Interrogating the multi-stress tolerance of *Oryza australiensis* using novel genomic and phenomic strategies

Aaron L. Phillips

BAdvSc (Biology), BPhil, MRes

A thesis submitted to The University of Adelaide in fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Adelaide
Faculty of Sciences
School of Agriculture, Food and Wine
October 2022

This thesis is presented in loving memory of

Poppy Don Phillips (1930-2022)

and

Nanny Joan Phillips (1931-2020)

“I did it my way”

COVID-19 Impact Statement:

During 2020, 2021, and even into 2022, the work of this PhD project was greatly affected by the COVID-19 pandemic. When I embarked on my PhD journey in February 2019, I was working with eight lines of transformed domestic rice to define the role of a gene, vacuolar pyrophosphatase isoform 3 (OVP3), in plant growth and yield. In 2019, I ran a major growth experiment at Macquarie University with the help of my co-supervisor Prof. Brian Atwell, and around the same time started sequencing the genome of the wild rice relative *Oryza australiensis* at the Australian National University. At the time, this was to be a side project for me. In late 2019, I was the recipient of the Australian Plant Phenomics Facility PhD internship award, which allowed me to perform three major experiments investigating the response of the transformed domestic rice lines to salt, low soil N and drought. On a whim and purely out of interest, I decided to include *O. australiensis* in these experiments.

In total I experienced 12 weeks of home quarantine due to COVID during my candidature. As a result of the significant interruptions to experiments, in mid-2021 I had little choice but to abandon the OVP story and focus more on the response of the wild rice to stress events and the sequencing of its genome. Further to this, I experienced four hospitalisations between 2019 and 2022 due to health conditions and had to have surgery in mid-2022 to address this. Following surgery, my health was unstable for a few months as my doctors worked to find a medication regime that worked for me. Shortly after recovering from surgery, I contracted COVID and had to enter another week of isolation. These events further impacted the progression of my PhD. Fortunately, the inclusion of *O. australiensis* in the 2019 experiments meant I had access to detailed phenotypic data, and the genome sequencing (and other informatics projects) could be performed from home during my time away from campus.

Declaration:

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I acknowledge that copyright of published works contained within the thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Aaron L. Phillips

12th October 2022

Acknowledgments:

I acknowledge the Traditional Custodians of the lands on which I worked during my PhD – the Kaurna people of the Adelaide Plains, the Wattamattagal clan of the Darug nation on which Macquarie University is built, and the Ngunnawal people of Canberra – and pay my respects to their Elders past and present. I also acknowledge the use of a wild, native Australian plant species in this project, and the collectors and curators of the seed collections I made use of.

The PhD journey has been one of the most trying periods of my life to date. Without the support, guidance, and encouragement I received from the wonderful people I am surrounded by I would not have been able to complete this marathon.

I am deeply grateful to have worked with Prof. Rachel Burton as my primary supervisor for the last 3.5 years. Rachel is an incredible scientist and an even better person. She gave me every opportunity to learn and grow in her group professionally, which I sometimes took for granted, and was always ready to provide support. Rachel managed to foster a sense of community in her lab even as it continued to grow and was cognisant of the value each member brought to the table. She was never shy to remind me of my potential when I was feeling lost or not performing at my best, which helped motivate me to dig deeper and keep going. Even now, at the end of my PhD, she continues to provide support and opportunity.

Dr. Nathan Watson-Haigh came along rather unexpectedly, when I realised that the genome assembly work was perhaps going to be a bit more involved than originally thought – how naïve I was. Having zero bioinformatics experience, I was completely lost and often in a state of disrepair as I tried to learn concept after concept. But Nathan was an invaluable guiding light throughout. His patience and kindness towards me and belief that I could make sense of the constant stream of gobbledegook spewing from the HPC provided me with a perfect learning environment that I am grateful for.

There'll never be a way I can adequately thank Prof. Brian Atwell for all that he has done for me. He gave me my first internship 10 years ago after opening my eyes to the complexity and wonders of plant biology in the second year of my undergraduate studies. Since then, we have worked closely together on several species, each more interesting than the last. As time went on and we built a rapport, Brian was always just a phone call away. For work, yes, but also to just have a chat, to check in on each other and have a laugh. Brian has been the single most supportive person I have ever met. He encouraged me to take a chance on moving to Adelaide to pursue a PhD, something that has changed my life forever, and has always given a helping hand even as he faced troubles of his own. I am eternally indebted to Brian for everything he has done for me, professionally and otherwise. I dedicate this thesis to you and, going forward, part of all the work I do will be in honour of you, Brian.

Throughout my PhD I have met and worked with a number of people to whom I am grateful for their help and guidance. Dr. Andrew Scafaro has been a wonderful person to work with since my undergraduate days. Prof. Justin Borevitz, Dr. Ashley Jones, and Scott Ferguson of the Australian National University were instrumental in the sequencing of the wild rice genome. Special thanks go to Scott for his patience while I learned some of the ins and outs of informatics. I am deeply grateful to the staff of The Plant Accelerator (TPA) for their continued support. In particular, I thank Lidia Michis, Nicole Bond and Fiona Norrish for their kindness and willingness to help as I balanced running my experiments with the complications of life. I also thank Dr. Bettina Berger and Dr. Darren Plett of TPA for their guidance before, during and after my time at TPA. I also extend my thanks to the members of the Centre of Excellence in Plant Energy Biology. I am grateful to **all members** of the Burton lab (which, at the time of writing, is 35 members strong!) but in particular I would like to thank Sandy Khor, Renee

Phillips, Carolyn Schultz, and Lisa O'Donovan. You guys are the powerhouse of the lab, and we would all be lost without you.

I am grateful to lab members and fantastic friends Dr. James Cowley, Dr. Emma Aspin, Jakob Schulz, Ali Gill, Dr. Tina Bianco-Miotto, Jacquelin Barsby, Amy Doan, Zane Marks, Lina Herliana, and Mel Ford for providing a beautiful community both within and outside of the lab and helping me escape the rigours of the PhD from time to time. I thank the Adelaide Frontrunners – an LGBT+ running group – for creating the most vibrant and accepting community and for helping me stay healthy. Thank you to Justin and Dr. Tanya Ayris (and now little Hali Ayris), Ben Ayris, Nick Dreyer and Jimmy Smit for all the game nights and quality time. Thank you to Sam Riley and Sheldon Skanes for the wonderful outdoor adventures we have had that helped me stay grounded.

I must also acknowledge the continued love, support, and friendship of those people I left behind in Sydney. First and foremost, I thank my immediate family. It has been a rough couple of years, but we've made it through together as we always will. I love you guys unconditionally and thank you for helping me reach this point. To my mum and dad, Susan and David, and my sisters, Ashley and Lauren, your unwavering support has been instrumental over the course of my life but in particular in the last four years. In everything I do, I try to make you guys proud of the person you helped form. To my best mate and brother, the Pharoah to my Yugi, the Yusei to my Leo, Jarrod Kelly – you and I will always be there for each other, a universal constant. I can't describe what you have done for me. I need to also thank Matthew Alfonzetti, Belinda Fabien, and Katie Berthon, my very good friends and science soundboards.

Finally, to my wonderful partner, Tom Porta. Even though you annoy me to no end, the sunshine you bring to each, and every day gives me the motivation to get up when I'm down and be a better person. Thank you for your dedication to us and for being my home.

