotype.

E – Ratio of mesophyll length to bundle sheath length by genotype.

While we did see some significant variation in mesophyll cell and bundle sheath cell anatomy, this did not necessarily reflect the original data published by Wang *et al*. Thus, we also investigated differences in vasculature although these were not apparent in the original dataset. Neither the total vascular bundle (Figure 4.6A) nor total vein (Figure 4.6B) sizes vary by genotype in the plants used in this experiment. JL82A had a significantly smaller interveinal distance (Figure 4.6C) than all other lines ($p<0.0033$). The ratio of vascular bundle to vein size remained fairly constant throughout all lines examined (Figure 4.6D). All significance tables can be found in Appendix 4.2.

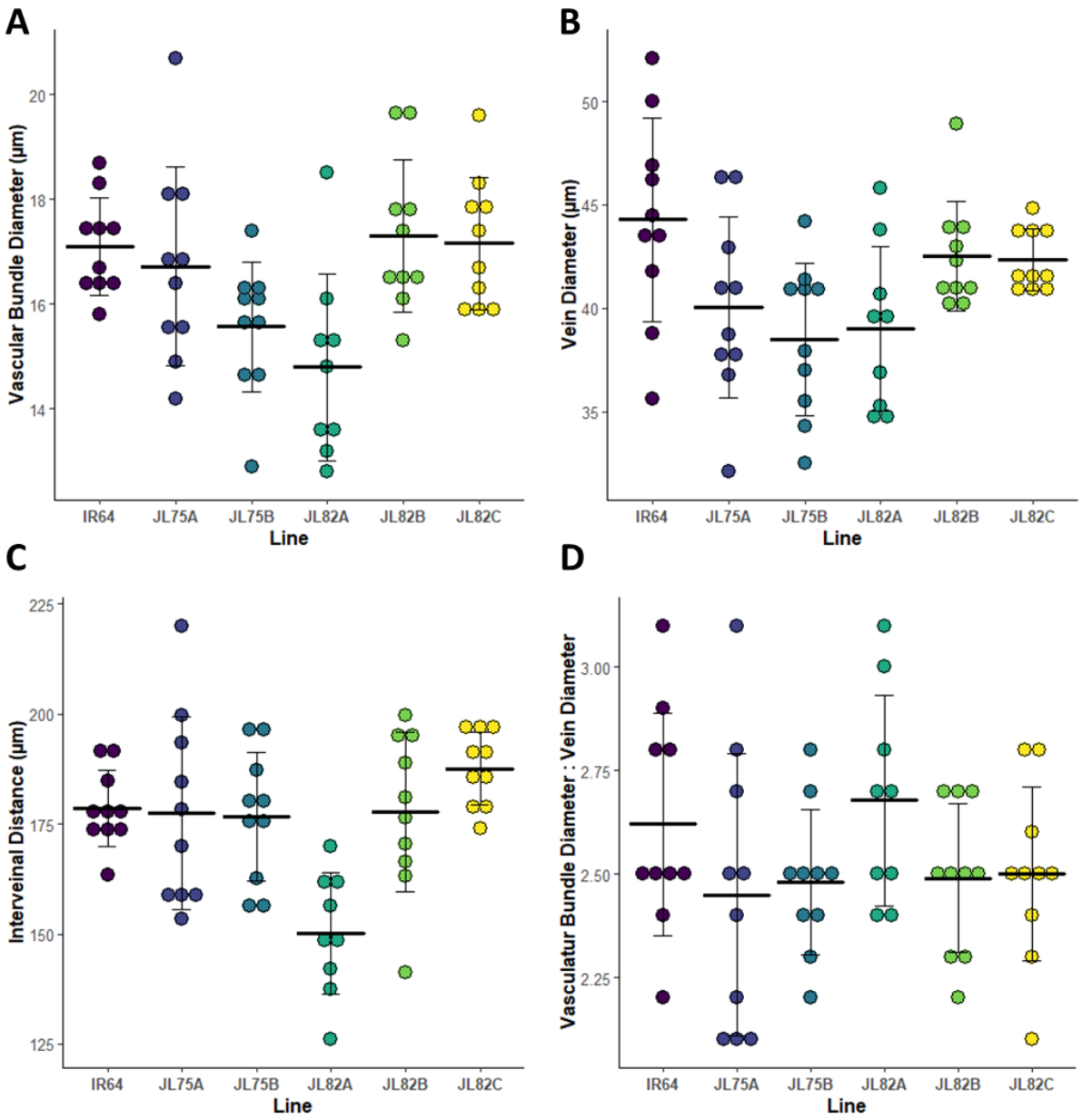


Figure 4.6 – Vascular phenotypes in leaf four

For every graph, each point represents an individual plant. Horizontal bars show the mean value for the genotype, and error bars show standard deviation.

The WT line is IR64.

A – Vascular bundle diameter (measured as the horizontal diameter between the bundle sheath cells) by genotype.

B – Vein diameter (measured as the horizontal diameter outside the maximum width of the bundle sheath cells) by genotype.

C – Interveinal distance (measured as the distance between the midpoint of two adjacent minor veins) by genotype.

D – Ratio of vascular bundle diameter to total vein diameter by genotype.

4.3.3 In-Depth Histology

Having screened ten plants per line using simplified structural parameters, five plants per line were carried forward for in-depth histology and gas exchange analysis. For each transgenic line, the plants with the most extreme phenotype were selected based on data for leaf four. For the IR64 wild-type, plants were preferentially selected if they exhibited very little deviation from the mean.

To investigate the irregular cell shape and uneven cell size distribution in the rice mesophyll, several additional structural parameters have been added to our analysis. One key parameter is lobing, which allows us to examine the 2D shape of the mesophyll cell. For this work, we examine this in two ways. The first measure is “lobiness” (shown in Figure 4.7A). Lobiness is defined as the ratio between the mesophyll cell perimeter (shown in black in Figure 4.7A) and the convex hull perimeter (in pink), with a value of 1 describing a cell which is completely devoid of lobes (thus, the cell perimeter equals the convex hull perimeter), and increasingly small numbers describing a greater amount of lobing. Our second parameter is lobing area, defined by the convex hull area less the cell area, normalised to the cell area (blue in Figure 4.7B). This acts as a proxy for number of, or depth of, lobes. In Sheffield, colleagues have previously observed a correlation between increased mesophyll cell lobiness and increased photosynthetic performance in similar JL lines in a Kitaake background, thus we were particularly interested in gathering more data about this parameter.

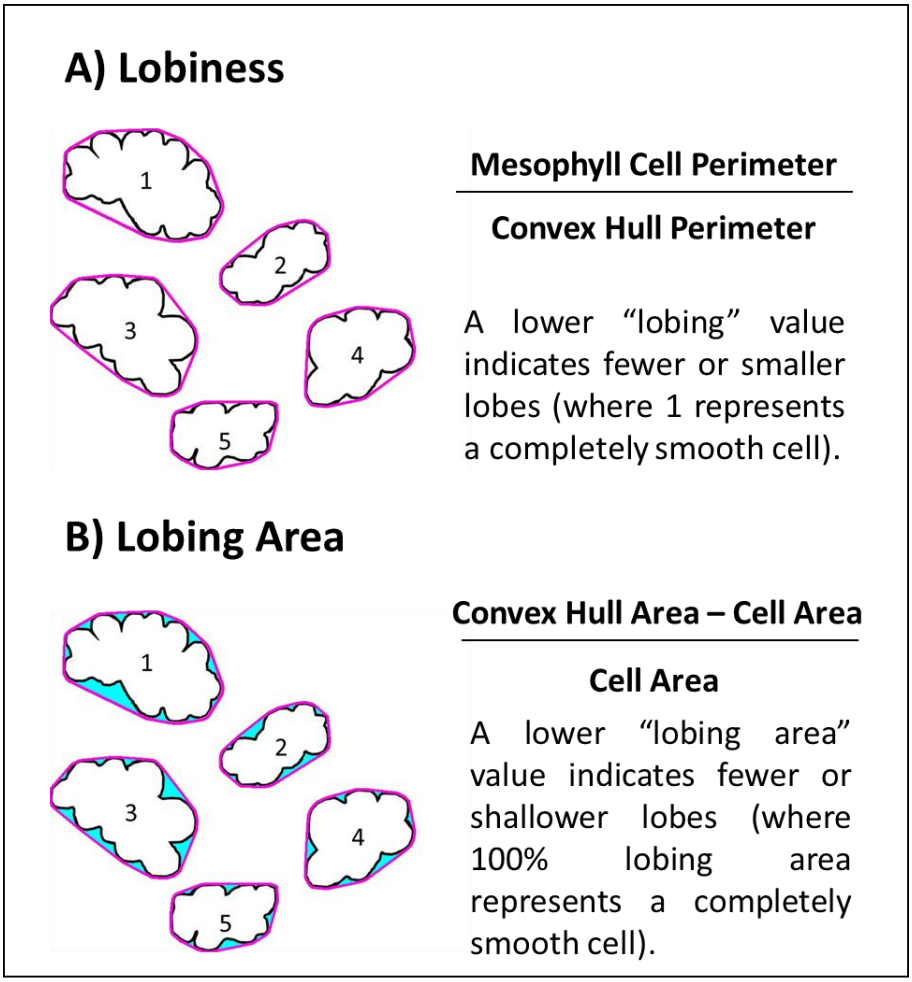


Figure 4.7 – Lobiness and lobing area

Schematic to represent how both lobiness and lobing area are defined in this analysis. Adapted from original graphics by Jen Sloan.

Of the lines tested in this project, lobing area (Figure 4.8A) was found to differ by genotype in an ANOVA ($p=0.01$), however no line differed significantly from another by t-test after correcting for multiple comparisons. Lobiness (Figure 4.8B) did not differ by genotype. All significance tables are shown in Appendix 4.3.

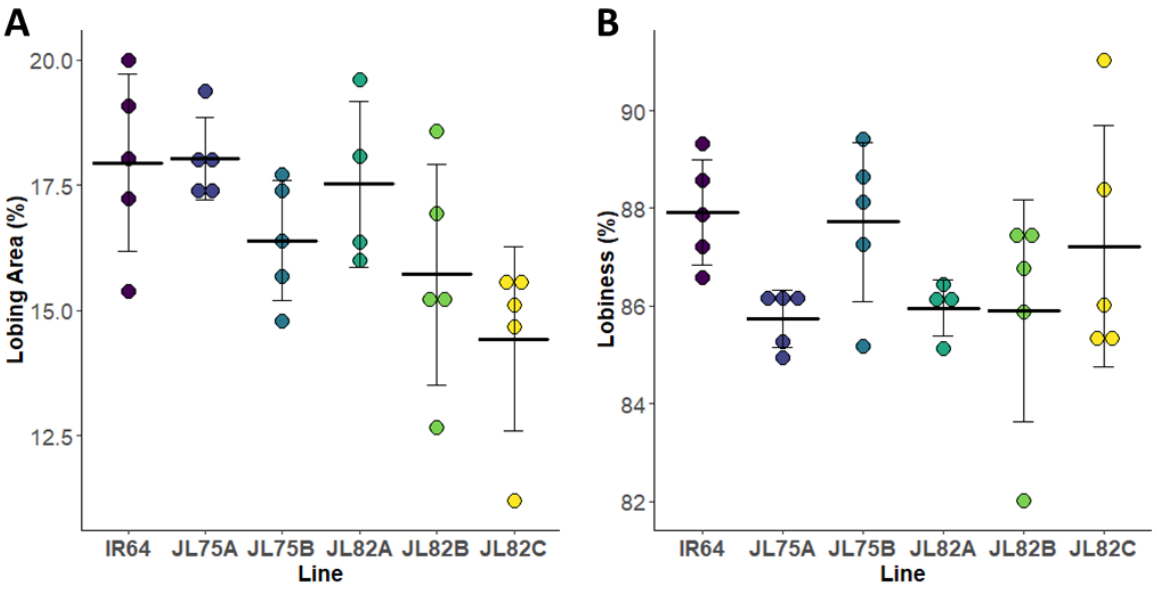


Figure 4.8 – Lobing phenotypes in leaf seven

For both graphs, each point represents an individual plant. Horizontal bars show the mean value for the genotype, and error bars show standard deviation.

The WT line is IR64.

A – Lobing Area by genotype.

B – Lobiness by genotype.

Further, while several different anatomical features varied significantly when looking at leaf four (Section 4.2), in leaf seven these structural changes are mostly not present (data shown in supplementary figure S4.1). Leaf seven anatomy was thoroughly investigated using the parameters described in Table 4.2. Mesophyll cell width, bundle sheath area and total area, interveinal distance, mesophyll cell number, vein area and annular radius were not found to differ by genotype in an ANOVA. Also, while mesophyll cell area, bundle sheath number and vein to vascular bundle ratio were shown to differ by ANOVA, no line differed significantly from another by t-test after correcting for multiple comparisons. In fact, the only significant differences by genotype were seen in mesophyll cell length, where JL75B is smaller than the wild type ($p=0.0005$), although these did not significantly differ at the leaf four stage; and in vascular bundle area, where JL82B has a larger vascular bundle area than JL75B ($p=0.0028$). All significance tables can be found in Appendix 4.3.

4.3.4 Gas Exchange Analysis

In addition to determining structural variation between genotypes, we were interested in which, if any, genotypes would vary in their photosynthetic efficiency. To observe the photosynthetic performance of these lines, *A*/*Ci* curves were produced for each line under varying [CO_2]. While leaf seven structural parameters were measured on five plants per genotype, several plants had to be excluded from gas exchange analysis due to instances of

rice blast damaging the leaves. Of all the genotypes tested, there is no significant difference in assimilation (Figure 4.9). Although JL82C appears to have a slightly higher assimilation rate than the other genotypes tested, more plants of this genotype were affected by rice blast than other lines, resulting in only two individuals giving useable photosynthetic data, making drawing conclusions from data for this line impossible.

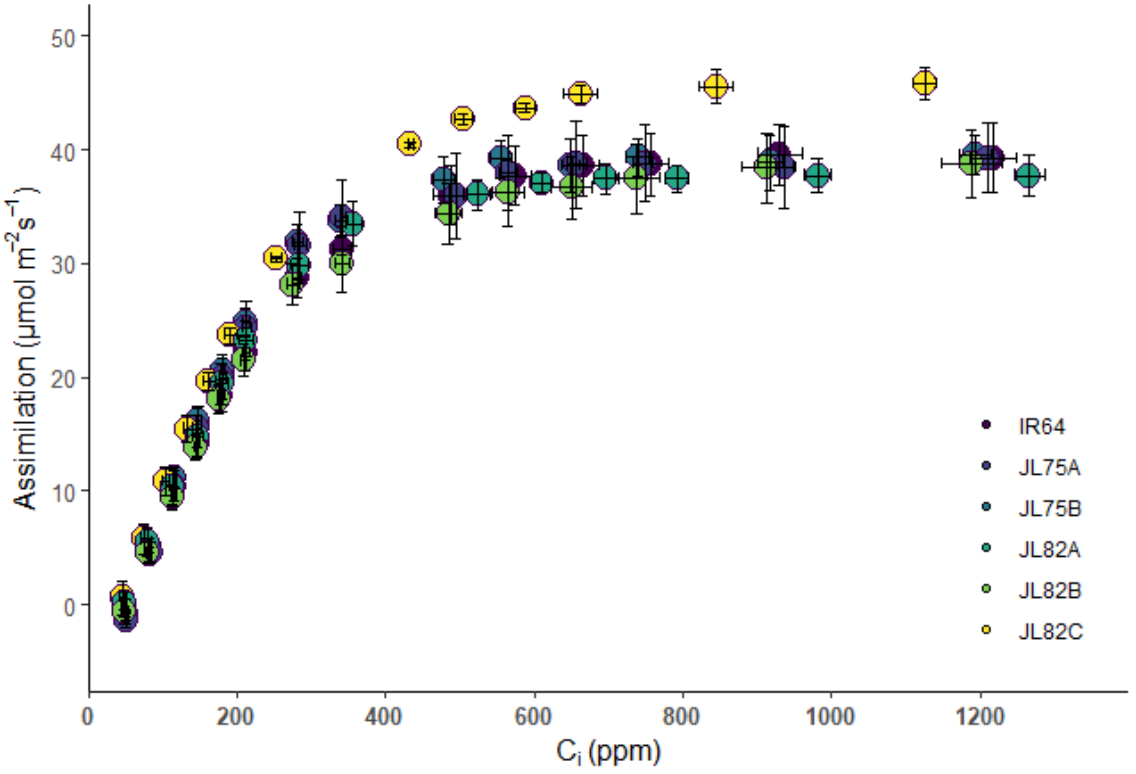


Figure 4.9 – *A/Ci* Curves for leaf seven

Each individual point represents an average of all plants within each genotype (n=4 for IR64, JL75A, JL75B and JL82B; n=3 for JL82A; n=2 for JL82C). Error bars represent SEM.

Furthermore, after fitting photosynthetic parameters to the FvCB model, there is no significant difference between any genotype for V_{Cmax} (Figure 4.10A), *A* at 1000 PAR (Figure 4.10B), *J* (Figure 4.10C) or *TPU* (Figure 4.10D). Significance tables can be found in Appendix 4.4.