Table of Contents:

Chapter 1 – General Introduction	1
Thesis Introduction	
Thesis Structure	
Chapter 2 – Literature Review	9
General Introduction	
How Are Rice Crops Affected by Different Stresses?	
Heat	
Drought	
Low N	
Utilising Intraspecific Variation to Improve Domestic Rice Cultivars	
Wild Rice Relatives as a Route to Multi-Stress Tolerance in Rice	
Oryza australiensis is a Wild Rice Relative with Multi-Stress Tolerance	
The Role of Wild Species' Genomes in Improving Stress Tolerance	
Concluding Remarks	
Chapter 3 – Photosynthetic traits of Australian wild rice (*Oryza australiensis*) confer tolerance to extreme daytime temperatures	49
Abstract	
Introduction	
Methods	
Results	
Discussion	
Chapter 4 – Responses of *Oryza australiensis* to individual and combined drought and low nitrogen stress	119
Abstract	
Introduction	
Materials and Methods	
Results	
Discussion	
Chapter 5 – The first long-read nuclear genome assembly of *Oryza australiensis*, a wild rice from northern Australia	163
Abstract	
Background	

Methods	
Results	
Discussion	
Chapter 6 – *CLAW*: An automated Snakemake workflow for the assembly of chloroplast genomes from long-read data	233
Abstract	
Introduction	
Methods	
Results	
Discussion	
Chapter 7 – Summary and Future Directions	263
Appendix I – Career and Research Training Skills (CaRST) Certification	275
Appendix II – Vacuolar pyrophosphatase literature review and research proposal	279
Appendix III – Rice cold tolerance literature review and research proposal	343
Appendix IV – Epigenetic mechanisms involved in intrauterine growth restriction and aberrant kidney development and function	387
Appendix V – Conference presentations and posters	389

Note to Readers:

The Appendices of this thesis are included to highlight the significant change of topic that I had to make mid-candidature (see Appendix II), and to show other academic pursuits I followed during my candidature in response to the restrictions imposed by the COVID-19 pandemic.

Out of my own interest and, hopefully, to make the experience of reading this thesis more enjoyable, each chapter is accompanied by a piece of digital art generated by A.I. using an online application called Wombo. Wombo accepts a 100-character prompt from the user, from which the art is generated after the user selects an artistic theme. The prompts used for each chapter contain keywords from each chapter.

List of Figure and Tables (in order of appearance in each chapter):

Chapter 1 – General Introduction

Figure 1.1, Page 7 – Schematic overview of research questions

Figure 1.2, Page 8 – Schematic overview of thesis structure

Chapter 2 – Literature Review

Figure 2.1, Page 47 – Distribution of *Oryza australiensis* in Australia

Figure 2.2, Page 48 – Temperature and precipitation patterns for Akita (Japan) and Kununurra (Australia)

Chapter 3 – *Oryza australiensis* heat tolerance

Table 3.1, Page 67 – Gas exchange measurements and calculations derived from heat stressed *O. sativa* and *O. australiensis* plants

Figure 3.1, Page 88 – Shoot dry mass of *O. sativa* and *O. australiensis* plants grown at 25, 35, and 45°C

Figure 3.2, Page 89 – Effect of different growth temperatures on the elongation rates of leaves of wild and domestic rice

Figure 3.3, Page 90 – The effect of high temperature (shock) on soluble sugar production in the leaves of *O. sativa* and *O. australiensis* plants

Figure 3.4, Page 91 – Leaf elongation rates (LERs) and shoot dry weights of *O. sativa* and *O. australiensis* plants grown at 30 and 45°C under elevated (700 ppm) atmospheric CO_2 concentrations

Figure 3.5, Page 92 – Effect of long- and short-term heat exposure on respiration in the light (R_{light}), CO_2 compensation point in the absence of mitochondrial respiration (Γ^*), and electron transport rates (ETR) in *O. sativa* and *O. australiensis*

Figure 3.6, Page 93 – RACiR curves for *O. sativa* and *O. australiensis* grown under variable temperature and CO_2 regimes

Figure 3.7, Page 94 – Effect of heat on CO_2 assimilation and the abundance of RbcL and Rca in *O. sativa* and *O. australiensis*

Supplementary Figure 3.8, Page 95 – Root dry weight of *O. australiensis* and *O. sativa* grown at 25, 35 and 45°C

Supplementary Figure 3.9, Page 96 – Rca:RbcL in temperature-treated *O. australiensis* and *O. sativa* leaves throughout the photoperiod

Supplementary Figure 3.10, Page 97 – Root to shoot ratio (dry weights) of *O. australiensis* and *O. sativa* grown at 25, 35 and 45°C

Supplementary Figure 3.11, Page 98 – Gas exchange data from *O. australiensis* and *O. sativa* exposed to 30°C or 45°C.

Supplementary Figure 3.12, Page 99 – Gas exchange data from *O. australiensis* and *O. sativa* plants grown at 30°C or 45°C and 400 or 700 ppm CO_2

Chapter 4 – Responses to combined drought and low soil nitrogen

Table 4.1, Page 151 – ANOVA summary table

Figure 4.1, Page 152 – Representative images of *O. sativa* and *O. australiensis* plants grown under different combinations of water and nitrogen availability

Figure 4.2, Page 153 – PSA, RGR, and greenness of *O. sativa* and *O. australiensis* plants grown under different combinations of water and nitrogen availability

Figure 4.3, Page 154 – Physiological responses of *O. sativa* and *O. australiensis* plants to different combinations of water and nitrogen availability

Figure 4.4, Page 155 – Root architecture responses of *O. sativa* and *O. australiensis* plants grown under different combinations of water and soil N availabilities

Figure 4.5, Page 156 – Water use responses of *O. sativa* and *O. australiensis* plants grown under different combinations of water and soil N availabilities

Supplementary Figure 4.6, Page 157 – Representative images of *O. sativa* and *O. australiensis* plants grown under either drought or low soil N

Supplementary Figure 4.7, Page 158 – PSA, RGR, and greenness of *O. sativa* and *O. australiensis* plants grown under either drought or low soil N

Supplementary Figure 4.8, Page 159 – Physiological responses of *O. sativa* and *O. australiensis* plants to drought or low soil N

Supplementary Figure 4.9, Page 160 – Relative changes in PSA, RGR, and greenness for *O. sativa* and *O. australiensis* plants to drought or low soil N

Supplementary Figure 4.10, Page 161 – Relative changes in PSA, RGR, and greenness for *O. sativa* and *O. australiensis* plants grown under different combinations of water and soil N availabilities

Supplementary Figure 4.11, Page 162 – Ratio of lateral root number to axial root number in *O. sativa* and *O. australiensis* plants grown under individual and combined low soil N and drought conditions

Supplementary Figure 4.12, Page 162 – Correlation between shoot dry weight and projected shoot area for *O. sativa* and *O. australiensis* plants

Chapter 5 – Nuclear genome assembly for *Oryza australiensis*

Table 5.1, Page 182 – Summary of reads used for each step of genome assembly

Table 5.2, Page 182 – Genome assembly statistics for *O. australiensis* Keep River

Table 5.3, Page 186 – Summary statistics of the Pac Bio Iso-Seq data

Table 5.4, Page 186 – Repeat elements in the *O. australiensis* KR scaffolds

Figure 5.1, Page 206 – Locations of *O. australiensis* seed collection sites

Figure 5.2, Page 207 – Alignment of scaffolded *O. australiensis* KR contigs to the *O. sativa* Nipponbare reference genome

Figure 5.3, Page 208 – Genetic distance between *Oryza* species/accessions derived from Illumina short read libraries

Supplementary Figure 5.4, 209 – High-quality transcripts derived from Iso-Seq data

Supplementary Figure 5.5, Page 210 – Long-read coverage of the *O. australiensis* Keep River contigs

Supplementary Figure 5.6, Page 211 – Whole genome alignment of *O. australiensis* Keep River contigs and *O. sativa* Nipponbare chromosomes

Supplementary Figure 5.7, Page 212 – Long-read coverage of the *O. australiensis* Keep River pseudomolecules

Supplementary Figure 5.8, Page 213 – Location of contig-contig boundaries, and assessment of, the structural variants (SVs) detected in *O. australiensis* KR pseudomolecules

Supplementary Figure 5.9, Page 214 – Inspection of Chr2 for a previously reported inversion

Supplementary Figure 5.10, Page 215 – Insertion times for the major LTR-RT superfamilies identified in the scaffolded *O. australiensis* nuclear genome assembly

Supplementary Figure 5.11, Page 216 – SNVs identified between the assembled *O. australiensis* Keep River genome and *O. sativa* Nipponbare reference genome

Supplementary Figure 5.12, Page 217 – Short-read coverage plot

Chapter 6 – *CLAW*: a chloroplast genome assembly tool

Table 6.1, Page 246 – Information on long reads used as input for *CLAW* and the *Flye*-generated chloroplast genome assembly statistics