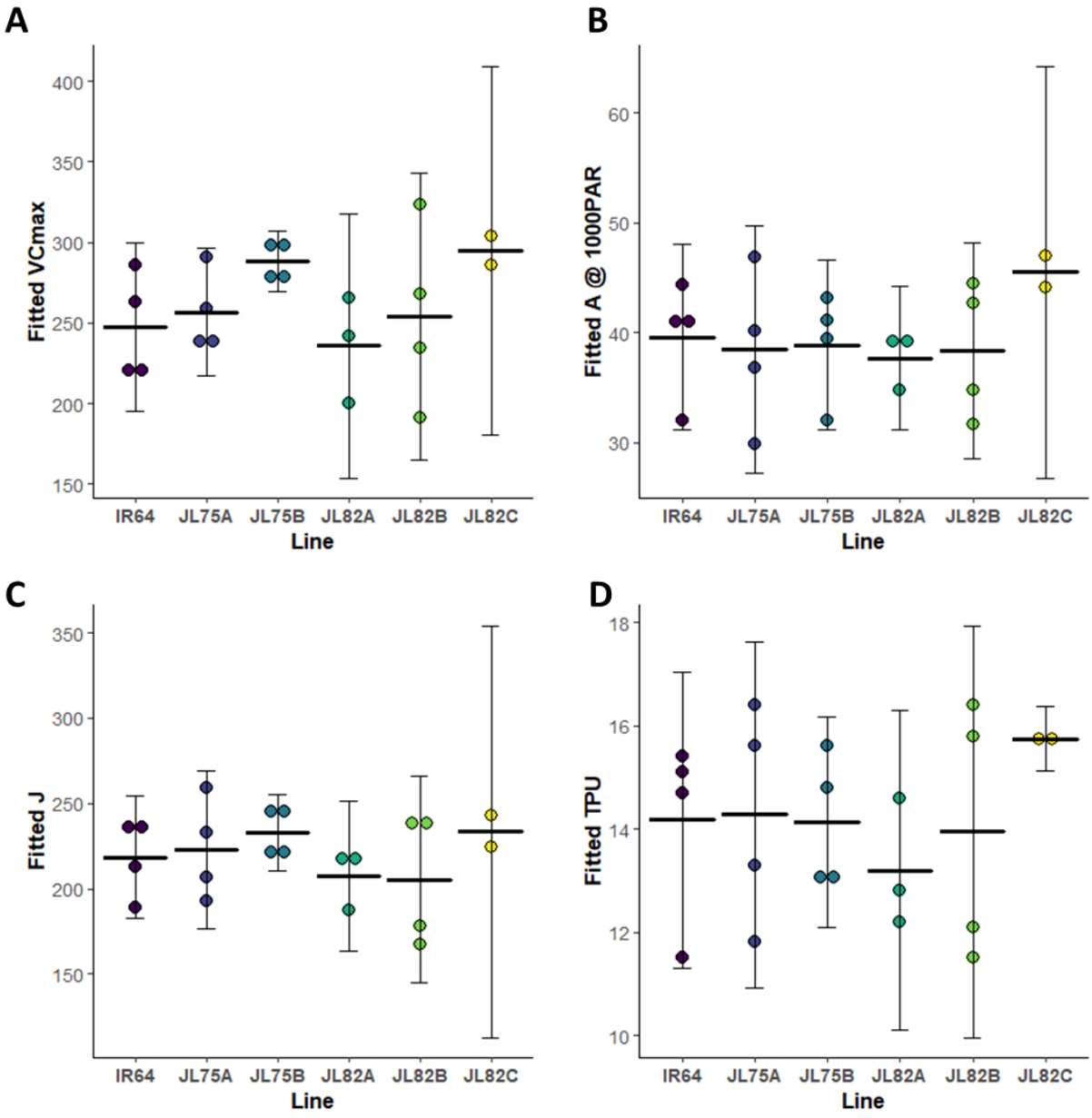


Figure 4.10 – Fitted parameters from gas exchange analysis

For every graph, each point represents an individual plant. Horizontal bars show the mean value for each line, and error bars show standard deviation.

A – Fitted V_{Cmax}

B – Fitted *A* @ 1000PAR

C – Fitted *J*

D – Fitted TPU

4.3.5 Structure-Function Correlations

While there was no significant difference in photosynthetic performance by genotype, there was sizeable variation in structural parameters between individuals.

Plants have several strategies to increase mesophyll cell surface area (and thus, chloroplast surface area aligned with the mesophyll cell surface): elongation of the cells, elongation coupled with cell division, producing more but smaller cells, and increased lobing. By examining the correlations between V_{Cmax} and various mesophyll structural parameters (Figure 4.11), we can examine if any of the individuals in our population have adopted these strategies. Neither increasing mesophyll cell area (Figure 3.11A) nor perimeter (Figure 4.11B) significantly impacted V_{Cmax}, although there is a small trend towards cells with increased perimeters having lower V_{Cmax}. While non-significant, increasing width (Figure 4.11C) and decreasing length (Figure 4.11D) both slightly correlate with increased V_{Cmax}, suggesting that mesophyll elongation is not a strategy adopted to increase surface area in these plants. Similarly, using our two measures of lobing, lobiness doesn't correlate with V_{Cmax} (Figure 4.11E) and there appears to be a slight negative correlation between lobing area and V_{Cmax} (Figure 4.11F), again implying that increased lobing is not a strategy adopted to increase surface area in these cells.

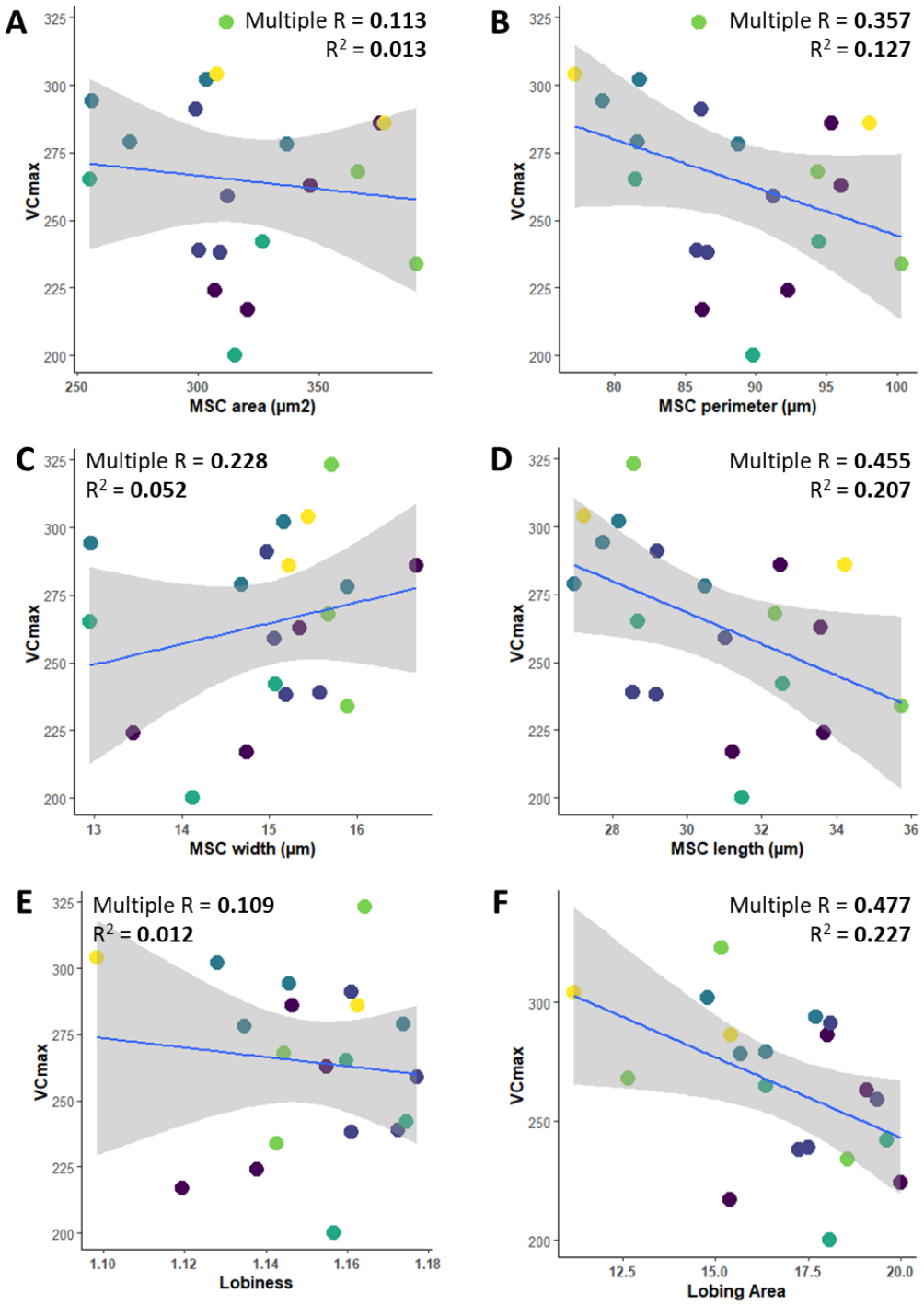


Figure 4.11 – Correlations between mesophyll parameters and V_{Cmax}

Each point represents an individual plant, coloured by genotype. The blue lines show the linear regression for the data, with a 95% confidence interval. Both R^2 and multiple R for the regression are shown.

A – Mesophyll area correlated with V_{Cmax}

B – Mesophyll perimeter correlated with V_{Cmax}

C – Mesophyll width correlated with V_{Cmax}

D – Mesophyll length correlated with V_{Cmax}

E – Lobiness correlated with V_{Cmax}

F – Lobing area correlated with V_{Cmax}

There was also no strong correlation between V_{Cmax} and any bundle sheath or vasculature trait (Figure 4.12 Panels A-G).

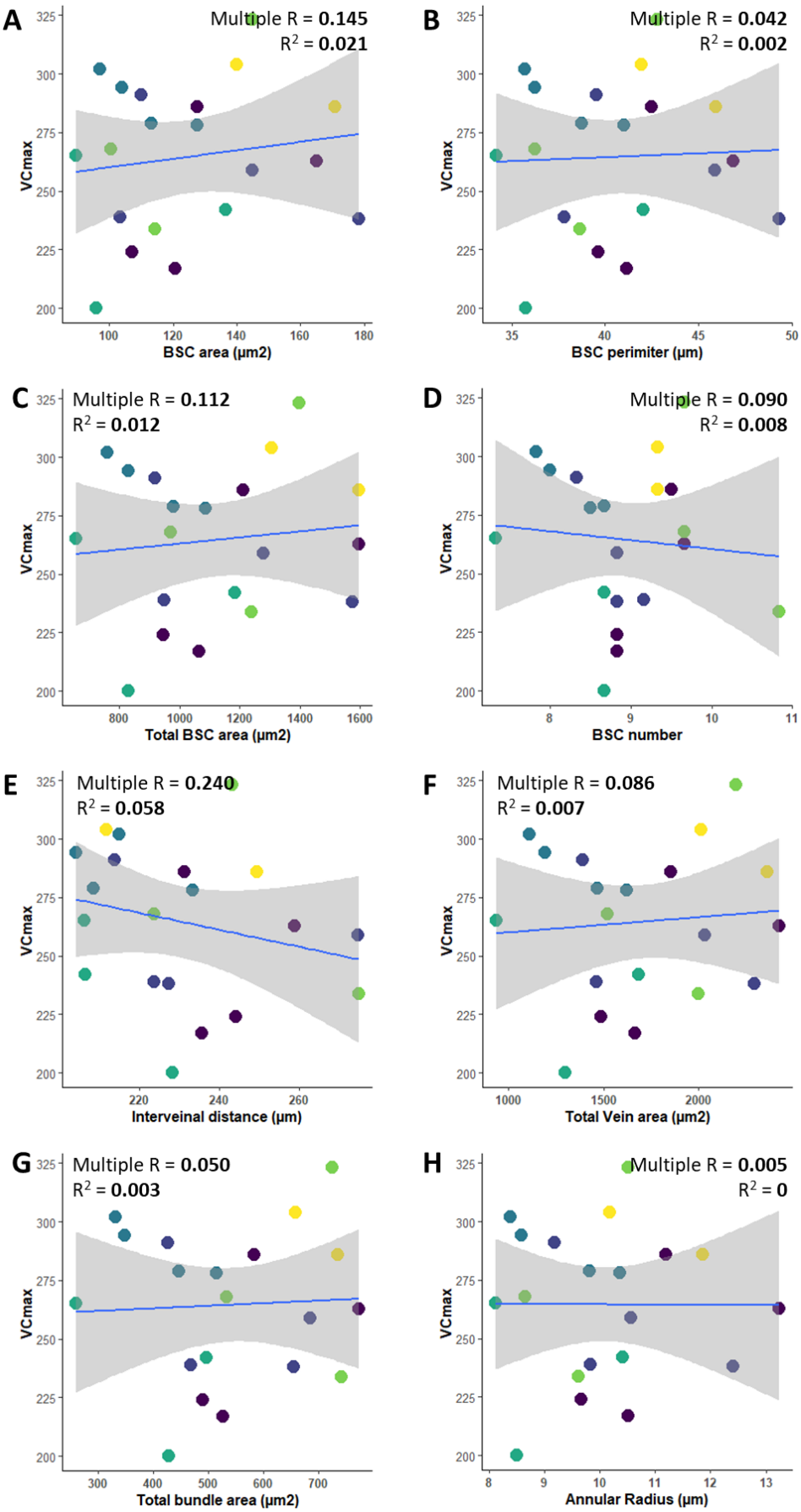

A
Multiple R = 0.145
R2 = 0.021
VCmax
BSC area (µm2)
B
Multiple R = 0.042
R2 = 0.002
VCmax
BSC perimiter (µm)
C
Multiple R = 0.112
R2 = 0.012
VCmax
Total BSC area (µm2)
D
Multiple R = 0.090
R2 = 0.008
VCmax
BSC number
E
Multiple R = 0.240
R2 = 0.058
VCmax
Interveinal distance (µm)
F
Multiple R = 0.086
R2 = 0.007
VCmax
Total Vein area (µm2)
G
Multiple R = 0.050
R2 = 0.003
VCmax
Total bundle area (µm2)
H
Multiple R = 0.005
R2 = 0
VCmax
Annular Radius (µm)

Figure 4.12 – Correlations between bundle sheath cell and vascular parameters and V_{Cmax}

Each point represents an individual plant, coloured by genotype. The blue line shows the linear regression for the data, with a 95% confidence interval. Both R^2 and multiple R for the regression are shown.

A – Average bundle sheath area correlated with V_{Cmax}

B – Average bundle sheath perimeter correlated with V_{Cmax}

C – Total bundle sheath area correlated with V_{Cmax}

D – Bundle sheath number correlated with V_{Cmax}

E – Interveinal distance correlated with V_{Cmax}

F – Total vein area correlated with V_{Cmax}

G – Total bundle area correlated with V_{Cmax}

H – Annular radius correlated with V_{Cmax}

4.4 Discussion

4.4.1 Experimental Limitations

There were several factors limiting the interpretation of the experiments laid out in this chapter.

Firstly, these plants were grown in glasshouses in the Philippines, not in controlled environment facilities. Whilst this is more representative of a farm environment where rice is grown, and is representative of the original conditions these lines were tested in, problems such as disease were harder to avoid. Several of the plants developed rice blast around the time of sampling for gas exchange measurements, and time constraints meant it was impossible to repeat the experiments with more replicates. This resulted in a smaller number of biological replicates than originally planned, and limited my statistical analyses.

Secondly, while there was an abundance of LI-COR 6400 machines available, these were not equipped with fluorescence heads. This meant that it was not possible to collect chlorophyll fluorescence data alongside gas exchange measurements. Fluorescence measurements would have made it possible to fit the gas exchange data with a more up-to-date tool, for example Moualeu-Ngangue et al., 2017, which incorporates a more accurate measure of CO_2 compensation point. As it stands, that data had to be fitted to an older tool (Sharkey et al., 2007), which is configured for tobacco, and estimates more parameters from the FvCB model. As such, several parameters had to be approximated, or modified to fit the model.

While both of these problems are specific to the environment the research was performed in, it is important to remember that a sizeable proportion of rice research is performed in the countries in which rice is grown, where overcoming limitation through expensive technological solutions may not be possible. If we wish the tools we are developing for rice

research to be useful globally, not just in affluent institutions, we must consider the likely experimental conditions, and limit the compulsory prerequisites to use the tools.

4.4.2 Efficiency of Approach

There have been several large-scale investigations examining the correlation between photosynthetic performance and complex traits such as mesophyll conductance (reviewed in Flexas et al., 2008), which span multiple phylogenetic groups. However, mesophyll conductance is a complex trait, and can be affected by multiple biochemical and structural parameters (Flexas et al., 2012). Comparing across species creates added layers of complexity, as they have adapted for growth in differing environments. Within a more homogenous rice population, forward genetic screens have demonstrated that narrow leaves with increased vein densities have higher photosynthetic capacity compared to wild type plants (Feldman et al., 2017). This work by Feldman and colleagues shows that multiple anatomical parameters have changed in these lines, with increased vein density correlating with a reduction in interveinal mesophyll cell number. In this chapter, we approached this structure-function correlation from a reductionist perspective: could lines with changes in only an individual anatomical parameter be identified, and would changes in this parameter correlate with differing photosynthetic performance? To this end, we performed more detailed histological measurements, for example measuring cross-sectional area, perimeter and shape of mesophyll cells, rather than just length. The data presented in this chapter do not support a correlation between any individual structural trait measured and increased photosynthetic activity. This is not to say that no individual trait does influence photosynthetic performance, particularly due to the low replicate number for the photosynthetic measurements presented in this chapter, however it does raise an interesting consideration when examining structure function relationships, namely that several anatomical and biochemical changes are likely acting together to promote photosynthetic improvement.

The aim of this project was to find a line or lines of rice that could be used as a tool in order to parameterise the 3D model of rice photosynthesis which is under development (Yi et al., in preparation). While a wide range of different anatomical parameters were included within the dataset, there was very low correlation to photosynthetic efficiency. This could suggest that the differences in leaf anatomy examined here are not a primary driver for photosynthetic improvement. It could also mean that in order to explore the relationship between anatomy and photosynthetic performance effectively, improved anatomical data, for example analysing cellular structures in three dimensions, would be needed. However, this approach would be both more time-intensive and costly, and rely on access to specialist

equipment. There is sizeable variation observed between wild-type individuals, and thus a larger sample size would be necessary to account for this. It is possible that the small correlations observed with the low n possible in this experiment would also be clearer with a larger sample size. This variation in the wild-type line is not unique to the experiments performed in this chapter. As shown in Figure 4.4, all three groups of wild-type rice showed sizeable variation for all parameters measured. This phenomenon is also not unique for field-grown plants, and is reflected in data from plants grown by colleagues in controlled environment facilities (personal communication). Given that IR64 is a highly inbred strain of rice, it is perhaps surprising that there is so much variation between individuals. However, this variation within a genotype is important to bear in mind when planning experiments: how many plants will be needed to be confident that a phenotype is real? Unless phenotypes are very extreme, it is possible they will be eclipsed by natural variation. However, simply growing more plants would require substantially more space (which particularly for plants grown in controlled environments may not be possible, or financially feasible), and substantially more hours of work to analyse the plants, which again, may not be possible due to time and financial limitations. One approach to help alleviate this problem is to perform pilot experiments, which is the approach taken in this chapter.
As the aim of this project was to identify rice lines with clear differences in anatomy and photosynthesis for further use in contributing to the photosynthetic model, a larger dataset would not have furthered this goal. The data presented in this chapter were sufficient to convincingly rule out these lines as candidates, and to allow colleagues to shift their focus to alternative approaches to model parameterisation.

Furthermore, the more interesting data from my analysis involved changes to the shape and size of mesophyll cells. The rice mesophyll is not, however, a ubiquitous layer, with the cells closest to the veins measuring twice or more the length of those in the middle layer. Cells nearer the epidermis are also smaller than those in the middle row. This structural variation in mesophyll cell size needs to be taken into account for modelling approaches.

Using transgenic plants instead of exploiting existing biodiversity could still be an advantageous strategy for several reasons. As the plants are identical in genetic background with the exception of changes to a single parameter (for example, cell wall constitution), it is more reasonable to attribute changes in photosynthetic performance to that anatomical parameter. When comparing genetically distinct lines, there are many more variables which could alter performance. If this approach were to continue, it would perhaps be more efficient to identify mutations in genes homologous to those known to fundamentally alter leaf anatomy in other species, and manipulate those genes in rice to create a novel study population.

4.4.3 Implications for the C_4 Rice Project

The lines used in this chapter were a subset of those originally generated in the hope of engineering C_4 traits into rice (Wang et al., 2017a). The overexpression of these individual transcription factors, identified as candidate regulators of Kranz anatomy, did not yield sizeable anatomical changes, and the authors concluded that the exploitation of different promoters and analysis of gene combinations will be necessary in future efforts. Work to attempt to explore additional promoter elements was carried out during this PhD project (described in Appendix 4.5), but did not successfully yield any new candidates.

It has been predicted that a transition to C_4 photosynthesis will require, at minimum, the alteration of tens of genes simultaneously (Pick et al., 2011). Until recently, stacking multiple genes in a single transgenic individual has required crossing single gene transformants until a line with all desired insertions is obtained. This is both time and labour intensive. More recent advances in cloning technology, such as Golden Gate cloning (Weber et al., 2011), allow for the transformation of up to twelve genes in a single construct, vastly reducing the number of generations needed to stack multiple traits, and allowing for more rapid development of plants with alterations to multiple genes. While this technology is useful for ongoing multigene transformation experiments, there is still an urgent need for more rice-specific or monocot-specific promoter modules, as existing publicly-available promoter modules are heavily focussed towards the transformation of Arabidopsis and other dicots (Engler et al., 2014).

Furthermore, while the use of constitutive promoters is well established in plant research, their use may not always be suitable for altering developmental trajectories. In other model plant species, such as Arabidopsis, constitutive overexpression of key developmental genes has caused significant disruption to the structure and function of the plant, whilst restricting overexpression of developmental genes to specific phases of development has resulted in more subtle phenotypes (e.g. Efroni et al., 2008). The use of alternative promoters has been less well studied in rice and other monocots, likely due in part to the experimental challenges of doing so. One significant example is recent work by Sun and colleagues, who have demonstrated that constitutive overexpression of PLASTOCHRON1 (a known regulator of leaf initiation and maturation (Mimura et al., 2012)) in maize resulted in slow-growing plants that would not flower, while restricted overexpression resulted in plants with both higher biomass and improved seed yield (Sun et al., 2017). Notable promoters available in rice are the cell-specific promoters PEPC (which targets expression to the mesophyll (Matsuoka et al., 1994)), and various bundle sheath-specific promoters. These have been used successfully in rice to direct cell-specific gene expression. In these cases, expression under

these cell-specific promoters serves to reinforce or alter gene expression in cells which have developed their identity and are already differentiating. While initiation of the vasculature occurs as early as the P2 stage of leaf development (Itoh et al., 2005), it continues to develop and mature over time (van Campen, 2016, thesis), thus it is possible that much of this tailored gene expression will occur later in leaf development when the leaf anatomy is already largely set. As the anatomy of the rice leaf seems to be determined very early in development (discussed in Chapter 2), it is possible that for effective, co-ordinated anatomical changes required for the introduction of C_4 traits, genes of interest should be targeted to much earlier in leaf development.

An alternative approach is to extend the phase of leaf development amenable to transformation. Indeed, overexpression of morphogenic genes *ZmBABYBOOM* and *ZmWUSCHEL2* improved transformation rates in maize, sorghum, sugarcane and rice (described in Hofmann, 2016), although these expression cassettes needed excising prior to callus regeneration, due to the negative growth impacts of constitutive overexpression of early development genes. An initial aim of this PhD project was to investigate the possibility of overexpressing similar morphogenic genes in a temporally specific manner (comparable to Sun et al., 2017) to create a stable rice line more amenable to genetic engineering. However, this would have required the identification of novel, phase-specific promoters. Unfortunately, as described in Appendix 4.5, this was not achieved, although a method to investigate promoter function was developed. If this work were to be continued, it could be interesting to examine if testing promoters cloned from maize would be able to overcome some of these issues.

Of course, promoters are not the only regulatory elements mediating plant gene expression. Enhancer, insulator and silencer sequences have important roles in mediating this gene expression (reviewed in Bilas et al., 2016). Furthermore, it has also been demonstrated that in many grasses, there is preferential binding of transcription factors to motifs within coding sequences, including for those transcription factors directing the cell-specific gene expression needed for the development of C_4 traits (Burgess et al., 2019). In addition to a need for a larger range of rice-compatible promoters for use in genetic modification, it would be interesting to consider if other cis elements could be exploit it to manipulate gene expression, and potentially provide more tools to help make C_4 rice a reality.

Chapter 5 –
Final Discussion

5.1 Overview

This thesis aimed to investigate and document *de novo* development in rice leaves. The work follows on from a previous PhD project between Sheffield and IRRI, which showed that the P3/P4-1 transition is an important stage in leaf development, where photosynthetic competence is first established (van Campen et al., 2016; van Campen, thesis, 2016). I aimed to complement existing genetic (RNAseq) and physiology (chlorophyll fluorescence) data with novel biochemical (metabolomics) and microscopy (SEM) approaches, and collected data on both the established P3, P4-1 and P5 stages, and those intermediate to P3 and P4-1. By optimising these techniques to collect data on individual primordia, I have been able to provide novel data on temporal rice leaf development, and document the P3/P4-1 transition for the first time.

To date, the majority of studies of leaf development in rice and other monocot crops have considered the classical spatial developmental trajectory. The work presented in this thesis adds to a growing body of work looking at temporal development in the rice leaf, summarised in Table 5.1, and is of interest from a fundamental developmental biology point of view. Further, this thorough characterisation of “normal” leaf development will be a useful benchmark when assessing the impact of alterations to leaf development, for example by assessing the impact of a transgene. The techniques developed are highly sensitive, require very small amounts of sample, and have the potential to be built upon for targeted screens.

Finally, the study of the relationship between leaf anatomy and photosynthetic performance described in this thesis did not demonstrate correlation between any individual anatomical trait and photosynthetic performance. This highlights the flexibility of the developing leaf to compensate for minor changes in structure, and raised useful considerations for ongoing development of a novel photosynthetic model.

Table 5.1 – Summary of temporal rice leaf development.

	P3	P3/P4-1 Transition	P4	P5
Epidermal phenotype Gross phenotype (Itoh et al., 2005) Detailed phenotype (This thesis, Chapter 2)	**Gross**: Initiation of epidermal cells. **Detailed**: No visible epidermal differentiation. Cells beginning to organise in files.	**Detailed**: Rapid differentiation of macrohairs (II), stomata (III), prickle hairs (IV) then papillae (V) in basipetal direction. All cells organised in files.	**Gross**: Differentiation of epidermal cells. **Detailed**: All cells mature at the tip of the leaf. Basipetal developmental gradient.	**Gross**: Maturation of epidermal cells.
Photosynthetic phenotype (van Campen et al., 2016; Kusumi et al., 2010)	Undetectable or negligible levels of electron transport, localised to leaf tip.		From P4-1, measurable electron transport, strongest at the leaf tip. Electron transport increases from P4-1 through P4-12, in a basipetal gradient.	Strong, measurable electron transport.
Genetic phenotype (van Campen et al., 2016; van Campen, thesis, 2016)	Low expression of photosynthesis genes. High expression of genes implicated in cell division, transcription, and of growth and development transcription factors. Expression of epidermal cell initiation genes.		Increasing expression of photosynthesis genes. High expression of genes implicated in cell wall development and remodelling. Expression of epidermal cell differentiation genes.	High expression of photosynthesis genes. Low expression of genes involved in growth and development. Low expression of epidermal cell genes.
Metabolic phenotype (This thesis, Chapter 3)	Low citrate, malate, fumarate. Mid oxaloacetate. High α-ketoglutarate.	Low citrate, malate, fumarate. α-ketoglutarate decreases sharply after IV. Oxaloacetate peaks at III. Metabolic fingerprinting distinguishes I to early IV from late IV to VI.	Low citrate, malate, fumarate (although higher than P3), low α-ketoglutarate, high oxaloacetate. Citrate and malate detectable by MS/MS. Metabolic fingerprinting distinguishes P4-1 from P4-2 and P4-3 primordia.	High citrate, malate, fumarate. Low α-ketoglutarate, oxaloacetate. Citrate and malate detectable by MS/MS.

5.2 A Temporal Approach to Rice Leaf Development

The division of leaf development into distinct stages is a useful tool in developmental biology (Sylvester et al., 1996). Traditionally, this staging has been based on tracking leaf morphological features (Poethig, 1997), and microdissections of developing primordia in dicots have proven a useful tool to follow these changes. However, in monocots, the majority of leaf developmental studies exploit the developmental gradient present along a maturing leaf, where the distal tip of the leaf is fully differentiated, and moving towards the base of the leaf the tissue is less differentiated. There are many advantages to this approach: a complete gradient of epidermal cells is present along a single leaf (Luo et al., 2012), and detailed transcriptomic and metabolomic data can be gleaned by dividing the leaf blade in to sections (Pick et al., 2011; Wang et al., 2014). The size of the leaf blade also means that light microscopy approaches can be used to visualise development, and that it is easy to sample sufficient material for downstream use in molecular studies. While a perfectly valid approach, focussing on this spatial view of leaf development alone has two limitations from a developmental biology standpoint. Firstly, the spatial resolution of such studies is often in the range of mms to tens of mms, whereas many developmental signals act over much shorter distances, in the region of tens to hundreds of µm. Secondly, in a maturing leaf blade, differentiation will inevitably be informed by the already mature distal tissue. This mature tissue will already be physiologically active, thus generating a wide range of signals which will influence the process of patterning in proximal 'naïve' tissue downstream. Further, pre-patterning of this tissue will inform its differentiation.

Taking a temporal approach to studying rice (and other monocot) leaf development has the potential to bridge some of the fundamental knowledge gaps in this area. Leaf primordia which have formed *de novo* do not possess this same degree of pre-patterning, and while they can receive long-distance signals from older leaves, internal developmental signalling within the growing leaf is restricted. Thus, studying temporal leaf development is closer to addressing the fundamental question of how naïve tissue co-ordinates the development of the biochemical, physiological and structural processes necessary to form a functioning leaf. Microdissections of developing leaves have led to a broad grouping of leaf development in to six stages: P1 to P6 (Itoh et al., 2005). Approaches such as SEM and *in situ* hybridisation have proven useful to study the very earliest (pre-P3) stages of rice leaf development, detailing leaf initiation (Kawakatsu et al., 2006; Miyoshi et al., 2004; Satoh et al., 1999) and early vascular development (Scarpella et al., 2003; Yasui et al., 2018). A temporal approach has also been used to examine chloroplast biogenesis in the P4 stage of development (Kusumi et al., 2010), where the leaf blade elongates to around ten times its original length. Broader transcriptomic approaches encompassing multiple developmental stages have also

been performed in the foliar and husk leaves of maize (Wang et al., 2013) and developing rice leaves (van Campen et al., 2016). In the context of engineering alterations to leaf development, these are key studies, showing that genetic regulators of Kranz anatomy in the maize leaf are acting between the P2 and P5 stages, after which leaf anatomy is set; and that photosynthetic competence in the rice leaf is first established during the P3/P4-1 transition phase. While it seems apparent that these temporal transitions are key stages in leaf development, there remains little additional information in these areas. In this thesis, I have focussed on the P3/P4-1 transition in rice, and attempted to fill some of these knowledge gaps.

5.3 Epidermal Differentiation in Early Leaf Development

In Chapter 2, I use SEM to examine the histogenesis of the epidermis of microdissected primordia. These data show that between the P3/P4-1 transition development can be further subdivided into six discrete stages, termed I to VI, by epidermal morphology. Interestingly, the tip of the growing primordium has fully mature epidermal cells before even the P4-1 stage. This poses an interesting question: even at P4-1, the leaf primordium is completely encapsulated by the sheath of older leaves. While capable of photosynthetic activity, this occurs only at low levels, and the leaf is not autotrophic at this stage (Kusumi et al., 2010), (van Campen, 2016, thesis). Similarly, it is unlikely that young leaf primordia, growing as they are in a humid environment, are responsible for regulation of transpiration in the plant, yet mature stomata are visible on the tips of the primordia as early as the IV stage of the P3/P4-1 transition (see Chapter 2).

In the mature monocot epidermis, stomata appear in cell files flanking underlying veins (Stebbins and Shah, 1960). As the vascular bundle begins to differentiate as early as the P2 stage (Itoh et al., 2005), while immature stomata do not appear until the P3/P4-1 transition (this thesis, Chapter 2), the observed pattern suggests either a mechanism of inhibition in epidermal cells directly contacting the vasculature, or an induction of the stomatal lineage for cell files beyond the vein. Recent observations that overexpression of *ZmSHORTROOT1* (*ZmSHR1*) in rice bundle sheath cells results in additional stomatal rows (Schuler et al., 2018), and that knocking out its rice homologues (*osshr1 osshr2* double mutant) results in a reduction in stomatal density (Wu et al., 2019), suggests that vascular-derived signals may be partially responsible for stomatal lineage identity in rice. In maize, the interaction between SHR and SCARECROW (SCR) has been shown to be vital in the establishment of normal vascular patterning and the establishment of Kranz anatomy (Slewinski et al., 2014). In rice, SCR is expressed in immature stomata (Kamiya et al., 2003), and is directly activated by SPEECHLESS (Wu et al., 2019), the transcription factor which controls the first asymmetric

division into the stomatal lineage. Thus, it has been hypothesised that the interaction between SCR and SHR creates a zone to establish stomatal fate (Nunes et al., 2020). This hypothesis fits with gene expression data which show that *OsSHR1* (LOC_Os07g39820), *OsSCR1* (LOC_Os11g03110) and *OsSCR2* (LOC_Os12g02870) levels all decrease over the course of development from P3 to P5 (van Campen et al., 2016). It is possible that this early stomatal differentiation is not a result of the need for functional stomata, but rather to set up early cell identity, allowing for the classical linear differentiation pattern observed as the leaf matures.

In addition to producing this novel image dataset, I also created a developmental index to accurately stage leaf primordia for use in future experiments. While it is possible to broadly characterise these young leaf primordia by eye, my data show a total epidermal remodelling at the tips of primordia between 1mm and 6mm in length which, particularly at the early stages, results in a very small window of each developmental stage. This index allowed me to stage primordia accurately by measuring their total length during microdissection, and later assigning a developmental stage based on my available data. This in turn led to my being able to use entire primordia for future molecular experiments. While similar indices were not produced for wild rice lines, my data suggest that the epidermal patterning process is conserved over a range of species. Thus, this technique could be an appropriate method to assess the developmental stage of leaf primordia across the *Oryza* genus.

5.4 Molecular Differentiation of Early Development

While morphology-based classification is interesting, and a useful starting point, it is not appropriate to classify developmental stage solely in this manner. Using length as an absolute marker is clearly inappropriate, as total leaf length will be affected by when the leaf emerged, and by environmental conditions (Humphries and Wheeler, 1963). Further, morphological features can be affected by genetic and environmental factors (Bar and Ori, 2014). For example, a mutation may affect primordia length but not the developmental stage of the leaf. Thus, it was important to combine morphological data with at least one more alternative measure of maturation.

In *Arabidopsis thaliana*, transcripts with distinct expression profiles throughout development have been used to characterise leaf developmental maturity. For example, Efroni et al., 2008 compared wild-type plants with “normal” leaf development to mutants with a reduction in levels of core growth regulators and aberrant leaf growth phenotypes. Theoretically, this approach could be translated to rice. Indeed, when studying spatial development in maize and rice leaves, transcript levels have proven useful markers for gauging maturity (Wang et

al., 2013). Further, RNAseq has been carried out on staged P3, P4-1 and P5 primordia (van Campen et al., 2016), and this data could be used to narrow down potential marker genes. However, for this approach to be effective, the expression of at least tens of genes must be measured simultaneously, which requires sufficient material to be extracted. Due to the tiny size of individual rice leaf primordia, tens must be pooled to gain sufficient RNA from a single extraction. In order to test the gene profiles of the six potential transition stages described above, the time taken to collect the samples would be prohibitive and impractical. While I feel a transcriptomics-based approach would be interesting, the methodology would need to be optimised for much smaller samples if it were to be practical. Further, it would realistically be a task for multiple people to undertake, particularly at the sample collection stage. Finally, transcripts are not the final product of gene expression, and do not necessarily correlate to changes in the levels of proteins or metabolites which are directly involved in cell function.

Thus, I explored an alternative approach in order to map molecular markers to developing primordia through mass spectrometry, as mass spectrometry is highly sensitive, and can be performed on very small samples. In the experiments described in this thesis, I was able to successfully develop an approach which allowed for metabolic fingerprinting of individual leaf primordia. This allowed me to explore the P3/P4-1 transition intermediate stage in a way which would not have been possible through a transcriptomics approach, or even through a similar GC-MS approach to existing metabolic studies of leaf development. I thus performed metabolite fingerprinting to produce the first global look at the metabolome of developing rice leaves.

Metabolite fingerprinting is a powerful tool that has proven useful in several areas of biology, including disease modelling, drug toxicity studies and medical diagnostics (Kosmides et al., 2013). It has also been used in plants for a variety of functions, including examining population level differences (Mumm et al., 2016; Overy et al., 2005; Silveira-Sotelo et al., 2015) and tracking the progress of pathogen infection (Parker et al., 2009). In the field of developmental biology, metabolite fingerprinting has been useful to differentiate between different embryonic stages in both *Drosophila melanogaster* and zebrafish (Dhillon et al., 2019; Thuy An et al., 2014).

Unlike targeted panels, where certain metabolic pathways are examined, metabolite fingerprinting is an untargeted and unbiased approach. Given this experiment represents a novel metabolic screen, it was important to gain this global overview in order to inform future experiments. Further, while there are subsets of metabolites which would certainly be interesting to look at, especially in terms of photosynthetic development (for example, Calvin-Benson cycle metabolites (Arrivault et al., 2009) or those known to differ between C3

and C4 plants), changes in individual pathways are not necessarily indicative of changes in the overall progress of leaf development. Where there have been studies looking at more global metabolic changes in monocots, a GC-MS approach has been applied to maize and to rice to simultaneously quantify and identify a broad subset of metabolites in the sample (Pick et al., 2011; Wang et al., 2014). However, this technique uses substantially larger sample sizes than could be produced in this project, with the original protocol using at least 50mg of tissue for a single extraction (Roessner et al., 2000), compared to the individual primordia weighing under 1mg used in the experiments presented in this thesis. Both the GC-MS approach used in previous studies, and the fingerprinting technique described in this thesis attempt to give a global picture of the metabolome. It is important to note that due to the chemical diversity of metabolites, no single protocol can measure every metabolite in a sample. To gain the most global overview possible, compromise conditions which ionise as many metabolites as possible are used. This differs from other “omics” approaches such as transcriptomics, where the chemical homogeneity of the sample allows for a truly global overview.

As discussed in Chapter 3, a substantially lower proportion of metabolites change in expression levels (up or down) between P3 and P4-1 than transcripts. Conversely, there is a much higher proportion of metabolites changing in expression (up or down) between P4-1 and P5 than there are transcripts. So while we do not see the same overall pattern in the two datasets, this is to be expected as the two approaches are addressing different, but complementary, questions. While transcriptomics data does provide a developmental snapshot, it also informs of the changes the leaf is preparing to make, as transcript accumulation does not lead to instantaneous protein translation or, indeed, downstream metabolic effects. Conversely, metabolomics data is an instantaneous look at which metabolic processes were already occurring at the time of sampling. It reflects the products of gene expression changes which have already had the time to affect leaf development. This data fits with observations that many key differentiation events, such as the acquisition of photosynthetic competence and the establishment of epidermal patterning, occur during the P3/P4-1 transition, as fundamental structures and biochemical machinery necessary for leaf function must be set down before metabolic processes associated with a functioning leaf can occur.