Table 6.2, Page 247 – Performance details of *CLAW* runs

Figure 6.1, Page 254 – Circularisation of linear reference genomes

Figure 6.2, Page 255 – Representative reference-sample chloroplast genome alignments

Figure 6.3, Page 256 – Repeat graphs for the 19 chloroplast genomes assembled by *CLAW*

Supplementary Table 6.3, Page 257 – Information on long reads used as input for *CLAW* and the *Unicycler*-generated chloroplast genome

List of publications – published and proposed:

The following publications relevant to this thesis are listed in order of submission:

1. **Phillips, A.**, Ferguson, S., Watson-Haigh, N., Jones, A., Borevitz, J., Burton, R., Atwell, B. 2022. The first long-read nuclear genome assembly of *Oryza australiensis*, a wild rice from northern Australia. *Plant. Mol. Biol.* DOI: 10.1038/s41598-022-14893-5

2. **Phillips, A.**, Scafaro, A., Atwell, B. 2022. Photosynthetic traits of Australian wild rice (*Oryza australiensis*) confer tolerance to extreme daytime temperatures. *Plant Mol. Biol.* DOI: 10.1007/s11103-021-01210-3

3. **Phillips, A.**, Ferguson, S., Watson-Haigh, N., Burton, R. Unpublished. *CLAW*: An automated Snakemake workflow for the assembly of chloroplast genomes from long-read data. Target Journal: *PLoS Comput. Biol.*

4. **Phillips, A.**, Plett, D., Berger, B., Burton, R., Atwell, B. Unpublished. Responses of *Oryza australiensis* to individual and combined drought and low nitrogen stress. Target Journal: *J. Exp. Bot.*

Chapter 1

Thesis Introduction

Wombo Prompt

"*Oryza australiensis* Northern Territory, Australia"

Wombo Theme

"Comic"

Thesis Summary/Abstract:

Elite high-yielding rice cultivars are, by necessity, grown under high-input scenarios, requiring substantial nitrogen and water inputs to achieve their yield potential. As a result of intensive breeding for performance in the relatively benign conditions of the tropics, they are also susceptible to a number of abiotic stresses including heat wave events and drought. To keep pace with increasing food demand, domestic rice will need to be augmented such that its realised yield is close to its potential yield. Recently there has been great interest in exploring novel germplasm for traits associated with increased nitrogen use efficiency (NUE), water use efficiency (WUE), and stress tolerance. Part of this exploration includes the exploitation of wild *Oryza* species. There are ~24 recognised wild *Oryza* species, each occupying and adapted to different ecological niches. Amongst them, *O. australiensis* has received attention because of its tolerance to leaf brown hopper and extreme heat and its hypothesised tolerance to drought conditions. Native to the northern savannahs of Australia, this species has evolved in a relatively hot and seasonally dry environment, with high rainfall variability and poor soil nutrition. Because of this, we hypothesised that this species might also possess both high WUE and high NUE. However, knowledge of the mechanisms behind its tolerance to extreme heat are lacking, and no reliable experiments have been performed to substantiate claims of drought tolerance in *O. australiensis*. Further, the International *Oryza* Map Alignment Project, an initiative that aimed to sequence *Oryza* genomes for the benefit of domestic rice breeding, called for the genome of *O. australiensis* to be sequenced in 2003. However, there has been no publicly available *O. australiensis* genome since. This thesis presents the results of several novel phenotyping techniques including RACiR gas exchange and state of the art imaging techniques to resolve questions of both heat and drought tolerance in the wild rice. I also present the first long-read assembly of the *O. australiensis* genome, and a novel

tool, *CLAW* (Chloroplast Long-read Assembly Workflow), developed for the assembly of chloroplast genomes derived from larger long-read libraries. Having identified several traits of interest in *O. australiensis* under stress conditions and having provided significant tools (the genome assembly and a tool for the assembly of chloroplast genomes), this thesis sets the stage for future studies hoping to identify candidate genes for the improvement of domestic rice performance under stress conditions.

Thesis Structure:

This thesis is comprised of seven chapters: a general thesis introduction (Chapter 1), a literature review (Chapter 2), two published research articles (Chapters 3 and 5), two unpublished research manuscripts (Chapters 4 and 6), and a final general discussion (Chapter 7).

Chapter 3 is a published research article that details experiments performed to dissect the mechanisms of extreme heat tolerance in *Oryza australiensis*, a wild relative of the domestic rice species *O. sativa* (Phillips et al. 2022, *Plant Molecular Biology*, doi: 10.1007/s11103-021-01210-3). Heat tolerance in *O. australiensis*, and heat sensitivity in *O. sativa*, had been previously established, and indicated that the photosynthetic reactions were a major driver for the observed species differences. However, the physiological mechanisms behind this tolerance, and the consequences for growth, were largely unknown. The experiments described in Chapter 3 reveal somewhat counter-intuitive responses of the wild rice to both long and short-term extreme heat that confer tolerance.

Chapter 4 is an unpublished research manuscript that details experiments exploring the responses of *O. australiensis* and *O. sativa* to drought and low nitrogen stress, both

individually and combined. Given the range of this species in the seasonally dry environments of northern Australia, drought tolerance has been hypothesised for *O. australiensis* for decades. Also, given that the soils of Australia are low in nitrogen (they contain about one tenth of the N that induces low nitrogen symptoms in *O. sativa*), it was hypothesised that low nitrogen tolerance would occur in *O. australiensis* too. However, no reliable studies have been undertaken to explore these hypotheses. A previous PhD project attempted to address the research questions covered in this chapter, though it became obvious that those plants were exposed to toxic levels of soil nitrogen and so the observed plant responses to drought were confounded. The issue of soil nitrogen concentration was fixed in the present study, and so this chapter presents the first reliable account of a detailed physiological investigation into drought and low nitrogen tolerance in *O. australiensis*.

Chapter 5 is a published research article in which we present the first publicly available long-read assembly for the *O. australiensis* nuclear genome (Phillips et al. 2022, *Scientific Reports*, doi: 10.1038/s41598-022-14893-5). Almost ten years ago this genome was identified by the International *Oryza* Map Alignment Project (IOMAP) as an important resource for domestic rice breeding. However, all efforts to assemble a high-quality genome have failed since. This is largely due to the inability of short-read technology to resolve repeat-rich regions of genomes, of which the *O. australiensis* genome is comprised of ~74%. Recent advances in long-read based sequencing platforms made it possible for us to assemble a high-quality genome for this species, which is now publicly available on NCBI (accession number: JAIFGZ000000000). This assembly of the *O. australiensis* genome will serve as the reference genome of the species until a more contiguous assembly can be produced. Thus, it is expected that

this assembly will assist researchers around the world in the identification of genes and/or regulatory elements associated with tolerance traits.

Chapter 6 is an unpublished manuscript that describes a user-friendly fully automated Snakemake based workflow for the assembly of chloroplast genomes from reads contained in larger read libraries. When scientists sequence the genome(s) of species it is common for leaf material to be the tissue of choice to obtain DNA from. Photosynthetically active leaves contain many thousands of chloroplasts, and each chloroplast contains hundreds of copies of its own DNA-based genome. Thus, in sequencing the DNA extracted from leaf tissue, scientists are also inadvertently sequencing chloroplastic DNA. Up to 10% of a sequencing read library derived from leaf tissue will be made up of chloroplast sequences. This is something that has been known since the advent of shotgun sequencing technologies, and researchers often have in-house scripts for the extraction and, sometimes, utilisation of these reads. Often these scripts are not released to the public or have been written in complicated ways and do not take into consideration the skill level of those accessing it. This presents as a significant barrier to entry for the biologist with little to no bioinformatics training, as I was when I started my PhD. Thus, this automated tool, which requires minimal user input, was designed with such people in mind. When the tool completes successfully the user is left with a chloroplast genome that can be used for downstream applications. It is hoped that, once published, this tool will allow biologists the take advantage of the reads that they inadvertently sequenced, adding extra value to sequencing projects.

Chapter 7 presents a summary and discussion of the findings of this thesis. Future directions for research are also suggested in this chapter.