In Chapter 2, I describe my observations that the P3/P4-1 transition can be further sub-divided from two stages to six based on morphological differences. However, several of these transitions represent fairly subtle changes, thus it was not clear if their global metabolic profiles would be distinct. In Chapter 3, I show that four groups can be distinguished by their global metabolite profiles: P3, an early transition group comprising

grades I to early IV, a late transition group comprising grades late IV to VI, and P4-1. This metabolic distinction between the transitioning primordia occurs at the same time as the most drastic epidermal changes, where specialised cells are rapidly maturing and papillae begin to appear, suggesting this is indeed a key point in cell remodelling and differentiation.

These data suggest that there may be key temporal transitions in early rice leaf development. Other metabolomic and transcriptomic work on developing monocot leaf blades concluded that differentiation occurs as a result of morphogenic gradients (Pick et al., 2011; Wang et al., 2014), and that there is no sudden temporal switch comparable to the rapid cell cycle arrest front seen in *Arabidopsis thaliana* (Andriankaja et al., 2012). I see this more gradual developmental gradient reflected along P4 developing leaf blades, with the complete transition from undifferentiated to differentiated epidermal cells, and the metabolomic fingerprints of these samples sitting intermediate to P4-1 and P5 leaves. However, when examining the P3/P4-1 transition there is certainly a point during epidermal differentiation where dramatic and rapid changes can be seen (IV/V), and there is also a visible difference in the clustering of these primordia by metabolomic analysis. If a cell cycle arrest front also occurred in rice, I would expect to see a sudden increase in cell size to reflect a switch from cell proliferation to cell elongation. This was not something that I observed in the primordia imaged in these studies, however cell size was not a parameter I measured, and my focus was only on epidermal cells, not on those within the leaf. Further experiments would be needed to conclusively investigate this possibility. It certainly seems that there is a very early differentiation of the cells at the distal tip of the leaf primordium, presumably reflecting exit from cell division, and I speculate that this is needed to organise and give identity to the cell files which appear later in development along the more immediate proximal region.

5.5 The Relationship Between Leaf Anatomy and Photosynthesis

The initial goal of the experiments described in Chapter 4 was to identify lines with alterations in individual anatomical traits, and to see if these changes correlated with altered photosynthetic performance in order to better parameterise a 3D model of photosynthesis currently under development. However, for the lines tested in this study, there was no correlation between any structural trait and photosynthetic ability. While this is perhaps unsurprising, it raises some important considerations when interpreting data correlating leaf anatomy with photosynthetic function. As discussed in Chapter 4, anatomical changes are often multifactorial – for example, decreased interveinal distance may be accompanied by decrease in mesophyll cell number – making it difficult to specifically identify which attributes are beneficial, and thus, difficult to know which traits to target in breeding programs and

genetic studies. Furthermore, photosynthesis is reliant on the interplay between leaf anatomy and biochemistry, thus anatomical changes do not occur in isolation. In fact, recent data have shown that in some situations, plants with 'improved' leaf anatomy will compensate for this by lowering the activity of core photosynthetic enzymes, re-equilibrating their photosynthetic performance to that of wild-type plants (Yi et al., in preparation / personal communication).

If similar compensation methods are commonplace in response to other minor anatomical changes, this is perhaps a key area for future research focus. Projects such as the C_4 rice project require a total anatomical and biochemical remodelling of the leaf, but many approaches, such as those utilising transformation of a single gene, may encounter sizeable problems if the leaf is hardwired to set its photosynthetic performance at a predetermined level.

5.6 Future Work

There are several experiments which would either add to, or lead on from the work presented in this thesis.

In Chapter 2, I began to explore the conservation of epidermal patterning amongst other diverse *Oryza* species. It would be interesting to further explore this evolutionary angle, initially by examining further wild rice species. This could be further expanded to compare to other agronomically important monocots, including maize, wheat and barley, for a more comprehensive evolutionary development view in monocots. A similar atlas of normal temporal development in other crop plants could also be a useful tool informing genetic improvement in these species.

A more detailed study of the pattern of cell size and division in these small developing primordia would also be interesting. Cell volume data can be obtained by reconstructing 3D images of the plant tissues from confocal microscopy images (Wuyts et al., 2010), and using software to identify these cells and calculate structural information. Such an approach has been useful to study epidermal cells in *Arabidopsis thaliana* and *Cardamine hirsute* primordia (de Reuille et al., 2015), and tailoring this technique to young rice leaves could produce a wealth of useful developmental information.

As mentioned in Chapter 3, the coronavirus pandemic curtailed the remaining experiments planned for this chapter, and these would be a logical place to continue this work. After using MS/MS to identify key marker metabolites for each developmental stage, there are many biochemical pathways which could be further explored using a metabolomics approach. To further explore the establishment of photosynthesis, development of a protocol to target

Calvin-Benson cycle metabolites would be particularly interesting, and existing literature of such panels in *Arabidopsis thaliana* would form a good starting point. It would also be interesting to perform a comparative study between rice and maize primordia, to identify when key marker metabolites change during the development of C3 and C4 plants. While it would be particularly interesting to take a GC-MS approach (see Pick et al., 2011) to simultaneously fingerprint samples and identify compounds, this would require prohibitively large sample sizes. One alternative would be to design an LC-MS-based targeted screen, as there is a lower prerequisite sample size. The data presented in Chapter 4 could inform which compounds would be interesting to investigate further due to their differential expression over the course of development.

Another interesting application of mass spectrometry would be utilising MS imaging (MSI) to visualise the distribution of certain metabolites in the developing primordium. MSI is becoming an increasingly popular tool to investigate the spatial distribution of metabolites, and in plants matrix-assisted laser desorption/ionisation (MALDI) imaging has proved most popular to date (reviewed in Heyman and Dubery, 2016). Recent work in the Fleming lab has also utilised desorption electrospray ionisation (DESI) imaging – which does not require the same intensive sample preparation as MALDI imaging – to examine TCA cycle metabolites in potato apical meristems (Tout, 2019, thesis), which also has the potential to be applied to leaf primordia in rice.

5.7 Final Remarks

In this thesis, I describe the work I have performed to document the P3/P4-1 transition in developing rice leaves. In Chapter 2, I have presented a novel SEM image dataset, and described how this transition can be further sub-divided based on epidermal morphology. In Chapter 3, I have integrated this information on developmental transition with the first global metabolic study of temporal development in the rice leaf. I have also tested the hypothesis that certain anatomical changes will translate to changed photosynthetic performance in Chapter 4, with the negative results informing colleagues' work in the development of a 3D model of photosynthesis.

This work adds to the growing body of knowledge concerning the temporal aspects of rice leaf development, and I hope it will serve as a useful framework to assess leaf developmental changes in future studies.

6. Bibliography

Agilent Technologies (2011). Time-of-Flight Mass Spectrometry Technical Overview.

Allahverdiyeva, Y., Mustila, H., Ermakova, M., Bersanini, L., Richaud, P., and Ajlani, G. (2013). Flavodiiron proteins Flv1 and Flv3 enable cyanobacterial growth and photosynthesis under fluctuating light. Proc. Natl. Acad. Sci. U. S. A. *110*, 4111–4116.

Alonso, A., Marsal, S., and Julià, A. (2015). Analytical methods in untargeted metabolomics: State of the art in 2015. Front. Bioeng. Biotechnol. *3*, 1–20.

Ananyev, G., Gates, C., Kaplan, A., and Dismukes, G.C. (2017). Photosystem II-cyclic electron flow powers exceptional photoprotection and record growth in the microalga Chlorella ohadii. BBA - Bioenerg. *1858*, 873–883.

Andriankaja, M., Dhondt, S., DeBodt, S., Vanhaeren, H., Coppens, F., DeMilde, L., Mühlenbock, P., Skirycz, A., Gonzalez, N., Beemster, G.T.S., et al. (2012). Exit from proliferation during leaf development in Arabidopsis thaliana: A Not-So-Gradual Process. Dev. Cell *22*, 64–78.

Annesley, T.M. (2003). Ion suppression in mass spectrometry. Clin. Chem. *49*, 1041–1044.

Arrivault, S., Guenther, M., Ivakov, A., Feil, R., Vosloh, D., Van Dongen, J.T., Sulpice, R., and Stitt, M. (2009). Use of reverse-phase liquid chromatography, linked to tandem mass spectrometry, to profile the Calvin cycle and other metabolic intermediates in Arabidopsis rosettes at different carbon dioxide concentrations. Plant J. *59*, 826–839.

Asai, K., Satoh, N., Sasaki, H., Satoh, H., and Nagato, Y. (2002). A rice heterochronic mutant, mori1, is defective in the juvenile-adult phase change. Development *129*, 265–273.

Banerjee, S., and Mazumdar, S. (2012). Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. Int. J. Anal. Chem. *2012*, 1–40.

Bar, M. and Ori, N. (2014). Leaf development and morphogenesis. Development. *141*. 4219-4230.

Barton, M.K. (2010). Twenty years on: The inner workings of the shoot apical meristem, a developmental dynamo. Dev. Biol. *341*, 95–113.

Bellasio, C., Farquhar, G.D., and Bellasio, C. (2019). A leaf-level biochemical model simulating the introduction of C2 and C4 photosynthesis in C3 rice : gains , losses and metabolite fluxes. New Phytol. *223*, 150–166.

Bilas, R., Szafran, K., Hnatuszko-Kona, K. and Kononowicz, A. K. (2016). Cis-regulatory elements used to control gene expression in plants. Plant Cell Tiss Organ Cult. *127*. 269-287.

Breeding, P., and Hall, E. (2007). The complex history of the domestication of rice. Ann. Bot. *100*, 951–957.

Brown, W. V. (1975). Variations in anatomy, associations and origins of Kranz tissue. Am. J. Bot. *62*, 395–402.

Brown, R.H., and Byrd, G.T. (1993). Estimation of bundle sheath cell conductance in C4 species and O2 insensitivity of photosynthesis. Plant Physiol. *103*, 1183–1188.

Buckley, T.N., and Warren, C.R. (2014). The role of mesophyll conductance in the economics of nitrogen and water use in photosynthesis. Photosynth. Res. *119*, 77–88.

Burgess, S.J., Reyna-Llorens, I., Jaeger, K., and Hibberd, J.M. (2019). Genome-wide transcription factor binding in leaves from C3 and C4 grasses. Plant Cell. *31*, 2297–2314.

von Caemmerer, S., Quick, W.P., and Furbank, R.T. (2012). The development of C4 rice: current progress and future challenges. Science (80-.). *336*, 1671–1672.

Caine, R.S., Yin, X., Sloan, J., Harrison, E.L., Mohammed, U., Biswal, A.K., Dionora, J., Chater, C.C., and Coe, R.A. (2018). Rice with reduced stomatal density conserves water and has improved drought tolerance under future climate conditions. New Phyt. *221.* 371-384.

van Campen, J. (2016). Early leaf development in rice (Oryza sativa L.) : structure, physiology, and gene expression. Thesis.

van Campen, J., Yaapar, M.N., Narawatthana, S., Lehmeier, C., Wanchana, S., Thakur, V., Chater, C., Kelly, S., Rolfe, S. a, Quick, W.P., et al. (2016). Combined chlorophyll fluorescence and transcriptomic analysis identifies the P3 / P4 transition as a key stage in rice leaf photosynthetic development. Plant Physiol. *170*, 1655–1674.

CGIAR. New irrigation technique can ease drought effects for rice farmers. (2014). Accessed at https://ccafs.cgiar.org/fr/blog/new-irrigation-technique-can-ease-drought-effects-rice-farmers on 27/08/19

Chaffey, N.J. (1983). Epidermal structure in the ligule of rice (Oryza sativa L.). Ann. Bot. 13–21.

Chen, M., and Blankenship, R.E. (2011). Expanding the solar spectrum used by photosynthesis. Trends Plant Sci. *16*, 427–431.

Chen, Y., Mao, H., Wu, N., Din, A. M. U., Khan, A., Zhang, H. and Yuan, S. (2020). Salicylic acid protects photosytem II by alleviating photoinhibition in Arabidopsis thaliana under high light. Int. J. Mol. Sci. *21*, 1-17.

Chiou, T.-J., and Bush, D.R. (1998). Sucrose is a signal molecule in assimilate partitioning. Proc. Natl. Acad. Sci. U. S. A. *95*, 4784–4788.

Clark, S.E., Running, M.P., and Meyerowitz, E.M. (1995). CLAVATA3 is a specific regulator of shoot and floral meristem development affecting the same processes as CLAVATA1. Development *121*, 2057–2067.

Cobb, J.N., Juma, R.U., Biswas, P.S., Arbelaez, J.D., Rutkoski, J., Atlin, G., Hagen, T., Quinn, M., and Ng, E.H. (2019). Enhancing the rate of genetic gain in public-sector plant breeding programs: lessons from the breeder' s equation. Theor. Appl. Genet. *132*, 627–645.

Collard, B.C.Y., and Mackill, D.J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *363*, 557–572.

Collard, B.C.Y., Gregorio, G.B., Thomson, M.J., Vergara, G. V, Laborte, A.G., Nissila, E., Kretzschmar, T., Cobb, J.N., Breeding, F.P., Division, B., et al. (2019). Transforming rice breeding: re-designing the irrigated breeding pipeline at the International Rice Research Institute (IRRI). Crop Breeding, Genet. Genomics *1*, e190008.

Demmig-Adams, B., and Adams III, W.W. (1992). Photoprotection and other responses of

plants to high light stress. Annu. Rev. Plant Physiol. Plant Mol. Biol. *43*, 599–626.

Dhillon, S.S., Torell, F., Donten, M., Lundstedt-Enkel, K., Bennett, K., Rännar, S., Trygg, J., and Lundstedt, T. (2019). Metabolic profiling of zebrafish embryo development from blastula period to early larval stages. PLoS One *14*, 1–14.

Du, B., Kruse, J., Winkler, J.B., Alfarray, S., Schnitzler, J.P., Ache, P., Hedrich, R., Rennenberg, H., and Foyer, C. (2019). Climate and development modulate the metabolome and antioxidative system of date palm leaves. J. Exp. Bot. *70*, 5959–5969.

Duursma, R.A. (2015). Plantecophys - An R package for analysing and modelling leaf gas exchange data. PLoS One *10*, 1–13.

Eastmond, P.J., Van Dijken, A.J.H., Spielman, M., Kerr, A., Tissier, A.F., Dickinson, H.G., Jones, J.D.G., Smeekens, S.C., and Graham, I.A. (2002). Trehalose-6-phosphate synthase 1, which catalyses the first step in trehalose synthesis, is essential for Arabidopsis embryo maturation. Plant J. *29*, 225–235.

Efroni, I., Blum, E., Goldshmidt, A., Eshed, Y., Efroni, I., Blum, E., Goldshmidt, A., and Eshed, Y. (2008). A protracted and dynamic maturation schedule underlies Arabidopsis leaf development. Plant Cell *20*, 2293–2306.

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.M., Werner, S., Jones, J.D.G., Patron, N.J., and Marillonnet, S. (2014). A Golden Gate modular cloning toolbox for plants. ACS Synth. Biol. *3*, 839–843.

Eveland, A.L., and Jackson, D.P. (2012). Sugars, signalling, and plant development. J. Exp. Bot. *63*, 3367–3377.

Evenson, R.E., and Gollin, D. (2003). Assessing the impact of the green revolution, 1960 to 2000. Science (80-.). *300*, 758–762.

Farquhar, G.D., and Sharkey, T.D. (1982). Stomatal conductance and photosynthesis. Annu. Rev. Plant Physiol. *33*, 317–345.

Farquhar, G.D., von Caemmerer, S., and Berry, J.A. (1980). A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. Planta *149*, 78–90.

Feldman, A. B., Leung, H., Baraoidan, M., Elmido-Mabilangan, A., Canicosa, I., Quick, W. P., Sheehy, J. and Murchie, E. H. (2017). Increasing leaf vein density via mutagenesis in rice results in an enhanced rate of photosynthesis, smaller cell sizes and can reduce interveinal mesophyll cell number. Fronteirs in Plant Sci. *8*, 1-10.

Fitzgerald, M.A., and Resurreccion, A.P. (2009). Maintaining the yield of edible rice in a warming world. Funct. Plant Biol. *36*, 1037–1045.

Flexas, J., Ribas-Carbó, M., Diaz-Espejo, A., Galmés, J., and Medrano, H. (2008). Mesophyll conductance to CO2: Current knowledge and future prospects. Plant, Cell Environ. *31*, 602–621.

Flexas, J., Barbour, M.M., Brendel, O., Cabrera, H.M., Carriquí, M., Díaz-Espejo, A., Douthe, C., Dreyer, E., Ferrio, J.P., Gago, J., et al. (2012). Mesophyll diffusion conductance to CO2: An unappreciated central player in photosynthesis. Plant Sci. *193–194*, 70–84.

Furbank, R.T., and Taylor, W.C. (1995). Regulation of photosynthesis in C3 and C4 plants: a molecular approach. Plant Cell *7*, 797–807.

Garnier, S. (2018). viridis: Default Color Maps from 'matplotlib'. R package version 0.5.1.

https://CRAN.R-project.org/package=viridis

Ge, S., Sang, T., Lu, B., and Hong, D. (1999). Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc. Natl. Acad. Sci. *96*, 14401–14405.

Giuliani, R., Koteyeva, N., Voznesenskaya, E., Evans, M.A., Cousins, A.B., and Edwards, G.E. (2013). Coordination of leaf photosynthesis, transpiration and structural traits in rice and wild relatives (genus Oryza). Plant Physiol. *162*, 1632–1651.

Global Rice Science Partnership (GRiSP) (2013). Rice Almanac.

Gowik, U., and Westhoff, P. (2011). The path from C3 to C4 photosynthesis. Plant Physiol. *155*, 56–63.

Gu, L., Pallardy, S.G., Tu, K., Law, B.E., and Wullschleger, S.D. (2010). Reliable estimation of biochemical parameters from C3 leaf photosynthesis-intercellular carbon dioxide response curves. Plant, Cell Environ. *33*, 1852–1874.

Guijas, C., Montenegro-Burke, J.R., Domingo-Almenara, X., Palermo, A., Warth, B., Hermann, G., Koellensperger, G., Huan, T., Uritboonthai, W., Aisporna, A.E., et al. (2018). METLIN: A technology platform for identifying knowns and unknowns. Ann. Chem. *90*, 3156–3164.

De Guzman, C., and Oard, J. (2018). Genetics and breeding system for cytoplasmic and genetic male sterility in rice. In Protecting Rice Grains in the Post-Genomic Era, p. e85191.

Hagemann, W., and Gleissberg, S. (1996). Organogenetic capacity of leaves: the significance of marginal blastozones in angiosperms. Plant Syst. Evol. *199*, 121–152.

Hanson, M.R., Lin, M.T., Carmo-silva, A.E., Parry, M.A.J., Building, B., Centre, E., and Kingdom, U. (2017). Towards engineering carboxysomes into C3 plants. Plant J. *87*, 38–50.

Haque, M., Pramanik, H.R., Biswas, J.K., Iftekharuddaula, K.M., and Hasanuzzaman, M. (2015). Comparative performance of hybrid and elite inbred rice varieties with respect to their source-sink relationship. Sci. World *2015*, e326802.

Hay, A., and Tsiantis, M. (2016). Cardamine hirsuta: A comparative view. Curr. Opin. Genet. Dev. *39*, 1–7.

Heyman, H.M., and Dubery, I.A. (2016). The potential of mass spectrometry imaging in plant metabolomics: a review. Phytochem. Rev. *15*, 297–316.

Hibino, H., Cabauatan, P.Q., Omura, T., and Tsuchizaki, T. (1985). Rice grassy stunt virus strain causing Tungrolike symptoms in the Philippines. Plant Dis. *69*, 538–541.

Ho, M.-Y., Shen, G., Canniffe, D.P., Zhao, C., and Bryant, D.A. (2016). Light-dependent chlorophyll f synthase is a highly divergent paralog of PsbA of photosystem II. Science (80-.). *353*, aaf9178-3.

Hofmann, N. R. (2016). A breakthrough in monocot transformation methods. Plant Cell. *30*. tpc.00696.2016.

Horgan, R.P., and Kenny, L.C. (2011). SAC review ‘ Omic ’ technologies: genomics, transcriptomics, proteomics and metabolomics. Obstet. Gynaecol. *13*, 189–195.

Humphries, E.C., and Wheeler, A.W. (1963). The physiology of leaf growth. Annu. Rev. Plant Physiol. *14*, 385–410.

Ishii, T., Brar, D.S., and Multani, D.S. (1994). Molecular tagging of genes for brown planthopper resistance and earliness introgressed from Oryza australiensis into cultivated rice , 0 . sativa. Genome *37*, 217–221.

IRRI Rice Genebank. https://www.irri.org/international-rice-genebank Accessed 26/05/2020

Itoh, J., Nonomura, K., Ikeda, K., Yamaki, S., Inukai, Y., Kitano, H., and Nagato, Y. (2005). Rice plant development : from zygote to spikelet. Plant Cell Physiol. *46*, 23–47.

Javelle, M., Vernoud, V., Rogowsky, P.M., and Ingram, G.C. (2011). Epidermis: the formation and functions of a fundamental plant tissue. New Phytol. *189*, 17–39.

Jin, H., Li, M., Duan, S., Fu, M., Dong, X., Liu, B., Feng, D., Wang, J., and Wang, H. (2016). Optimisation of light-harvesting pigment improves photosynthetic efficiency. Plant Physiol. *172*, 1720–1731.

Jönsson, H., Heisler, M.G., Shapiro, B.E., Meyerowitz, E.M., and Mjolsness, E. (2006). An auxin-driven polarized transport model for phyllotaxis. Proc. Natl. Acad. Sci. U. S. A. *103*, 1633–1638.

Kalve, S., De Vos, D., and Beemster, G.T.S. (2014). Leaf development: a cellular perspective. Front. Plant Sci. *5*, 362.

Kamiya, N., Itoh, J.I., Morikami, A., Nagato, Y., and Matsuoka, M. (2003). The SCARECROW gene's role in asymmetric cell divisions in rice plants. Plant J. *36*, 45–54.

Kamthan, A., Chaudhuri, A., Kamthan, M., and Datta, A. (2016). Genetically modified (GM) crops: milestones and new advances in crop improvement. Theor. Appl. Genet. *129*, 1639–1655.

Karki, S., Rizal, G., and Quick, W.P. (2013). Improvement of photosynthesis in rice (Oryza sativa L.) by inserting the C4 pathway. Rice *6*, 1.

Kawakatsu, T., Itoh, J., Miyoshi, K., Kurata, N., Alvarez, N., Veit, B., Nagato, Y., Kawakatsu, T., Miyoshi, K., Kurata, N., et al. (2006). PLASTOCHRON2 Regulates Leaf Initiation and Maturation in Rice. Plant Cell *18*, 612–625.

Khush, G.S. (1997). Origin, dispersal, cultivation and variation of rice. Plant Mol. Biol. *35*, 25–34.

Khush, G.S., and Ling, K.C. (1974). Inheritance of Resistance to Grassy Stunt Virus and Its Vector in Rice. J. Hered. *65*, 135–136.

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B.A., Thiessen, P.A., Yu, B., et al. (2019). PubChem 2019 update: Improved access to chemical data. Nucleic Acids Res. *47*, D1102–D1109.

Kosmides, A.K., Kamisoglu, K., Calvano, S.E., Corbett, S.A., and Androulakis, I.P. (2013). Metabolomic fingerprinting: Challenges and opportunities. Crit. Rev. Biomed. Eng. *41*, 205–221.

Kusano, M., Yang, Z., Okazaki, Y., Nakabayashi, R., Fukushima, A., and Saito, K. (2015). Using metabolomic approaches to explore chemical diversity in rice. Mol. Plant *8*, 58–67.

Kusumi, K., Chono, Y., Shimada, H., Gotoh, E., Tsuyama, M., and Iba, K. (2010). Chloroplast biogenesis during the early stage of leaf development in rice. Plant Biotechnol. *27*, 85–90.

Kusumi, K., Hirotsuka, S., Kumamaru, T., and Iba, K. (2012). Increased leaf photosynthesis caused by elevated stomatal conductance in a rice mutant deficient in SLAC1, a guard cell anion channel protein. J. Exp. Bot. *63*, 5635–5644.

Langdale, J. Development of Kranz Anatomy. Accessed at https://langdalelab.com/research-2/kranz-anatomy/ on 25/08/19

Langdale, J.A. (2011). C4 cycles : past, present, and future research on C4 photosynthesis. Plant Cell *23*, 3879–3892.

Laux, T., Mayer, K.F., Berger, J., and Jürgens, G. (1996). The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. Development *122*, 87–96.

Lawas, L.M.F., Li, X., Erban, A., Kopka, J., Jagadish, S.V.K., Zuther, E., and Hincha, D.K. (2019). Metabolic responses of rice cultivars with different tolerance to combined drought and heat stress under field conditions. Gigascience *8*, 1–21.

Leegood, R.C. (2008). Roles of the bundle sheath cells in leaves of C 3 plants. J. Exp. Bot. *59*, 1663–1673.

Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S.L., Kebrom, T.H., Provart, N., Patel, R., Myers, C.R., et al. (2010). The developmental dynamics of the maize leaf transcriptome. Nat. Genet. *42*, 1060–1067.

Li, S., Yang, D., and Zhu, Y. (2007). Characterization and use of male sterility in hybrid rice breeding. J. Inte *49*, 791–804.

Lin, Z., Zhang, X., Wang, Z., Jiang, Y., Liu, Z., Alexander, D., Li, G., Wang, S., and Ding, Y. (2017a). Metabolomic analysis of pathways related to rice grain chalkiness by a notched-belly mutant with high occurrence of white-belly grains. BMC Plant Biol. *17*, 1–15.

Lin, Z., Yan, J., Yan, H., and Wang, F. (2017b). Characterization of a strong green tissue-specific motif in rice photosystem I gene promoter Ppsak. Plant Biotechnol. Rep. *11*, 87–95.

Liu, J. and Last, R. L. (2015). MPH1 is a thylakoid membrane protein involved in protecting photosystem II from damage in land plants. Plant Sig. Behav. *10*, e1076602.

Long, S.P., and Bernacchi, C.J. (2003). Gas exchange measurements, what can they tell us about the underlying limitations to photosynthesis? Procedures and sources of error. J. Exp. Bot. *54*, 2393–2401.

Lundgren, M.R., Osborne, C.P., and Christin, P. (2014). Deconstructing Kranz anatomy to understand C4 evolution. J. Exp. Bot. *65*, 3357–3369.

Luo, L., Zhou, W., Liu, P., Li, C., and Hou, S. (2012). The development of stomata and other epidermal cells on the rice leaves. Biol. Plant. *56*, 521–527.

Lysenko, V., Guo, Y., and Chugueva, O. (2016). Cyclic electron transport around photosystem II: mechnaisms and methods of study. Am. J. Plant Physiol. *12*, 1–9.

Maes, L., and Goossens, A. (2010). Hormone-mediated promotion of trichome initiation in plants is conserved but utilizes species- and trichome-specific regulatory mechanisms. Plant Signal. Behav. *5*, 205–207.

Majeran, W., Friso, G., Ponnala, L., Connolly, B., Huang, M., Reidel, E., Zhang, C., Asakura, Y., Bhuiyan, N.H., Sun, Q., et al. (2010). Structural and metabolic transitions of C4 leaf development and differentiation defined by microscopy and quantitative proteomics in maize. Plant Cell *22*, 3509–3542.

Mann, D.G.J., Lafayette, P.R., Abercrombie, L.L., King, Z.R., Mazarei, M., Halter, M.C., Poovaiah, C.R., Baxter, H., Shen, H., Dixon, R.A., et al. (2012). Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (Panicum virgatum L.) and other monocot species. Plant Biotechnol. J. *10*, 226–236.

Martinez, C.C., Chitwood, D.H., Smith, R.S., and Sinha, N.R. (2016). Left - Right leaf asymmetry in decussate and distichous phyllotactic systems. Philos. Trans. R. Soc. B Biol. Sci. *371*.

Matsuoka, M., Kyozuka, J., Shimamoto, K., and Kano-Murakami, Y. (1994). The promoters of two carboxylases in a C4 plant (maize) direct cell-specific, light-regulated expression in a C3 plant (rice). Plant J. *6*, 311–319.

Melorose, J., Perroy, R., and Careas, S. (2015). World population prospects. United Nations *1*, 587–592.

Menguer, P.K., Sperotto, R.A., and Ricachenevsky, F.K. (2017). A walk on the wild side : Oryza species as source for rice abiotic stress tolerance. Genet. Mol. Biol. *40*, 238–252.

Mimura, M., Nagato, Y., and Itoh, J.I. (2012). Rice PLASTOCHRON genes regulate leaf maturation downstream of the gibberellin signal transduction pathway. Planta *235*, 1081–1089.

Ming, R., VanBuren, R., Wai, C.M., Tang, H., Schatz, M.C., Bowers, J.E., Lyons, E., Wang, M.-L., Chen, J., Biggers, E., et al. (2015). The pineapple genome and the evolution of CAM photosynthesis. Nat. Genet. *47*, 1435–1442.

Miyoshi, K., Ahn, B.-O., Kawakatsu, T., Ito, Y., Itoh, J.-I., Nagato, Y., and Kurata, N. (2004). PLASTOCHRON1, a timekeeper of leaf initiation in rice, encodes cytochrome P450. Proc. Natl. Acad. Sci. *101*, 875–880.

Moualeu-Ngangue, D.P., Chen, T.W., and Stützel, H. (2017). A new method to estimate photosynthetic parameters through net assimilation rate−intercellular space CO2 concentration (A−Ci) curve and chlorophyll fluorescence measurements. New Phytol. *213*, 1543–1554.

Mumm, R., Hageman, J.A., Calingacion, M.N., de Vos, R.C.H., Jonker, H.H., Erban, A., Kopka, J., Hansen, T.H., Laursen, K.H., Schjoerring, J.K., et al. (2016). Multi-platform metabolomics analyses of a broad collection of fragrant and non-fragrant rice varieties reveals the high complexity of grain quality characteristics. Metabolomics *12*, 1–19.

Murchie, E.H., and Niyogi, K.K. (2011). Manipulation of photoprotection to improve plant photosynthesis. Plant Physiol. *155*, 86–92.

Murchie, E.H., Hubbart, S., Peng, S., Horton, P., and Sheffield, S. (2005). Acclimation of photosynthesis to high irradiance in rice : gene expression and interactions with leaf development. J. Biochem. Mol. Toxicol. *56*, 449–460.

Narawatthana, S. (2013). The regulation of leaf thickness in rice (Oryza sativa L.). Thesis.

Nelissen, H., Gonzalez, N., and Inzé, D. (2016). Leaf growth in dicots and monocots: So different yet so alike. Curr. Opin. Plant Biol. *33*, 72–76.

Nelson, T., and Langdale, J.A. (1989). Patterns of leaf development in C4 plants. Plant Cell *1*, 3–13.

Nunes, T.D.G., Zhang, D., and Raissig, M.T. (2020). Form, development and function of grass stomata. Plant J. *101*, 780–799.

Ohtsu, K., Smith, M.B., Emrich, S.J., Borsuk, L.A., Zhou, R., Chen, T., Zhang, X., Timmermans, M.C.P., Beck, J., Buckner, B., et al. (2007). Global gene expression analysis of the shoot apical meristem of maize (Zea mays L.). Plant J. *52*, 391–404.

Oi, T., Enomoto, S., Nakao, T., Arai, S., Yamane, K., and Taniguchi, M. (2017). Three-dimensional intracellular structure of a whole rice mesophyll cell observed with FIB-SEM. Ann. Bot. *120*, 21–28.

Okazaki, Y., and Saito, K. (2016). Integrated metabolomics and phytochemical genomics approaches for studies on rice. Gigascience *5*, 1–7.

Ort, D.R., Zhu, X., and Melis, A. (2011). Optimising antenna size to maximise photosynthetic efficiency. Plant Physiol. *155*, 79–85.

Ouk, R., Oi, T., and Taniguchi, M. (2019). Three-dimensional anatomy of mesophyll cells in rice leaf tissue by serial section light microscopy. Plant Prod. Sci. *00*, 1–11.

Overy, S.A., Walker, H.J., Malone, S., Howard, T.P., Baxter, C.J., Sweetlove, L.J., Hill, S.A., Quick, W.P., Sciences, P., and Road, S.P. (2005). Application of metabolite profiling to the identification of traits in a population of tomato introgression lines. J. Exp. Bot. *56*, 287–296.

Padilla-González, G.F., Frey, M., Gómez-Zeledón, J., Da Costa, F.B., and Spring, O. (2019). Metabolomic and gene expression approaches reveal the developmental and environmental regulation of the secondary metabolism of yacón (Smallanthus sonchifolius, Asteraceae). Sci. Rep. *9*, 1–15.

Parker, D., Beckmann, M., Zubair, H., Enot, D.P., Caracuel-Rios, Z., Overy, D.P., Snowdon, S., Talbot, N.J., and Draper, J. (2009). Metabolomic analysis reveals a common pattern of metabolic re-programming during invasion of three host plant species by Magnaporthe grisea. Plant J. *59*, 723–737.

Patrick, L.D., Ogle, K., and Tissue, D.T. (2009). A hierarchical Bayesian approach for estimation of photosynthetic parameters of C3 plants. Plant, Cell Environ. *32*, 1695–1709.

Peterhansel, C. (2011). Best practice procedures for the establishment of a C4 cycle in transgenic C3 plants. J. Exp. Bot. *62*, 3011–3019.

Peterhansel, C., Horst, I., Niessen, M., Blume, C., Kebeish, R., Kürkcüoglu, S., and Kreuzaler, F. (2010). Photorespiration. Arab. B. *8*, e0130.

Pick, T.R., Brautigam, A., Schluter, U., Denton, A.K., Colmsee, C., Scholz, U., Fahnenstich, H., Pieruschka, R., Rascher, U., Sonnewald, U., et al. (2011). Systems Analysis of a Maize Leaf Developmental Gradient Redefines the Current C4 Model and Provides Candidates for Regulation. Plant Cell *23*, 4208–4220.

Poethig, R.S. (1997). Leaf morphogenesis in flowering plants. Plant Cell *9*, 1077–1087.

Prado, J.R., Segers, G., Voelker, T., Carson, D., Dobert, R., Phillips, J., Cook, K., Cornejo, C., Monken, J., Grapes, L., et al. (2014). Genetically engineered crops: from idea to product. Annu. Rev. Plant Biol. *65*, 769–790.

Prasad, B., Eizenga, G.C., and Griff, O. (2008). Rice Sheath Blight Disease Resistance Identified in Oryza spp . Accessions. 1503–1509.

Price, J., Laxmi, A., St. Martin, S.K., and Jang, J.C. (2004). Global transcription profiling reveals multiple sugar signal transduction mechanisms in Arabidopsis. Plant Cell *16*, 2128–2150.

Quan, R., Wang, J., Hui, J., Bai, H., Lyu, X., and Zhu, Y. (2018). Improvement of salt tolerance using wild rice genes. Front. Plant Sci. *8*, 1–11.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ray, D.K., Mueller, N.D., West, P.C., and Foley, J.A. (2013). Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLoS One *8*.

de Reuille, P.B., Routier-Kierzkowska, A.L., Kierzkowski, D., Bassel, G.W., Schüpbach, T., Tauriello, G., Bajpai, N., Strauss, S., Weber, A., Kiss, A., et al. (2015). MorphoGraphX: A platform for quantifying morphogenesis in 4D. Elife *4*, 1–20.

Rice Association. http://www.riceassociation.org.uk/content/1/18/types-of-rice.html Accessed 26/05/2020

Rochfort, S. (2005). Metabolomics reviewed: A new “omics” platform technology for systems biology and implications for natural products research. J. Nat. Prod. *68*, 1813–1820.

Roessner, U., Wagner, C., Kopka, J., Trethewey, R.N., and Willmitzer, L. (2000). Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. Plant J. *23*, 131–142.

Royal Society. What GM crops are currently being grown and where?. (2016). Accessed at https://royalsociety.org/topics-policy/projects/gm-plants/what-gm-crops-are-currently-being-grown-and-where/ on 27/08/19

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

Sage, R.F., Khoshravesh, R., and Sage, T.L. (2014). From proto-Kranz to C4 Kranz : building the bridge to C4 photosynthesis. J. Exp. Bot. *65*, 3341–3356.

Sánchez, E., and Espinoza, A.M. (2005). Ultrastructure of Oryza glumaepatula, a wild rice species endemic of tropical America. Rev. Biol. Trop. *53*, 15–22.

Sánchez, E., Montiel, M., and Espinoza, A.M. (2003). Ultrastructural morphologic description of the wild rice species Oryza latifolia (Poaceae) in Costa Rica . Rev. Biol. Trop. *51*, 345–354.

Sánchez, E., Quesada, T., and Espinoza, A.M. (2006). Ultrastructure of the wild rice Oryza grandiglumis (Gramineae) in Costa Rica . Rev. Biol. Trop. *54*, 377–385.

Satoh, N., Hong, S.K., Nishimura, A., Matsuoka, M., Kitano, H., and Nagato, Y. (1999). Initiation of shoot apical meristem in rice: Characterization of four SHOOTLESS genes. Development *126*, 3629–3636.

Savaldi-Goldstein, S., Peto, C., and Chory, J. (2007). The epidermis both drives and restricts plant shoot growth. Nature *446*, 1–4.