I have included in the Appendices some academic pursuits that required significant time to complete during my candidature. While not all of these pursuits are directly related to my thesis, I believe they each represent new learnings and/or skills that aided me in the completion of my thesis. They are as follows:

- A certificate of completion for the mandatory 120-hour Career and Research Skills Training (CaRST) program that the University of Adelaide requires its postgraduate students to complete (Appendix I),
- The literature review and research proposal I wrote for my original PhD topic, in which I explore the role of vacuolar pyrophosphatases in plant growth, development, and yield (Appendix II),
- The literature review and research proposal I wrote for Agri-Futures Australia – a potential industry partner – in relation to their call for research into cold tolerance in Australian rice crops (Appendix III),
- A paper entitled “Epigenetic mechanisms involved in intrauterine growth restriction and aberrant kidney development and function” published during my PhD candidature on which I am a co-author for my contribution as statistician (Appendix IV), and
- Presentations and posters I made during my candidature for conferences and other gatherings (Appendix V).

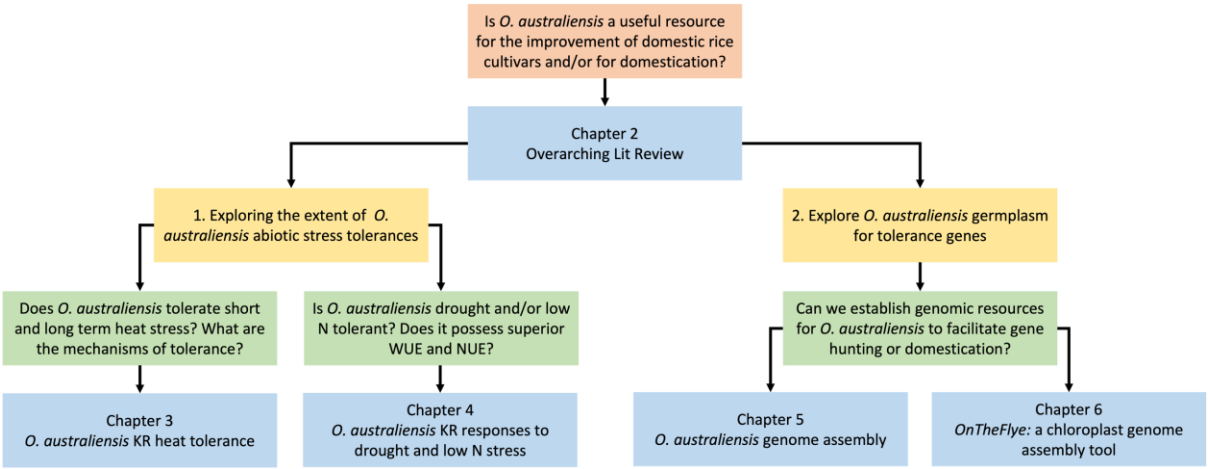


Figure 1.1. Overview of the research questions that guided thesis. The orange box indicates the overarching research question. The yellow boxes indicate the core themes derived from a review of the literature. The green boxes indicate the research questions, and the blue boxes indicate the relevant chapters of the thesis.

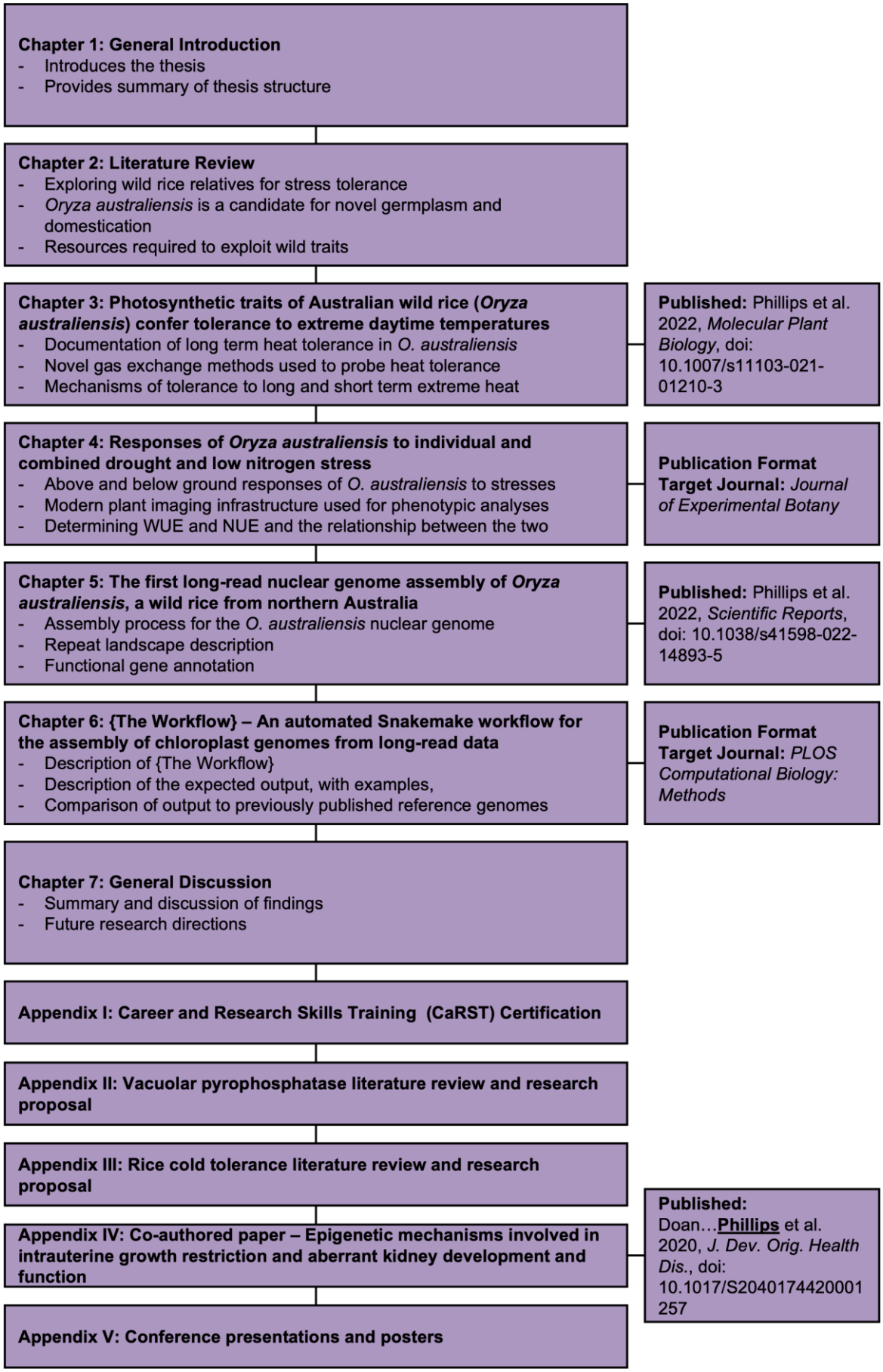


Figure 1.2. Overview of the chapters and their contents presented in this thesis.

Chapter 2

Literature Review

Wombo Prompt

“Threatened rice paddy”

Wombo Theme

“Comic”

General Introduction

Plant growth and yield are multivariate traits, determined by the interaction of various genetic and environmental factors (Kosaki *et al.*, 1989; Bibi *et al.*, 2015). Abiotic stresses (i.e. unfavourable conditions including drought, flooding, saline soils, heat, chilling, and low nitrogen soils) impose a significant yield gap on agricultural systems around the globe and, as a result, realised yield is often far below the potential yield of a crop (Powell *et al.*, 2012). While these systems currently produce enough food and fibre to adequately feed and clothe the global population (even though unequal distribution of resources exposes millions of people to famine), crop productivity will have to increase by 100–110% with fewer inputs (e.g. fertiliser) on ever shrinking suitable cropping land before 2050 to match increases in demand as the human population exceeds nine billion people (Tilman *et al.*, 2011). This feat is further complicated by the changes in climate observed around the world (Powell *et al.*, 2012). Extremes in temperature (both chilling and heat wave events) and water regimes (i.e. drought and flooding) are increasing in frequency, duration, and intensity, and it is expected that rising sea levels will impact many coastal agricultural systems by increasing soil salinity (Frich *et al.*, 2002; Tripathi *et al.*, 2016; Shahid *et al.*, 2018). Also of concern is the impact that these changes will have on crop performance when they occur in combination. When abiotic stressors interact, they can exacerbate the individual effect(s) of each concomitant stress, and/or present new challenges not faced when either stress is experienced individually (Sarkar *et al.*, 2019; Costa *et al.*, 2021). For example, drought conditions may amplify the negative impacts of heat stress because there is insufficient water for leaf cooling by evapotranspiration. Also, when drought conditions interact with low soil nitrogen (for example, in a scenario when fertiliser inputs are reduced) root architecture must respond to both stresses at the same time and there may be conflicts in the phenotypes that would best alleviate each

stress individually (e.g., thick, suberin-rich roots, vs investment in fibrous roots with high surface area). The compounding negative effects of combined stresses and the sometimes-conflicting responses that they elicit increase yield gaps further. As such, it is necessary to understand how plants respond to a multitude of stress types at different levels of biological organisation (i.e., from genes to whole-plant physiology), especially if we are to generate crops that can maintain growth and yield in the face of variable conditions.

How are Rice Crops Affected by Different Stresses?

Rice is a staple food for ~60% of the human population and demand for rice is expected to increase by 18% by 2050 due to population increase (see Yuan *et al.*, 2022). Major rice producing regions, such as Southeast Asia, which currently produces about 40% of global rice export, are at risk of increasing already substantial yield gaps, though this will likely occur in a region-specific manner (Yuan *et al.*, 2022). Stresses associated with climate change are predicted to reduce rice yield significantly. For example, rice yield is expected to decline by between 3.3 and 10% depending on when a crop is exposed to elevated temperature (Peng *et al.*, 2004; Zhao *et al.*, 2017). Further, rice yield in India is expected to fall from 2.13 t/ha to 1.67 t/ha under climate change (RCP 8.5 scenario; Debnath *et al.*, 2020). Changes in the supply of rice due to changes in climatic conditions or soil quality would negatively impact billions of people around the world. For example, increasing salinity in the coastal soils of Bangladesh due to rising sea levels is expected to cause human settlements to retreat inland in part due to loss of coastal agricultural (including rice crop) productivity (Chen and Mueller, 2018). Thus, efforts must be made to predict and ameliorate the impact of climate change on rice production.

Due to their evolutionary history, most species in the genus Oryza are adapted to and occur in low latitude environments, typically characterised by warm temperatures and plentiful water supply in the growing season. Unlike temperate crops such as wheat or barley, rice crops are susceptible to 'low' temperatures and are typically sensitive to drought conditions. Despite having evolved in more tropical climates, the upper temperature limit for rice crops is ~32^{o}C, and there is wide variation between high-yielding O. sativa cultivars in their tolerance to water logging. To understand how rice yield might be stabilised or perhaps even increased under sub-optimal or stress conditions it is necessary to understand how the plants are affected by and respond to stress.

Heat Stress

On top of increasing average day and night temperatures, climate models predict an increase in the frequency, duration, and intensity of heat wave events in the coming decades (IPCC, 2019, 2022). Temperature is the dominant influence on growth and development in all crops, including rice. Thus, temperature regimes have a significant impact on the final yield of a rice crop. For example, it has been predicted that the yield of a rice crop will decline by 10% for every 1^{o}C increase in night time temperature (Peng *et al.*, 2004), though yield losses between 21-80% have been observed under climate change scenarios (Li et al., 2003; Liu *et al.*, 2021). The reproductive phase is usually regarded as the most heat-sensitive phase of rice development. However, all stages are susceptible to heat stress (i.e., temperatures above ~32^{o}C) and plant responses to high temperatures at all stages will ultimately affect final yield. As such it is necessary to understand how rice plants respond to heat stress during both vegetative and reproductive phases.

The vegetative-growth phase encompasses the period of development in which seedlings rapidly elongate, leaves expand and mature, and tillers form. During a high temperature event, the surface temperature of soils can greatly exceed the temperature of the air (Kolb and Robberecht, 1996). Seedlings, which occupy this space, can therefore be exposed to temperatures far greater (e.g., 30°C air temperature vs 75°C soil surface temperature; Kolb and Robberecht, 1996) than the ambient air temperature. Given that seedlings do not have well-established root systems, their ability to mobilise water for evaporative cooling is restricted. In support of these observations, Wei et al. (2013) show that the effect of high temperature is much more severe on soil-grown rice seedlings than it is on hydroponically-grown seedlings. Given that the authors did not control for soil surface temperature, it is plausible that the reason soil-grown rice was more susceptible to high temperatures was because the seedlings were experiencing a higher growth temperature than their hydroponically grown counterparts. Thus, seedlings exposed to high temperatures have reduced growth rates and high mortality (Wahab *et al.*, 2020). Reduced plant number will ultimately lead to reduced yield. This may become a significant problem for dry-seeded rice crops as temperatures climb.

Rice plants that survive to maturity are also susceptible to heat. Mature rice leaves are not subject to the extreme soil surface temperatures discussed above, however their photosynthetic machinery is sensitive to temperatures above ~32ºC. Photoassimilates fuel processes associated with defence, growth, and reproduction, so reductions in photosynthetic capacity have negative impacts on crop yield (Bita and Gerats, 2013). Both the light and carbon fixation reactions of photosynthesis demonstrate sensitivity to heat. However, uncertainty over which process limits plant productivity under heat stress remains because both sets of reactions are co-dependent. Historically, the light

reactions were thought to be the main limitation to photosynthesis under heat stress. Changes in the stability of the thylakoid membrane under high temperature reduces the efficiency of photosystem I and II (PSI and PSII) and the electron transport chain (ETC) that together generate the energy (ATP) and reducing equivalents (NADPH) required for the carbon fixation reactions to take place (Horváth *et al.*, 2012). However, more recent evidence suggests that the enzymes directly involved in carbon fixation are inhibited by high temperatures (Salvucci and Crafts-Brandner, 2004*b*). The temperature sensitivity of the carbon fixation reactions is exemplified by the deactivation of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), the enzyme that fixes atmospheric CO2 into two molecules of 3-phosphoglyceric acid (3-PGA) by combining it with one molecule of ribulose-1,5-bisphosphate (RuBP; Salvucci and Crafts-Brandner, 2004*b*). It has been shown that the efficiency at which Rubisco fixes CO2 declines at temperatures below the thermal tolerance threshold of the PSII complexes (i.e., the temperatures required to reduced Rubisco activity are lower than those required to reduced PSII activity; Sharkey and Zhang, 2010; Ashraf and Harris, 2013), though it is worth noting that this is likely a species-specific phenomenon (Salvucci and Crafts-Brandner, 2004*c*; Scafaro *et al.*, 2012).

Rubisco is a multimer of eight large and eight small protein subunits (RbcL and RbcS, respectively). The large subunits are catalytically active – that is, they participate directly in CO_2 fixation (Andersson and Backlund, 2008). The genes that encode RbcL are highly conserved across photosynthetic organisms, which indicates that they have reached an evolutionary equilibrium and that variations that alter function have been purged (Parry *et al.*, 2003; Studer *et al.*, 2014). Indeed, studies into the sequences of RbcL in wild and domestic species of wheat revealed very little diversity in structure (Prins *et al.*, 2016). Despite this, there is great interest in improving Rubisco kinetics

with the goal of increasing CO_2 fixation (Carmo-Silva *et al.*, 2015). More recently, attention has turned to RbcS as these subunits are believed to moderate the activity of RbcL (Ishikawa *et al.*, 2011; Morita *et al.*, 2014). From this, amino acid residues have been identified in both RbcL and RbcS that could be modified for improved function. However, because of the low sequence diversity in the genes that encode RbcL, and, to a lesser extent RbcS, targeted modifications of the Rubisco structure are deviations from forms that have been maintained throughout evolutionary history and will therefore likely have reduced functionality (Parry *et al.*, 2006; Tcherkez *et al.*, 2006; Studer *et al.*, 2014).