Scarpella, E., Rueb, S., Boot, K.J.M., Hoge, J.H.C., and Meijer, A.H. (2000). A role for the rice homeobox gene Oshox1 in provascular cell fate commitment. Development *127*, 3655–3669.

Scarpella, E., Rueb, S., and Meijer, A.H. (2003). The RADICLELESS1 gene is required for vascular pattern formation in rice. Development *130*, 645–658.

Schiefelbein, J.W. (1994). Cell fate and cell morphogenesis. Curr. Opin. Genet. Dev. *4*, 647–

651.

Schindelin, J., Rueden, C.T., Hiner, M.C., and Eliceiri, K.W. (2015). The ImageJ ecosystem : an open platform for biomedical image analysis. Mol. Reprod. Dev. *529*, 518–529.

Schreier, T.B., and Hibberd, J.M. (2019). Variations in the Calvin-Benson cycle: Selection pressures and optimization? J. Exp. Bot. *70*, 1697–1701.

Schuler, M.L., Sedelnikova, O. V., Walker, B.J., Westhoff, P., and Langdalea, J.A. (2018). SHORTROOT-mediated increase in stomatal density has no impact on photosynthetic efficiency. Plant Physiol. *176*, 757–772.

Seck, P.A., Diagne, A., Mohanty, S., and Wopereis, M.C.S. (2012). Crops that feed the world 7: Rice. Food Secur. *4*, 7–24.

Sharkey, T.D. (2016). What gas exchange data can tell us about photosynthesis. Plant Cell Environ. *39*, 1161–1163.

Sharkey, T.D., Bernacchi, C.J., Farquhar, G.D., and Singsaas, E.L. (2007). Fitting photosynthetic carbon dioxide response curves for C3 leaves. Plant, Cell Environ. *30*, 1035–1040.

Shyh-Chang, N., Daley, G.Q., and Cantley, L.C. (2013). Stem cell metabolism in tissue development and aging. Dev. *140*, 2535–2547.

Silveira-Sotelo, M., Chauvin, A.L., Marsch-Martínez, N., Winkler, R., and de Folter, S. (2015). Metabolic fingerprinting of Arabidopsis thaliana accessions. Front. Plant Sci. *6*, 1–13.

Sitch, L.A., Nelson, R., Dalmaeio, R.D., Oliva, N.P., Aswidinnoor, H., and Leung, H. (1992). Transfer of bacterial blight and blast resistance from the tetraploid wild rice Oryza minuta to cultivated rice , Oryza sativa. Theor. Appl. Genet. *84*, 345–354.

Slewinski, T.L., Anderson, A.A., Price, S., Withee, J.R., Gallagher, K., and Turgeon, R. (2014). Short-root1 plays a role in the development of vascular tissue and Kranz anatomy in maize leaves. Mol. Plant *7*, 1388–1392.

Somssich, M., and Je, B. Il (2016). CLAVATA-WUSCHEL signaling in the shoot meristem. Development *143*, 3238–3248.

Song, Q., Wang, Y., and Ort, D.R. (2017). The impact of modifying photosystem antenna size on canopy photosynthetic efficiency — development of a new canopy photosynthesis model scaling from metabolism to canopy level processes. Plant, Cell Environ. *40*, 2946–2957.

Stebbins, G.L., and Shah, S.S. (1960). Developmental studies of cell differentiation in the epidermis of monocotyledons. II. Cytological features of stomatal development in the Gramineae. Dev. Biol. *2*, 477–500.

Sun, X., Cahill, J., Van Hautegem, T., Feys, K., Whipple, C., Novák, O., Delbare, S., Versteele, C., Demuynck, K., De Block, J., et al. (2017). Altered expression of maize PLASTOCHRON1 enhances biomass and seed yield by extending cell division duration. Nat. Commun. *8*, 14752.

Sweetlove, L.J., Beard, K.F.M., Nunes-Nesi, A., Fernie, A.R., and Ratcliffe, R.G. (2010). Not just a circle: Flux modes in the plant TCA cycle. Trends Plant Sci. *15*, 462–470.

Sylvester, A.W., Smith, L., and Freeling, M. (1996). Acquisition of identity in the developing leaf. Annu. Rev. Cell Dev. Biol. *12*, 257–304.

Sylvester, A.W., Parker-Clark, V., and Murray, G.A. (2001). Leaf shape and anatomy as indicators of phase change in the grasses: Comparison of maize, rice, and bluegrass. Am. J. Bot. *88*, 2157–2167.

Tagliabue, G. (2017). The EU legislation on " GMOs " between nonsense and protectionism : An ongoing Schumpeterian chain of public choices. GM Crop. Food *8*, 57–73.

Tanaka, N., Itoh, H., Sentoku, N., Kojima, M., Sakakibara, H., Izawa, T., Itoh, J.-I., and Nagato, Y. (2011). The COP1 ortholog PPS regulates the juvenile-adult and vegetative-reproductive phase changes in rice. Plant Cell *23*, 2143–2154.

Teng, N., Wang, J., Chen, T., Wu, X., Wang, Y., Teng, N., Wang, J., Chen, T., Wu, X., Wang, Y., et al. (2006). Elevated CO2 induces physiological , biochemical and structural changes in leaves of Arabidopsis thaliana. New Phytol. *172*, 92–103.

Terashima, I., Hanba, Y.T., Tazoe, Y., Vyas, P., and Yano, S. (2006). Irradiance and phenotype: Comparative eco-development of sun and shade leaves in relation to photosynthetic CO2 diffusion. J. Exp. Bot. *57*, 343–354.

Terashima, I., Hanba, Y.T., Tholen, D., and Niinemets, Ü. (2011). Leaf functional anatomy in relation to photosynthesis. Plant Physiol. *155*, 108–116.

Thuy An, P.N., Yamaguchi, M., Bamba, T., and Fukusaki, E. (2014). Metabolome analysis of Drosophila melanogaster during embryogenesis. PLoS One *9*.

Tolbert, N.E. (1997). The C2 oxidative photosynthetic carbon cycle. Annu. Rev. Plant Physiol. Plant Mol. Biol. *48*, 1–25.

Tsukaya, H. (2005). Leaf shape : genetic controls and environmental factors. Int. J. Dev. Biol. *555*, 547–555.

Walker, B.J., VanLoocke, A., Bernacchi, C.J., and Ort, D.R. (2016). The Costs of Photorespiration to Food Production Now and in the Future. Annu. Rev. Plant Biol. *67*, annurev-arplant-043015-111709.

Walker, B.J., Drewry, D.T., Slattery, R.A., Vanloocke, A., and Cho, Y.B. (2018). Chlorophyll can be reduced in crop canopies with little penalty to photosynthesis. Plant Physiol. *176*, 1215–1232.

Walker, H. (2013). Direct approaches to metabolite profling by Mass Spectrometry. Thesis.

Wang, L., Czedik-Eysenberg, A., Mertz, R.A., Si, Y., Tohge, T., Nunes-Nesi, A., Arrivault, S., Dedow, L.K., Bryant, D.W., Zhou, W., et al. (2014). Comparative analyses of C_4 and C_3 photosynthesis in developing leaves of maize and rice. Nat. Biotechnol. *32*, 1158–1165.

Wang, P., Kelly, S., Fouracre, J.P., and Langdale, J.A. (2013). Genome-wide transcript analysis of early maize leaf development reveals gene cohorts associated with the differentiation of C4 Kranz anatomy. Plant J. *75*, 656–670.

Wang, P., Karki, S., Biswal, A.K., Lin, H.C., Dionora, M.J., Rizal, G., Yin, X., Schuler, M.L., Hughes, T., Fouracre, J.P., et al. (2017a). Candidate regulators of Early Leaf Development in Maize Perturb Hormone Signalling and Secondary Cell Wall Formation When Constitutively Expressed in Rice. Sci. Rep. *7*, 1–15.

Wang, P., Khoshravesh, R., Karki, S., Tapia, R., Balahadia, C.P., Bandyopadhyay, A., Quick, W.P., Furbank, R., Sage, T.L., and Langdale, J.A. (2017b). Re-creation of a key step in the evolutionary switch from C3 to C4 leaf anatomy. Curr. Biol. *27*, 3278-3287.e6.

Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. PLoS One *6*.

Wei, J., Wang, A., Li, R., Qu, H., and Jia, Z. (2018). Metabolome-wide association studies for agronomic traits of rice. Heredity (Edinb). *120*, 342–355.

Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Wickham, H. and Henry, L. (2020). tidyr: Tidy Messy Data. R package version 1.0.2. https://CRAN.R-project.org/package=tidyr

Wickham, H., François, R., Henry, L. and Müller, K. (2020). dplyr: A Grammar of Data Manipulation. R package version 0.8.4. https://CRAN.R-project.org/package=dplyr

Wu, Z., Chen, L., Yu, Q., Zhou, W., Gou, X., Li, J., and Hou, S. (2019). Multiple transcriptional factors control stomata development in rice. New Phytol. *223*, 220–232.

Wuyts, N., Palauqui, J.-C., Conejero, G., Verdeil, J.-L., Granier, C., and Massonnet, C. (2010). High-contrast three-dimensional imaging of the Arabidopsis leaf enables the analysis of cell dimensions in the epidermis and mesophyll. Plant Methods *6*, 1–14.

Xiao, Y., Hepworth, C., Mathers, A., Sloan, J., Thorley, R., Baille, A., Jones, H., Chang, T., Osborne, C. P., Chen, X., Sturrock, C., Mooney, S., Fleming, A. J. and Zhu, X. eLeaf, An integrated spatial model of leaf photosynthetic physiology, identifies traits for improved rice leaf performance in elevated CO2. Manuscript in preparation.

Xie, K., Shen, J., Hou, X., Yao, J., Li, X., Xiao, J., and Xiong, L. (2012). Gradual Increase of miR156 Regulates Temporal Expression Changes of Numerous Genes during Leaf Development in Rice 1[C][W][OA]. Plant Physiol. *158*, 1382–1394.

Xu, F., Wang, K., Yuan, W., Xu, W., Liu, S., Kronzucker, H.J., Chen, G., Miao, R., Zhang, M., Ding, M., et al. (2019). Overexpression of rice aquaporin OsPIP1;2 improves yield by enhancing mesophyll CO2 conductance and phloem sucrose transport. J. Exp. Bot. *70*, 671–681.

Yamamoto, H., Takahashi, S., Badger, M.R., and Shikanai, T. (2016). Artificial remodelling of alternative electron flow by flavodiiron proteins in Arabidopsis. Nat. Plants *2*, 16012.

Yamamoto, Y.Y., Ichida, H., Matsui, M., Obokata, J., Sakurai, T., Satou, M., Seki, M., Shinozaki, K., and Abe, T. (2007). Identification of plant promoter constituents by analysis of local distribution of short sequences. BMC Genomics *8*, 67.

Yamanaka, S., Takeda, H., Komatsubara, S., and Ito, F. (2009). Structures and physiological functions of silica bodies in the epidermis of rice plants . Appl. Phys. Lett. *123703*, 1–4.

Yang, J., Yuan, Z., Meng, Q., Huang, G., and Périn, C. (2017). Dynamic Regulation of Auxin Response during Rice Development Revealed by Newly Established Hormone Biosensor Markers. Front. Plant Sci. *8*, 1–17.

Yasui, Y., Ohmori, Y., Takebayashi, Y., Sakakibara, H., and Hirano, H.Y. (2018). WUSCHEL-RELATED HOMEOBOX4 acts as a key regulator in early leaf development in rice. PLoS Genet. *14*, e1007365.

Yin, X., Karki, S., Francisco, A., Wanchana, S., Lin, H., Acoba, M., Tapia, R., Libiternos, R., ver Sagun, J., Coe, R., de Luna, A., Chadha-Mohanty, P., Slamet-Loedin, I., Bandyopadhyay, A. and Quick, W. P. “Safe harbours” for targeted insertion of foreign genes in rice identified from transgenic rice plants that express maize C4 photosynthetic genes. Poster proceedings.

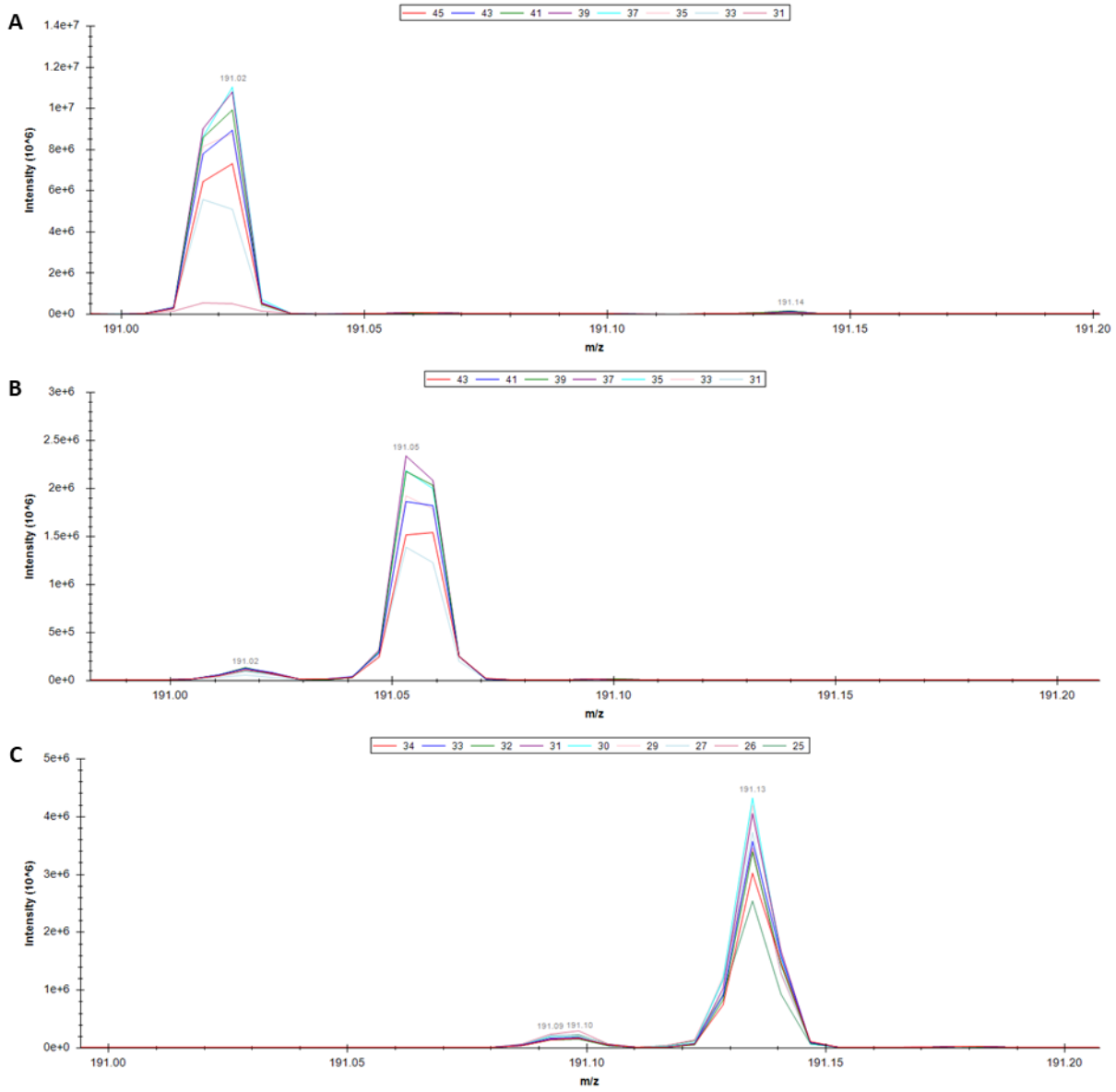
Zarzycki, J., Axen, S.D., Kinney, N., Kerfeld, A., and Rende, I.-A. (2013). Cyanobacterial-based approaches to improving photosynthesis in plants. J. Exp. Bot. *64*, 787–798.

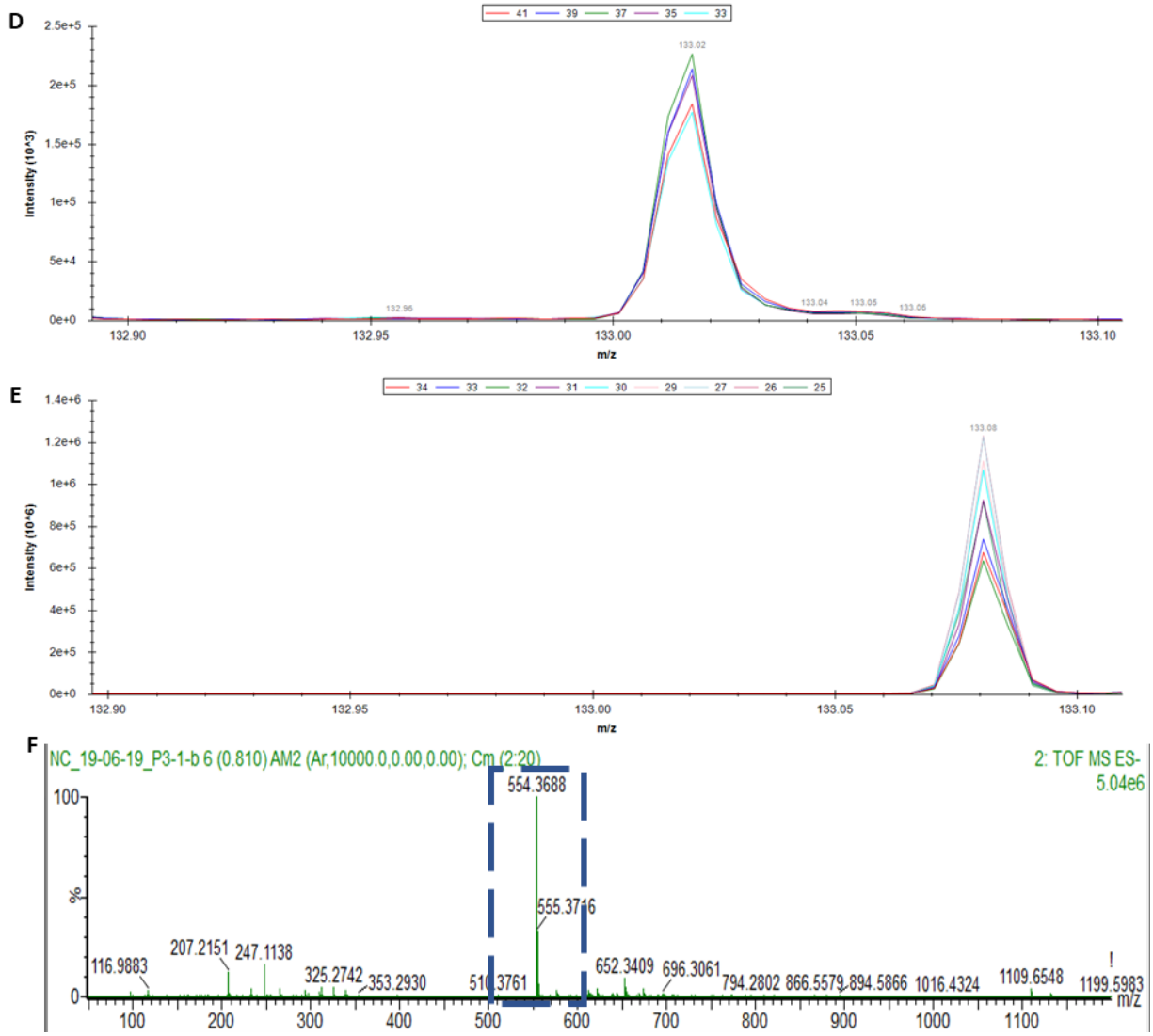
Zhang, F., and Xie, J. (2016). Genes and QTLs Resistant to Biotic and Abiotic Stresses from Wild Rice and Their Applications in Cultivar Improvements. In Rice - Germplasm, Genetics and Improvement, pp. 59–78.

Zhu, X.-G., Long, S.P., and Ort, D.R. (2010). Improving photosynthetic efficiency for greater yield. Annu. Rev. Plant Biol. *61*, 235–261.

Zhu, X.G., Long, S.P., and Ort, D.R. (2008). What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? Curr. Opin. Biotechnol. *19*, 153–159.

S3.1 Lockmass adjustment of spectra





Supplementary Figure S3.1 – Lockmass adjustment of spectra

A – MS spectrum peak showing citrate standard

B – MS spectrum peak showing confirmed citrate in P5 extract, sample run alongside A

C – MS spectrum peak showing putative citrate in P5 extract, sample run in separate experiment to A and B

D – MS spectrum peak showing confirmed malate in P5 extract, sample run alongside A and B

E – MS spectrum peak showing putative malate in P5 extract, sample run alongside C

F – MS spectrum showing lock mass m/z value (highlighted at 554.3688) for random sample run alongside C and E

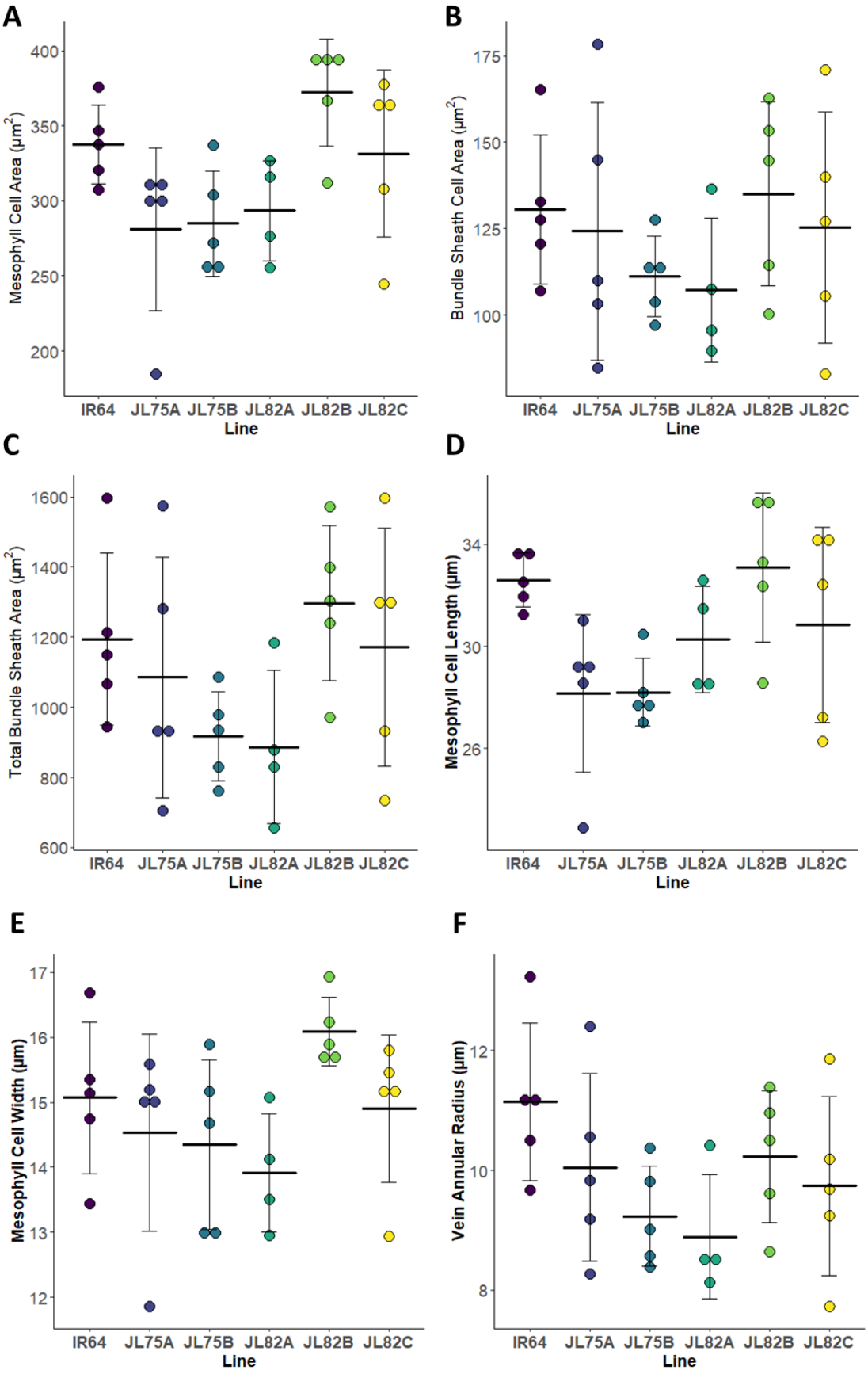
Appendix 3.1 – Transcript information for Figure 3.12

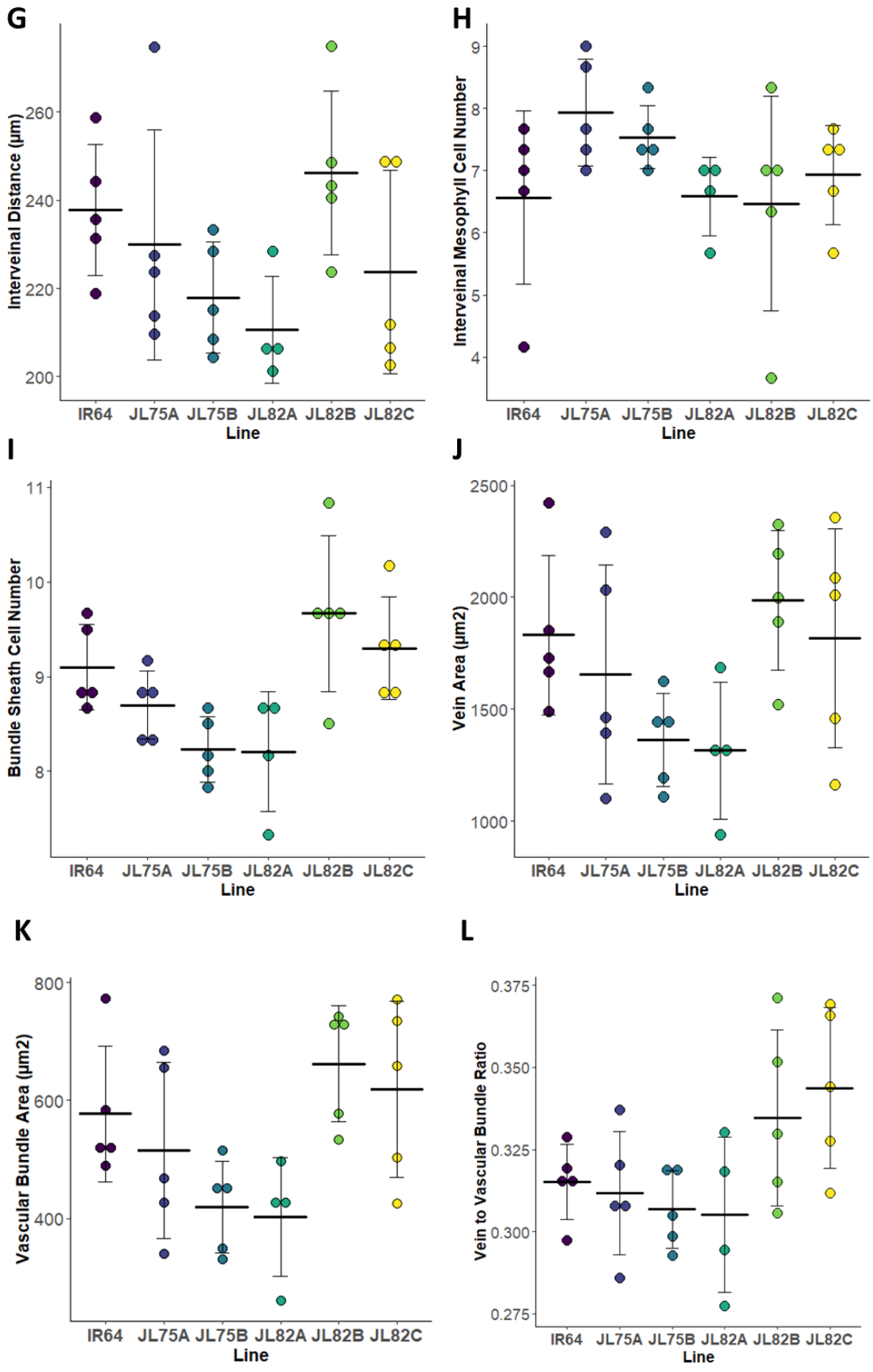
Gene	Protein	Pathway	MSU Gene ID	RAP Gene ID	Experimental Evidence (Uniprot)	Review Status (Uniprot)	Leaf Gradient Expression (Rice EFP Browser 'rice leaf gradient')	Expression (RNAseq)
	Citrate Synthase	TCA	LOC_Os11g33240	Os11g0538900	Transcript level	Computational	negligible	Down 2
	Aconitase	TCA	LOC_Os08g09200	Os08g0191100	Inferred from homology	Manual	higher at base	Down 2
	Aconitase	TCA	LOC_Os03g04410	Os03g0136900	Inferred from homology	Computational	higher at base	Down 2
	Aconitase	TCA	LOC_Os06g19960	Os06g0303400	Inferred from homology	Computational	negligible	No Expression
	Isocitrate Dehydrogenase	TCA	LOC_Os04g42920	Os04g0508200	Transcript Level	Computational	higher at base	Peak
	Isocitrate Dehydrogenase	TCA	LOC_Os01g46610	Os01g0654500	Transcript Level	Computational	higher at base	Down 2
	Isocitrate Dehydrogenase	TCA	LOC_Os05g49760	Os05g0573200	Inferred from homology	Computational	negligible	negligible
	Succinate--CoA Ligase α	TCA	LOC_Os07g38970	Os07g0577700	Transcript level	Manual	higher at base	Down 3

	Succinate--CoA Ligase β	TCA	LOC_Os02g40830	Os02g0621700	Transcript level	Manual	higher at base	Down 3
SDH 1	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial	TCA	LOC_Os07g04240	Os07g0134800	Protein Level	Manual	higher at base	Down 3
SDH 2-1	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit 1, mitochondrial	TCA	LOC_Os08g02640	Os08g0120000	Protein Level	Manual	high at base and tip	Down 3
SDH 2-2	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit 2, mitochondrial	TCA	LOC_Os09g20440	Os09g0370300	Protein Level	Manual	No expression	No Expression
SDH 3-1	Succinate dehydrogenase subunit 3-1 (mitochondrial)	TCA	LOC_Os02g02940	Os02g0121800	Protein Level	Manual	higher at base	Down 1
SDH 3-2	Succinate dehydrogenase subunit 3-2 (mitochondrial)	TCA	LOC_Os07g33680	Os07g0521000	Protein Level	Manual	Higher at base	Down 2
SDH 4	Succinate dehydrogenase	TCA	LOC_Os01g70980	Os01g0935900	Protein Level	Manual	higher at base	Down 2

	subunit 4 (mitochondrial)							
SDH 5	Succinate dehydrogenase subunit 5 (mitochondrial)	TCA	LOC_Os04g34100	Os04g04180 00	Protein Level	Manual	higher at base	Down 1
SDH 6	Succinate Dehydronase Subunit 6 (Mitochondrial)	TCA	LOC_Os08g02080	Os08g01128 00	Protein Level	Manual	higher at base	Down 3
SDH 7	Succinate Dehydronase Subunit 7 (Mitochondrial)	TCA	LOC_Os09g21470	Os09g03825 00	Protein Level	Manual	negligable	Down 1
SDH 8A	Succinate Dehydronase Subunit 8A (Mitochondrial)	TCA	LOC_Os06g03486	Os06g01249 00	Inferred from homology	Manual	Higher at base	Peak
SDH 8B	Succinate Dehydronase Subunit 8B (Mitochondrial)	TCA	LOC_Os06g03514	Os06g01251 32	Inferred from homology	Manual	higher at base	negligible
	Fumarase	TCA	LOC_Os03g21950	Os03g03379 00	Transcript level	Computational	higher at base	Down 3
	Malate Dehydrogenase (Cytoplasmic)	TCA	LOC_Os10g33800	Os10g04782 00	Protein Level	Manual	higher at base	Down 2

S4.1 Histology data for leaf seven





Supplementary Figure S4.1 – Histology data for leaf seven
For every graph, each point represents an individual plant. Horizontal bars show the mean value for the

genotype, and error bars show standard deviation.
The WT Line is IR64.
A – Average mesophyll cell area by genotype.
B – Average bundle sheath cell area by genotype.
C – Total bundle sheath cell area by genotype.
D – Average mesophyll cell length by genotype.
E – Average mesophyll cell width by genotype.
F – Vein annular radius by genotype.
G – Interveinal distance by genotype.
H – Average interveinal mesophyll cell number by genotype.
I – Average bundle sheath cell number by genotype.
J – Vein area by genotype.
K – Vascular bundle area by genotype.
L – Vein to vascular bundle ratio.

Appendix 4.1 – Lines used in Chapter 4

4.1.1 - Data for lines used:

Provided by Xiaojia Yin, published in Wang *et al.*, 2017. Reproduced from supplementary data available online, under a Creative Commons licence.

	4th leaf			Expanded 6th seedling leaf									
Line	Genotype (PCR+, WT, TC)	Copy number	Transgene expressed (high/ medium/ low/no)	Leaf Width (mm)	Leaf length (mm)	Total # veins	Distance between veins (um, mean)	Distance between veins (sd)	Cell # between veins (mean)	Cell # between veins (sd)	BS cell size (mean)	BS cell size (sd)	Veins per mm
JL75													
IR64-IRS948-019-07 "JL75A"	PCR+	3	high	5.5	247.0	31.0	183.3	18.2	5.8	1.3	11.2	2.8	5.6
IR64-IRS948-023-01 "JL75B"	PCR+	2	high	6.0	237.0	30.0	194.8	20.3	5.8	1.4	11.0	2.7	5.0
Wildtype Average (n=9)	WT	NA	NA	5.6	234.7	29.1	194.0	21.5	6.4	1.1	10.5	2.8	5.2
JL82													
IR64-IRS940-004-05 "JL82A"	PCR+	1	NA	7.0	288.0	28.0	194.7	15.8	8.0	1.2	10.4	2.4	4.0
IR64-IRS940-010-02 "JL82B"	PCR+	1	NA	7.0	329.0	30.0	218.0	16.0	7.9	0.7	9.0	2.2	4.3
IR64-IRS940-017-07	PCR+	1	NA	6.5	263.0	35.0	186.7	11.7	6.2	0.8	10.1	2.5	5.4

“JL82C”													
Wildtype Average (n=9)	WT	NA	NA	7.7	242.6	34.2	207.6	17.4	7.3	1.1	11.2	2.9	4.5

Appendix 4.2 – ANOVAs and significance tables for Section 4.3.2

Cells highlighted yellow indicate significant differences. Cells highlighted orange indicate a non-significant ANOVA.

4.2.1 ANOVA and Significance Table for Interveinal Distance

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	7398.242	5	1479.648	6.521888	8.55E-05	2.389444
Within Groups	12024.33	53	226.8742			
Total	19422.58	58				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.925789	0.004918	0.981976	0.198961	0.894579
JL75B			0.000826	0.894046	0.060509	0.743789
JL82A				0.001658	9E-06	0.00014
JL82B					0.14064	0.901141
JL82C						0.028919
WT						

4.2.2 ANOVA and Significance Table for Vascular Bundle Width

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	49.80124	5	9.960247	4.653492	0.001355	2.389444
Within Groups	113.4402	53	2.140381			
Total	163.2414	58				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.127168	0.03473	0.461502	0.561043	0.574886
JL75B			0.289932	0.011093	0.011466	0.006131
JL82A				0.004404	0.005161	0.004501
JL82B					0.819611	0.738472
JL82C						0.929754
WT						

4.2.3 ANOVA and Significance Table for Total Vein Width

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	255.4072	5	51.08143	3.743879	0.005627	2.389444
Within Groups	723.1312	53	13.64398			
Total	978.5383	58				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.407823	0.597405	0.148941	0.139562	0.056413
JL75B			0.780139	0.013001	0.009688	0.008499
JL82A				0.043355	0.038301	0.019387
JL82B					0.887846	0.332486
JL82C						0.263864
WT						

4.2.4 ANOVA for Vascular Bundle Width : Total Vein Width

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.360669	5	0.072134	1.275892	0.28809	2.389444
Within Groups	2.99641	53	0.056536			
Total	3.35708	58				

4.2.5 ANOVA for Mesophyll Cell Number

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	4.806257	5	0.961251	2.108554	0.078695	2.389444
Within Groups	24.16173	53	0.455882			
Total	28.96798	58				

4.2.6 ANOVA and Significance Table for Mesophyll Cell Length

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	229.4785	5	45.8957	7.93328	1.23E-05	2.389444
Within Groups	306.6162	53	5.785211			
Total	536.0946	58				

Post-Hoc Testing

Adjusted P-value = 0.003333

JL75A	JL75B	JL82A	JL82B	JL82C	WT

JL75A	0.21994	0.052554	0.004492	0.000187	0.43929
JL75B		0.019725	0.290569	0.03323	0.576029
JL82A			0.000212	1.29E-05	0.02853
JL82B				0.139538	0.061637
JL82C					0.003241
WT					

4.2.7 ANOVA and Significance Table for Bundle Sheath Cell Number

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2.717682	5	0.543536	2.63423	0.033619	2.389444
Within Groups	10.9358	53	0.206336			
Total	13.65348	58				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.333442	0.306852	0.582987	0.061566	0.208217
JL75B			0.017005	0.637701	0.26488	0.740428
JL82A				0.061621	0.000941	0.005472
JL82B					0.125254	0.421952
JL82C						0.40768
WT						