Despite the strong selection pressures that have directed the structure of Rubisco, the active sites maintain the ability to fix oxygen as well as CO_2 in a process termed photorespiration (Maurino and Peterhansel, 2010). The oxygenase activity that gives Rubisco part of its name is thought to aid in photoprotection by providing a sink for excess electrons, though this is still not certain. Whatever the function, the cost of 3.5 ATP and 2 NADPH per oxygenase event results in an estimated 20% yield loss in American wheat farms (Walker *et al.*, 2016). Importantly, the rate of photorespiration increases as ambient temperatures increase because Rubisco loses affinity for CO_2 (see Dusenge *et al.*, 2019). Thus, it is expected that the cost of photorespiration will increase as average daily temperatures increase.

Along with increases in rates of photorespiration at high temperature, the rate of CO_2 fixation by Rubisco is reduced (Salvucci and Crafts-Brandner, 2004*a*,*b*,*c*). For example, domestic rice (*O. sativa* japonica) leaves exposed to 45^{o}C capture ~50% less CO_2 than when grown under optimal conditions (Scafaro *et al.*, 2016). This occurs, in part, because the Rubisco active sites bind to the sugar phosphates they produce and

so become unavailable to fix more CO_2 (Salvucci and Crafts-Brandner, 2004*a*). The phenomenon of Rubisco deactivation under heat is common across all tested plant taxa, though there are species-specific differences in the temperature sensitivity of Rubisco deactivation (Salvucci and Crafts-Brandner, 2004*c*). The relatively low diversity of RbcL and RbcS genes again means that improvements in Rubisco kinetics under heat stress are unlikely to come from natural variation in these genes nor from targeted modifications of Rubisco – some, though not all (Whitney *et al.*, 2009), Rubisco transformants suffer from faulty Rubisco assembly (Gutteridge and Gatenby, 1995). Instead, much work has been devoted to exploring the role that Rubisco's accessory proteins can play in ameliorating heat-induced reductions in CO_2 fixation. Key amongst these proteins is the chaperone Rubisco activase (RCA; Parry *et al.*, 2013).

RCA is part of the AAA+ ATPase protein family and is found in all plant species (Portis, 2003). In its active form, it is a hexameric enzyme that hydrolyses ATP to remove the inhibitory sugar phosphates from Rubisco active sites (Wachter and Henderson, 2015). As such, RCA works to maintain the number of Rubisco active sites present in a cell (Mueller-Cajar, 2017). There are several conserved amino acid motifs that make up the RCA monomers. These sites are involved in the Rubisco-RCA interaction. The remaining domains of RCA are more diverse between species (Mueller-Cajar *et al.*, 2014). It may be that because RCA is an accessory protein (i.e., it does not participate directly in photosynthesis) it is more likely to accrue mutations that can then be acted on by natural selection (Hsu *et al.*, 2016; Alvarez-Ponce *et al.*, 2016). Thus, the overall sequence diversity of RCA is greater than that of Rubisco. The ubiquity of RCA across the plant kingdom, its high sequence diversity, and its ability to activate Rubisco makes it a prime candidate for improving photosynthetic capacity (Parry *et al.*, 2013).

However, for many (though not all) species that experience heat stress Rubisco deactivation is compounded by a reduction in the ability of RCA to remove sugar phosphates from the Rubisco active sites, so bound active sites remain unusable until temperatures fall (Crafts-brandner and Salvucci, 2000). Even at moderate temperatures RCA hexamers begin to aggregate and become non-functional (Salvucci *et al.*, 2001). However, the temperature sensitivity of RCA differs between species and depends on the environments they have evolved in (Salvucci and Crafts-Brandner, 2004*c*). For example, the maximum activity of RCA from plants adapted to cold-to-temperate conditions (e.g., spinach and Antarctic hairgrass) was between 25-27^{o}C, while species from tropical and desert environments had maximum RCA activity between 33 and 35^{o}C (Henderson *et al.*, 2013). Further, RCA from *Gossypium hirsutum*, *Nicotiana tabacum*, and creosote bush (species adapted to warmer environments) denature only after temperatures exceed ~40^{o}C. Scafaro et al. (2016) showed that RCA purified from the leaves of *O. sativa* was able to maintain Rubisco in an active state up to ~34^{o}C. However, RCA from wheat was unable to maintain Rubisco in an active at temperatures exceeding ~25^{o}C (Degen *et al.*, 2020). Because both Rubisco and RCA can be sensitive to heat, and because the sequence diversity of RCA is higher than that of Rubisco, RCA is the prime candidate for improving the thermotolerance of C3 photosynthesis by maintaining the availability of Rubisco active sites.

The number of active Rubisco active sites, and therefore the amount of CO_2 that can be assimilated, in the chloroplast is tightly regulated by balancing the expression of Rubisco and RCA. For example, if the number of active Rubisco active sites declines it is typical to see increases in the expression of RCA (Fukayama *et al.*, 2012, 2018).

In doing so, the cell ensures that the Rubisco active sites are maintained in an active state. Likewise, if the abundance of RCA is reduced in a cell, the expression of Rubisco large and small subunits will increase such that even though the active sites are not maintained there are simply more active sites present (Fukayama *et al.*, 2012, 2018). This interaction between Rubisco and RCA expression patterns presents the possibility of using changes in the stoichiometry of Rubisco and RCA under heat stress to help plants maintain photosynthetic capacity. For example, plants may increase the expression of RCA such that the abundance of total RCA (that is, the abundance of all RCA isoforms, e.g., alpha and beta) increases, or the abundance of total RCA may stay the same, but the ratios of different isoforms may change to favour those that are more heat stable. For example, Scafaro *et al*. (2016) showed that the beta isoform of RCA is more heat stable than the alpha isoform. The net effect of these responses is to increase the total number of active Rubisco active sites in the chloroplast, which can help a plant maintain photosynthesis in the face of short-term heat exposure.

However, under longer term heat stress (as in a heat wave that lasts several days) plants lacking RCA that can maintain Rubisco in an active state should suffer growth and ultimately yield penalties. Historically, the term “heat stress” has been used to describe temperature regimes that deal individually with short-term heat events, with very few studies exploring the impact that long-term heat stress has on the growth and physiology of plants. Jagadish et al. (2021) call for experiments that explore both kinds of temperature regime (both short and long term heat events) in a bid to understand plant responses to heat more generally.

Drought Stress

Drought conditions are associated with declines in plant growth and yield (Pandey and Bhandari, 2009). Drought tolerant plants respond to low-water stress at all levels of organisation. At the leaf level, when water becomes limiting, a series of signal cascades causes reduced stomatal conductance to conserve water (Hu *et al.*, 2017). As a result, gas exchange is limited and CO_2 fixation declines (Lauteri *et al.*, 2014). Further, drought-tolerant rice accessions tend to have lower stomatal density than drought-susceptible accessions (Caine *et al.*, 2019). In roots, cells can alter the expression of genes that lead to the accumulation of osmolytes and thus decrease the water potential of the root tissue (Kim *et al.*, 2020). Drought conditions usually favour those species that either have a well-established root system or that respond to low-water conditions by increasing the depth of their seminal roots (i.e., those the exhibit a plastic response to drought stress; Ahmadi *et al.*, 2014; Sandhu *et al.*, 2016). For example, the quantitative trait locus (QTL) DRO1 encodes an auxin-sensitive gene product that controls growth angle of rice root tips in a way that promotes downward growth (Uga *et al.*, 2013). This 'drilling' behaviour has been shown to increase the yield of commercial rice cultivars by 10% under optimal conditions (Arai-Sanoh *et al.*, 2014), and stabilises yield under moderate drought (Uga *et al.*, 2013). Drought-tolerant rice from the *aus* sub-species have thick roots (i.e., high average root diameter), which is correlated with grain yield under drought conditions (Liao *et al.*, 2022). Thus, deeper root systems allow plants to access more soil moisture and 'avoid' drought for a time, while thicker roots may help droughted plants penetrate drought-hardened soil to gain access to moisture lower in the soil profile (Klein *et al.*, 2020). This may come at the expense of root system surface area – i.e., investment in the fibrous, lateral roots declines. To this end, the kinds of roots that a plant invests in and the architecture of the root system both contribute to the drought tolerance of a species (Lynch, 2022).