4.2.8 ANOVA and Significance Table for Bundle Sheath Cell Width

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*

Between Groups	11.60426	5	2.320852	3.312428	0.011222	2.389444
Within Groups	37.13444	53	0.70065			
Total	48.7387	58				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.040209	0.569165	0.020528	0.519451	0.754037
JL75B			0.215951	0.000491	0.064271	0.076466
JL82A				0.029077	0.848725	0.483306
JL82B					0.000861	0.236197
JL82C						0.484034
WT						

4.2.9 ANOVA and Significance Table for Bundle Sheath Cell Width : Mesophyll Cell Width Ratio

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2.226074	5	0.445215	6.434452	9.68E-05	2.389444
Within Groups	3.667194	53	0.069192			
Total	5.893268	58				

Post-Hoc Testing

Adjusted P-value = 0.003333

JL75A	JL75B	JL82A	JL82B	JL82C	WT

JL75A		0.006524	0.453289	0.30089	0.000549	0.681667
JL75B			0.005663	0.061755	0.335043	0.027034
JL82A				0.143841	0.000771	0.321351
JL82B					0.006709	0.598412
JL82C						0.002932
WT						

Appendix 4.3 – ANOVAs and significance tables for Section 4.3.3

4.3.1 ANOVA for Annular Radius

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	14.89519	5	2.979037	1.871605	0.138626	2.639999
Within Groups	36.60915	23	1.591702			
Total	51.50434	28				

4.3.2 ANOVA and Significance Table for Vascular Bundle Area

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	263006.4	5	52601.28	3.743232	0.012614	2.639999
Within Groups	323204.5	23	14052.37			
Total	586210.9	28				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.251478	0.220848	0.108812	0.30434	0.482948
JL75B			0.790513	0.002825	0.038015	0.038297
JL82A				0.006848	0.036527	0.045625
JL82B					0.602685	0.244996
JL82C						0.636732
WT						

4.3.3 ANOVA for Total Vein Area

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1713004	5	342600.8	2.415653	0.06707	2.639999
Within Groups	3261983	23	141825.3			
Total	4974987	28				

4.3.4 ANOVA and Significance Table for Total Vein Area: Vascular Bundle Area Ratio

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.006094	5	0.001219	2.96967	0.03273	2.639999
Within Groups	0.00944	23	0.00041			
Total	0.015534	28				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.630427	0.663246	0.160904	0.051713	0.733473
JL75B			0.903531	0.081355	0.024196	0.280347
JL82A				0.123558	0.049686	0.473635
JL82B					0.591467	0.192526
JL82C						0.05959
WT						

4.3.5 ANOVA for Interveinal Distance

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.004187	5	0.000837	1.730709	0.167603	2.639999
Within Groups	0.011128	23	0.000484			
Total	0.015315	28				

4.3.6 ANOVA for Leaf Thickness

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	484.9672	5	96.99345	1.392533	0.263906	2.639999
Within Groups	1602.009	23	69.65255			
Total	2086.976	28				

4.3.7 ANOVA for Mesophyll Cell Width

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	13.20232	5	2.640465	2.016142	0.114146	2.639999
Within Groups	30.12223	23	1.309662			
Total	43.32455	28				

4.3.8 ANOVA and Significance Table for Mesophyll Cell Length

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	109.9932	5	21.99864	3.256702	0.02282	2.639999
Within Groups	155.3623	23	6.754885			

Total	265.3556	28

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.983529	0.264982	0.031905	0.258927	0.029648
JL75B			0.147123	0.015863	0.204385	0.000516
JL82A				0.133715	0.779902	0.108002
JL82B					0.328795	0.728232
JL82C						0.376008
WT						

4.3.9 ANOVA and Significance Table for Mesophyll Cell Area

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	32102.54	5	6420.509	3.67496	0.013688	2.639999
Within Groups	40183.21	23	1747.096			
Total	72285.75	28				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.903069	0.689732	0.016778	0.186552	0.082998
JL75B			0.718202	0.004565	0.159211	0.029386
JL82A				0.011975	0.248343	0.076479
JL82B					0.211175	0.12234

JL82C		0.82958
WT		

4.3.10 ANOVA for Lobiness

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.002386	5	0.000477	1.753004	0.162643	2.639999
Within Groups	0.00626	23	0.000272			
Total	0.008646	28				

4.3.11 ANOVA and Significance Table for Lobing Area

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	102.259	5	20.45179	3.723432	0.012916	2.639999
Within Groups	126.3327	23	5.492725			
Total	228.5916	28				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.039197	0.61551	0.076624	0.006121	0.949008
JL75B			0.30598	0.586051	0.079818	0.146005
JL82A				0.210535	0.03472	0.715829
JL82B					0.335555	0.116798
JL82C						0.014011
WT						

4.3.12 ANOVA for Interveinal Mesophyll Cell Number

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	8.846935	5	1.769387	1.485553	0.233062	2.639999
Within Groups	27.39444	23	1.191063			
Total	36.24138	28				

4.3.13 ANOVA for Average Bundle Sheath Cell Area

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2737.27	5	547.4541	0.760337	0.587513	2.639999
Within Groups	16560.34	23	720.0146			
Total	19297.61	28				

4.3.14 ANOVA for Total Bundle Sheath Cell Area

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	612376.3	5	122475.3	1.783815	0.156028	2.639999
Within Groups	1579161	23	68659.16			
Total	2191537	28				

4.3.15 ANOVA and Significance Table for Bundle Sheath Cell Number

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*

Between Groups	8.27227	5	1.654454	5.507087	0.001779	2.639999
Within Groups	6.909722	23	0.300423			
Total	15.18199	28				

Post-Hoc Testing

Adjusted P-value = 0.003333

	JL75A	JL75B	JL82A	JL82B	JL82C	WT
JL75A		0.070414	0.228556	0.057218	0.079694	0.161688
JL75B			0.946243	0.014042	0.008182	0.01011
JL82A				0.019727	0.033326	0.05978
JL82B					0.434626	0.224844
JL82C						0.54535
WT						

Appendix 4.4 – ANOVAs for Section 4.3.4

4.4.1 ANOVA for Fitted VCmax

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	8204.643	5	1640.929	1.437464	0.267523	2.901295
Within Groups	17123.17	15	1141.544			
Total	25327.81	20				

4.4.2 ANOVA for Fitted J

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2430.226	5	486.0452	0.741306	0.604452	2.901295
Within Groups	9834.917	15	655.6611			
Total	12265.14	20				

4.4.3 ANOVA for Fitted TPU

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	8.054881	5	1.610976	0.482738	0.783827	2.901295
Within Groups	50.0575	15	3.337167			
Total	58.11238	20				

4.4.4 ANOVA for Fitted A@1000PAR

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*

Between Groups	92.52809	5	18.50562	0.632748	0.677954	2.901295
Within Groups	438.6962	15	29.24641			
Total	531.2243	20				

Appendix 4.5 – Identification of Rice Early Leaf Promoters

Background

This appendix briefly describes ones of the initial focuses of my PhD, which involved examining rice promoters for their potential to be used in transformation efforts. The diagrams in this appendix have been reproduced from my confirmation review report, or from lab meeting presentations.

Early in leaf development, the cells in the primordium are pluripotent, thus have the potential to be directed towards a variety of fates. However, as development progresses, these cells enter a particular lineage, and no longer possess this broad plasticity. This increasing temporal restriction of cell plasticity may explain the observation that, in some cases, introduction of genes intended to direct rice to a C4 structure results in either partial reversion of traits back to a C3 state, or low initial genetic penetration (P. Quick, personal communication).

Rice transformation constructs usually rely on monocot constitutive promoters, such as maize or switchgrass ubiquitin, or rice actin. However, this approach overexpresses genes of interest in all tissues at all time. When focussing on temporal or spatial restriction, which will be required for the introduction of C4 photosynthesis into rice, constitutive overexpression is not necessarily appropriate. While promoters allowing for the restriction of expression to either bundle sheath cells or mesophyll cells are available, promoters which restrict expression temporally are not so readily available in rice. There is evidence from maize that this temporal restriction of gene expression can be effective to alter leaf function (Sun et al., 2017).

In this work, my aims were to:

1. Develop a biolistics-based transient expression system to assay reporter gene expression in rice leaves
2. Identify potential promoters which could be used to tailor gene expression to different stages of rice leaf development

Biolistics Assay Optimisation

Several parameters required optimisation for the biolistics assay.

Initially, there was the problem of sample preparation. The methodology I used was adapted from one used for one-inch square onion slices. As the width of young leaves is significantly

less than an inch, I experimented with creating small woven leaf mats to produce sufficient surface area (shown in Figure A.4.6.1). This proved an effective approach, and later mats mixed primordia of different ages, from young P4 to mature P6 (not pictured).

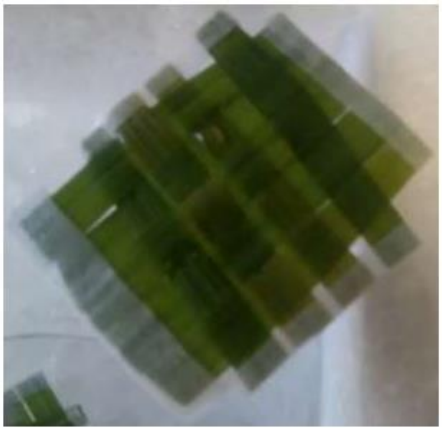


Figure A4.6.1 – Prototype rice mat used for biolistics assays.
Reproduced from confirmation review.

To test the system, I chose the pANIC5b plasmid, which had been designed for use in monocot transformation by biolistics (Mann et al., 2012). The reporter gene module of this plasmid consisted of a maize ubiquitin promoter, GUS reporter gene and *nos* terminator.

Several rounds of biolistics were carried out to optimise pressure needed for the rupture disk, distance between the rupture disk and sample, plasmid concentration, incubation time for GUS staining, and to check if primordia of all ages could be successfully transformed using this procedure. Eventually, I was able to achieve high levels of transformation in leaves (Figure A.4.6.2A), so was able to select conditions for future experiments, using 1800psi rupture disks, ~1$\mu g \mu l^{-1}$ plasmid, and a 24 hour GUS incubation period. I also briefly tested this protocol in root tissue, mindful that this comparison could be useful to examine if promoters of interest were expressed more strongly in either green tissue, or actively dividing tissue.

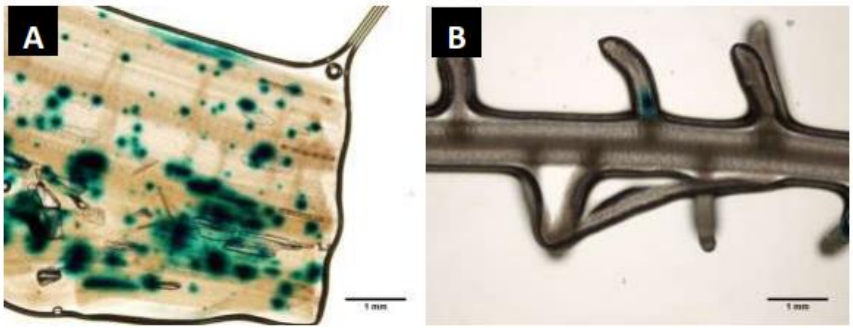


Figure A.4.6.2 – GUS staining of transformed tissue.
A – leaf tissue, B – root tissue. Scale bars 1mm.
Reproduced from confirmation review.

Cloning strategy and promoter selection

While the pANIC5b plasmid was clearly effective for transient transformation, it was not a suitable plasmid for testing different promoters. For this reason, I developed a cloning strategy using Golden Gate cloning, building constructs from the MoClo toolkits (Engler et al., 2014; Weber et al., 2011). I recreated the reporter module from the pANIC5b plasmid by cloning the maize ubiquitin promoter into an empty promoter module, and combining these with a GUS reporter gene and *nos* terminator. This plasmid was used as a control for my plasmid constructs.

I then selected a variety of promoters to test based on gene expression data of staged P3, P4-1 and P5 leaf primordia (van Campen et al., 2016). I was interested in genes which were expressed most highly at the P3 stage compared to other developmental stages, and also at a sufficiently high level that their promoters might drive gene expression. I also included the Ppsak promoter (Lin et al., 2017), which has been published as a strong green-tissue specific promoter to act as a control for rice.

As it is difficult to exactly determine the functional length of promoters, in most cases I cloned 2kb upstream of the gene start codon, except for when this overlapped with the coding sequence of the previous gene. I also wrote a pattern matching script in R to map the locations of common core and regulatory motifs in rice (motifs obtained from (Yamamoto et al., 2007)) to add supporting evidence that the promoter was likely encompassed in the cloned fragment. All promoters were cloned and sequenced before incorporation into the promoter module.

Cloned rice promoters did not express in rice

After demonstrating that the biolistics technique was robust, and that the cloning strategy was appropriate, I then tested constructs with both the Ppsak and various cloned promoters

from rice. For each round of biolistics, I included a maize ubiquitin control plasmid to account for experimental problems in an individual run. However, no rice promoter, including the green-tissue specific promoter, showed any GUS activity after transformation (Figure A.4.6.3). Data is not shown for the other test constructs, none of which showed any GUS expression.

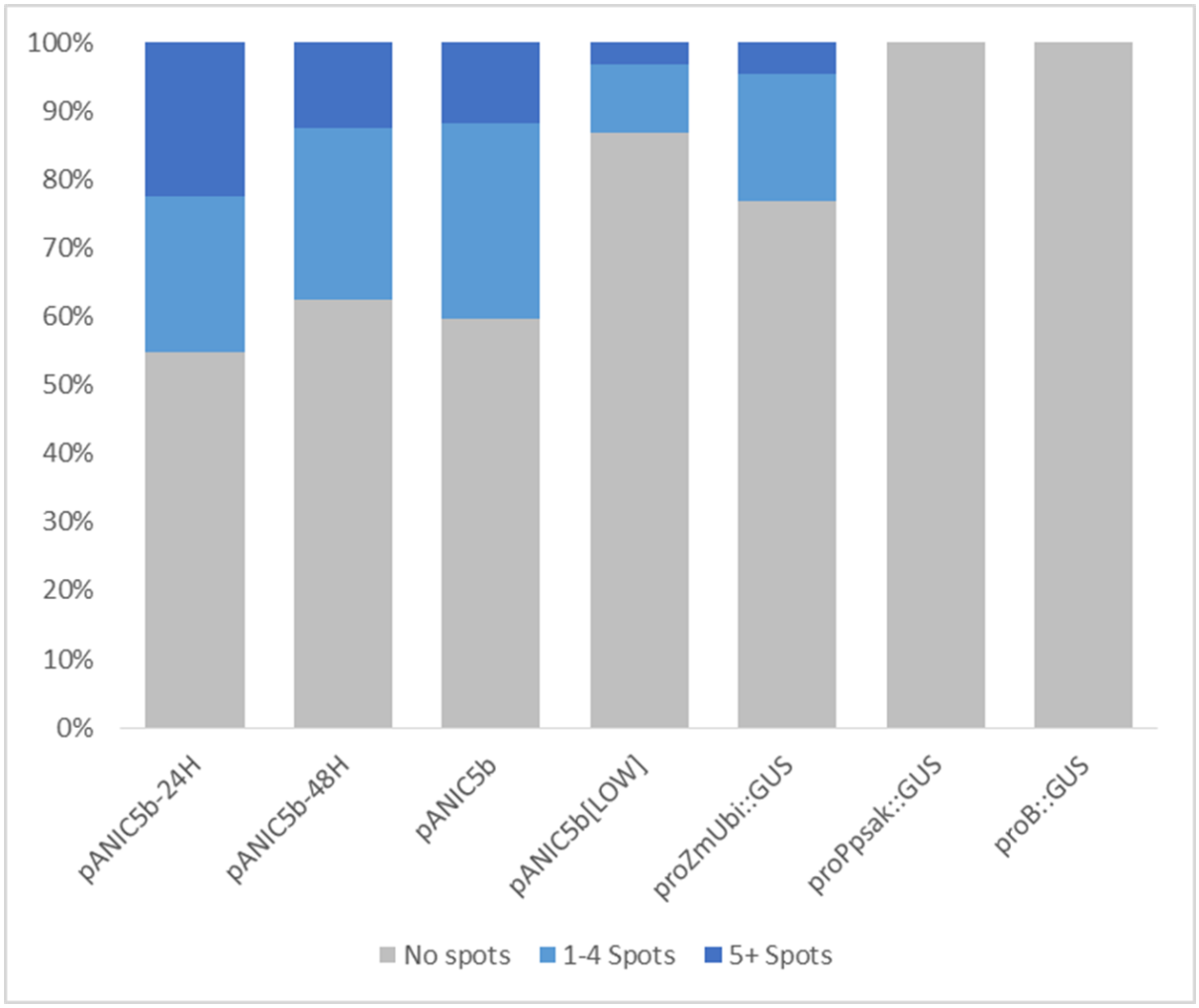


Figure A.4.6.3 – Results of biolistics experiments

Each bar represents GUS expression from a different plasmid. Grey represents no expression, light blue 1-4 spots on a single leaf fragment, and dark blue more than 5 sots on a single leaf fragment.

For the four pANIC5b bars, pANIC5b-24H and pANIC5b-48h were from the same experiment, and represent different incubation times used for GUS staining. pANIC5b and pANIC5b[LOW] were also from the same (but separate to pANIC5b-24H and pANIC5b-48H) experiment, and show expression from concentrations of ~1$\mu g \mu l^{-1}$ and ~300$ng \mu l^{-1}$.

References

van Campen, J., Yaapar, M.N., Narawatthana, S., Lehmeier, C., Wanchana, S., Thakur, V., Chater, C., Kelly, S., Rolfe, S. a, Quick, W.P., et al. (2016). Combined chlorophyll fluorescence and transcriptomic analysis identifies the P3 / P4 transition as a key stage in rice leaf photosynthetic development. Plant Physiol. 170, 1655–1674.

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.M., Werner, S., Jones, J.D.G., Patron, N.J., and Marillonnet, S. (2014). A Golden Gate modular cloning toolbox for plants. ACS Synth. Biol. 3, 839–843.

Lin, Z., Yan, J., Yan, H., and Wang, F. (2017). Characterization of a strong green tissue-specific motif in rice photosystem I gene promoter Ppsak. Plant Biotechnol. Rep. 11, 87–95.

Mann, D.G.J., Lafayette, P.R., Abercrombie, L.L., King, Z.R., Mazarei, M., Halter, M.C., Poovaiah, C.R., Baxter, H., Shen, H., Dixon, R.A., et al. (2012). Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (Panicum virgatum L.) and other monocot species. Plant Biotechnol. J. 10, 226–236.

Sun, X., Cahill, J., Van Hautegem, T., Feys, K., Whipple, C., Novák, O., Delbare, S., Versteele, C., Demuynck, K., De Block, J., et al. (2017). Altered expression of maize PLASTOCHRON1 enhances biomass and seed yield by extending cell division duration. Nat. Commun. 8, 14752.

Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. PLoS One 6.

Yamamoto, Y.Y., Ichida, H., Matsui, M., Obokata, J., Sakurai, T., Satou, M., Seki, M., Shinozaki, K., and Abe, T. (2007). Identification of plant promoter constituents by analysis of local distribution of short sequences. BMC Genomics 8, 67.