Low N Stress

Nitrogen stress is also characterised by reduced plant growth and yield (Ju *et al.*, 2015). Adequate nitrogen is required for the myriad synthesis reactions that occur in plant cells. For example, protein and chlorophyll synthesis reactions are large nitrogen sinks (Tegeder and Masclaux-Daubresse, 2018). One of the largest protein-, and therefore N-, sinks in leaves is the Rubisco enzyme (Staswick, 1997). Thus, it is no surprise that one of the symptoms of nitrogen deficiency is also reduced photosynthetic capacity, which in turn impacts grain production (Tantray et al. 2020). As with drought stress, the type of roots a plant invests in, and the architecture of the root system contribute to low-nitrogen stress tolerance (Lynch, 2019, 2022; Ajmera *et al.*, 2022). Selection for high-yielding rice cultivars in high-input irrigated paddies has also inadvertently selected for shallow rooting systems that take advantage of abundant nutrients in the topsoil (Ahmadi *et al.*, 2014). Conversely, those cultivars selected for high yield in upland, low input (i.e., rainfed with little to no N supplementation) farms tend to have deeper rooting systems that allow them to access water and N deeper in the soil profile. For rice, increasing the depth of the axial root(s) often comes at the expense of fibrous, lateral roots (Xiong *et al.*, 2021). Interestingly, DRO1, which doubles rooting depth in rice (Uga *et al.*, 2013), also increases leaf N, presumably by allowing rice plants to access nitrates deeper in the soil profile (Arai-Sanoh *et al.*, 2014). In spinach, one of the key traits negatively associated with above ground biomass accumulation under low-N conditions is average root diameter. Complementary results are reported for maize – low N reduced the average root diameter of young roots (GAO *et al.*, 2015; Yang *et al.*, 2019). Thus, plants with deeper but thinner roots can access more soil nitrogen and are therefore less likely to be affected by low nitrogen conditions (Awika *et al.*, 2021). However, it is worth noting that average root diameter is positively correlated with NUE in maize (Yang *et al.*, 2019).

Utilising Intraspecific Variation to Improve Domestic Rice Cultivars

The Asian domestic rice (*Oryza sativa*) germplasm is a genetically diverse resource (Wang *et al.*, 2014). The diversity of this germplasm is thought to be the result of both multiple independent domestication events in which different *O. rufipogon* genotypes were used for selection, and the extensive breeding efforts that have been carried out over the past 10,000 years (Londo *et al.*, 2006). These efforts have culminated in a collection of cultivars that are highly specialised to the region in which they were bred. As such, genetic markers can be used to differentiate the Asian rice germplasm into six subgroups - *indica*, *aus*, *aromatic*, temperate *japonica*, tropical *japonica*, and *rayada* (Wang *et al.*, 2014). Regional specialisation involves the selection of traits (whether inadvertent or targeted) that enhance tolerance to/performance in local conditions. In the case of the *aus* subgroup, traits including enhanced flood tolerance, salt tolerance, drought tolerance and heat tolerance have been revealed because this subgroup has been exposed to similar adverse conditions during its natural, and subsequent artificial, evolutionary history (Londo *et al.*, 2006; Casartelli *et al.*, 2018; Chen *et al.*, 2020). These traits can be used to alter the qualities of the final grain product (in the case of the aromatic rices) and to augment the tolerance of susceptible cultivars to several stresses.

The identification of quantitative trait loci (QTL) constitutes a major part of introducing tolerance traits into susceptible cultivars. QTLs are regions of a genome that have been associated with certain phenotypes via statistical analyses. The ultimate goal of QTL analyses is to identify genes that are associated with a desirable trait. QTL analyses have been used by rice biologists to investigate genomic regions associated with drought (Oo *et al.*, 2021), heat (Tian *et al.*, 2022), salinity (Kong *et al.*, 2021), and

low-nitrogen tolerance (Subudhi *et al.*, 2020), amongst others, in diverse *O. sativa* genotypes. The *aus* genotypes are promising for the improvement of tolerance to stresses likely to become more relevant under climate change scenarios. However, the upper limits of stress that these *O. sativa* accessions can tolerate often fall short of those of the wild *Oryza* species. For example, *O. meridionalis* and *O. australiensis* can tolerate temperatures up to 45°C (Scafaro *et al.*, 2012, 2016), while *O. sativa aus* genotypes have been shown to tolerate temperatures up to 41°C (Wahab *et al.*, 2020). Thus, there is a need to consider abiotic tolerances in the wild *Oryza* species (Atwell *et al.*, 2014).

Wild Rice Relatives as a Route to Multi-Stress Tolerance in Rice

The genus *Oryza* is made up of ~27 species that have been classified into 11 genome groups. Thus far, we have considered only those improvements that have been made using genetic variation that occurs within the AA genome clade. In this section we move our attention to the wild *Oryza* species that occupy all 11 genome clades (AA, BB, CC, BBCC, CCDD, EE, FF, GG, HHJJ, HHKK, and KKLL). Each of these species is adapted to different ecological niches and thus represent reservoirs of novel germplasm for the improvement of domestic rice cultivars (Atwell *et al.*, 2014). Given the success that has been realised using the *O. sativa* germplasm variation alone, it is expected that the untapped variation of the wild *Oryza* genomes will provide novel alleles that can be used to improve domestic rice stress tolerance (Gaikwad *et al.*, 2021). Many, if not all, crop species have suffered varying degrees of ‘genetic erosion’ through the intensive breeding processes that have given rise to modern high-yielding cultivars. For example, when compared with wild species, the diversity of the soybean (*Glycine max*) genome is reduced by ~50% and the presence of rare alleles reduced by ~60% (Hyten *et al.*, 2006). Such losses in allelic variation render agriculturally

important species susceptible to environmental stresses. While different cultivars of the same species can show variation in traits associated with tolerance to extreme conditions, it is much more likely that the functional diversity needed to improve modern cultivars will be found in wild species adapted to marginal environments where local adaptation to extreme conditions has occurred (Atwell *et al.*, 2014; Dempewolf and Guarino, 2015). For example, thermotolerant orthologs of RCA have been identified in *O. australiensis* and *O. meridionalis*, wild relatives of rice native to the Northern Territory of Australia (Scafaro *et al.*, 2016).

Due to differences in the ecological niches that the wild *Oryza* species occupy they are predicted to harbour tolerance to different abiotic stresses (Atwell *et al.*, 2014). For example, species from areas of high annual rainfall are likely to possess traits that facilitate survival in flooded conditions, though likely not for drought (Atwell *et al.*, 2014). Thus, if one is interested in using wild *Oryza* species to explore responses to agriculturally relevant abiotic stresses (like heat, drought, and low soil N), one should focus on species whose native range exposes them to such conditions.

Oryza australiensis is a Wild Rice Relative with Multi-Stress Tolerance

Within the genus *Oryza*, *O. australiensis* is remarkable for its uniqueness. The sole member of the EE genome clade, it is the most genetically distant member of the *O. officinalis* genome clade. *O. australiensis* plants co-occur with *O. meridionalis* and *O. rufipogon*, are perennial, and survive the dry season in the form of a rhizome (Henry *et al.*, 2010). The perennial nature of wild rice relatives, like *O. australiensis*, is of immense interest as they can help move rice cultivation towards a more sustainable future (Zhang *et al.*, 2022). *O. australiensis* usually form fewer tillers than domestic rice and have vastly different canopy structures that rival the light interception efficiency of

O. sativa (Taylor, 2015; Henry, 2019). The anatomical properties of *O. australiensis* leaves (e.g., larger mesophyll cells that have more surface area exposed to intracellular air spaces and which have more chloroplasts than domestic rice) make them more photosynthetically efficient than *O. sativa* (Mathan *et al.*, 2021*a*). Unlike in domestic rice, instead of directing the bulk of its photosynthate to developing grains, *O. australiensis* favours the synthesis of structural carbohydrates (Mathan *et al.*, 2021*b*). As a result of these contrasting source-sink relationships, *O. sativa* grain yield is higher than *O. australiensis*, but *O. australiensis* accumulates more biomass (Mathan *et al.*, 2021*b*). The grains of *O. australiensis* are short and red and have been proposed as a healthy food source because they contain slowly digestible starch (Tikapunya *et al.*, 2017; Henry, 2019). Also, the flour has properties desirable for a number of food products (Tikapunya *et al.*, 2017). Given its range, *O. australiensis* was hypothesised to possess traits associated with heat and drought tolerance.

The native range of *O. australiensis* extends across the northern savannahs of Australia, stretching from Western Australia, across the Northern Territory and into Queensland (Fig. 2.1) The harsh environments that *O. australiensis* occurs in are characterised by low soil nutrition, and relatively hot and seasonally dry conditions, especially when compared to the Akita Prefecture – one of the most productive rice growing regions in Japan (Fig. 2.2a & b). Over the course of a year, the temperature at Kununurra (a town in Western Australia, close to Keep River National Park, from which a majority of the O. australiensis seed used for stress tolerance studies originated; Scafaro *et al.*, 2011, 2012, 2016, 2019; Yichie *et al.*, 2018, 2019; Hamzelou *et al.*, 2020) is relatively stable, rarely dropping below 25°C and with a mean maximum of 38°C (Fig. 2.2a). Akita, on the other hand, has much more seasonally defined temperature changes with large temperature ranges (e.g., -2°C to 28°C). The opposite

is true for rainfall, where Kununurra has a strong seasonality to its rainfall, and Akita has relatively consistent rain year-round (though the irrigated rice crops of Akita have access to water through the growing season; Fig. 2.2). Over the course of the year, Akita receives ~1200 mm of rain, while Kununurra receives ~800 mm of rain. Within the growing season (October to March for Kununurra), there is large monthly variation in rainfall supply to *O. australiensis*. For example, in October and November, when *O. australiensis* seeds are germinating and seedlings are developing, the average monthly rainfall at Kununurra is ~21 mm and ~57 mm, respectively, while at the peak of the wet season (February), Kununurra receives ~206 mm of rain. There is also substantial between-year rainfall variability in Kununurra – that is, precipitation is not necessarily predictable from year to year early in the development of *O. australiensis* (Fig. 2.2c & d). Thus, in Kununurra, *O. australiensis* plants experience limited and somewhat variable water availability early in their development and grow on soils that have low soil N (~0.03% soil N; Nunn, 1965; Dixon, 1996), especially compared to traditional rice paddies (up to 0.15% total soil N; Roberts *et al.*, 2011). Taken together, the genetic remoteness and the harsh conditions of its native range make *O. australiensis* a prime candidate for investigating stress tolerance traits, especially those associated with high temperature, low N and drought stress tolerance.

Heat sensitivity in *O. sativa* has been well established and there is substantial evidence indicating superior thermotolerance in *O. australiensis* (Scafaro *et al.*, 2016). Leaf elongation rates of *O. australiensis* were not affected by short term heat exposure (45°C), but *O. sativa* leaves elongated slower at this temperature. Photosynthesis was also stable in *O. australiensis* under heat, though not in *O. sativa*. This photosynthetic thermotolerance has been attributed to the maintenance of Rubisco activity mediated by a thermostable ortholog of RCA – while *O. sativa* RCA was stable up to 36°C, *O.*

australiensis RCA was stable up to 42°C. However, given that heat events are becoming more frequent and have longer duration times, there is a need to understand how *O. sativa* and *O. australiensis* will respond to long term high temperatures as well (Jagadish *et al.*, 2021). Characterising the growth response of these species to long-term heat events alongside more in-depth physiological (e.g., exploring the role of photosynthetic electron transport rates, photorespiration, and respiration in heat tolerance) and molecular analyses (e.g., analysing sugar content, and measuring changes in the abundance of heat-responsive enzymes like RCA) under short-term heat exposure would be a major step forward in understanding the heat tolerance of *O. australiensis*.

Researchers have been hypothesising about drought tolerance in *O. australiensis* for decades (Chatuverdi *et al.*, 1996). Given that *O. australiensis* grows during summer in the tropics of Australia, where rainfall patterns have strong seasonality (i.e., dry winter and wet, monsoonal summer) one may not expect to identify drought tolerance in this species. Further, there are few mentions on the Atlas of Living Australia (Belbin *et al.*, 2021) of this species occurring away from a water source (i.e., depressions in clay-heavy soils on seasonal floodplains, swamps, rivers, creeks, ponds, and even in roadside drains), though some occurrences on dry riverbeds are noted. Thus, *O. australiensis* populations should have access to a ready supply of water throughout the vegetative phase. A notable exception may be during its early growth when rainfall variability is still quite high (Fig. 2.2) – if rainfall is unreliable during the seedling stage, then one might expect to see seedling level drought tolerance. Indeed, seedlings of *O. australiensis* rapidly accumulate proline in response to salt stress (Nguyen *et al.*, 2021), and different populations of *O. australiensis* have been found to have different levels of salt tolerance (Yichie *et al.*, 2018, 2019). Salt and drought stress are different

in nature, though they both exert osmotic stress on plant roots and thus they elicit some shared plant responses (Ma *et al.*, 2020). Further, moderately drought tolerant crosses between *O. sativa* and *O. meridionalis*, a wild rice that co-occurs with *O. australiensis*, have been successfully produced (Gouda *et al.*, 2012). Also, different accessions of *O. rufipogon* (a wild rice relative) have different levels of drought tolerance, while a single accession of *O. officinalis* (the type species for the *O. officinalis* complex of wild rice relatives that *O. australiensis* is part of) demonstrated tolerance to drought late in the vegetative stage (Feng *et al.*, 2012). It may therefore be reasonable to predict that *O. australiensis* possesses some level of drought tolerance.

While tolerance to other stresses has been reported by several authors, drought tolerance has not been extensively investigated for *O. australiensis*. Hamzelou et al. (2020) compared the response of *O. sativa*, *O. glaberrima* and *O. australiensis* to drought conditions and found that the wild rice was able to maintain leaf water potential while the domestic rice species did not. They also report on significant below-ground investment in the wild rice regardless of treatment, which suggests that *O. australiensis* may be primed to respond to water deficit conditions. 'High-yielding' drought-tolerant hybrids of *O. sativa* and *O. meridionalis* had longer roots and higher root volume than 'low-yielding' drought-sensitive hybrids (Gouda *et al.*, 2012). Based on these results it would be interesting to know how the architecture of *O. australiensis* root systems compares to that of *O. sativa* and whether drought elicits changes in root architecture. However, the physiological data presented by Hamzelou et al. (2020) must be interpreted with caution as the plants were likely experiencing N toxicity symptoms (Atwell and Plett, pers. comm.). Further, while the plants were droughted (maintained at 40% field capacity) for seven days, it is important to understand the impact of water

deficit over longer time frames. So, to the best of our knowledge, there are no reliable nor long term investigations into drought tolerance for *O. australiensis*.

There is reasonable cause to suspect tolerance to low nitrogen by *O. australiensis*. The soils of Australia are famously deficient in nutrient elements, including nitrogen (Nunn, 1965; Dixon, 1996). Bui et al. (2013) show that Australian grasslands, such as those that *O. australiensis* is commonly found in, have very low soil nitrogen, nitrogen:phosphorous, and carbon:phosphorous ratios. Further, decreasing soil nitrogen in degraded (i.e., those with elevated soil nitrogen and which therefore have a high proportion of annual species) grassland ecosystems allow native Australian perennial grassland species to re-establish dominance and grow. Perennial species, like *O. australiensis*, can benefit from low soil nitrogen as fast-growing annual species cannot obtain sufficient nitrogen to maintain their otherwise competitive growth (Hacker *et al.*, 2011). In these conditions, it is beneficial for perennial species to have high nitrogen use efficiency (NUE) traits (Ens *et al.*, 2015). This is powerfully demonstrated by the competition between the perennial grasses *Andropogon gayanus* Kunth (invasive) and *Alloteropsis semialata* (native) in the north Australian grasslands (Ens *et al.*, 2015). *A. gayanus* dominates what were once biodiverse habitats because it has high NUE in the dry season while the native perennial species senesces to avoid the deleterious impacts of seasonal water deficits (Ens *et al.*, 2015). However, *A. semialata* persists in the ecosystem because it has higher NUE during the wet season and so can compete with *A. gayanus*. Given that *O. australiensis* is a perennial species that occurs on low N soils, it may have NUE traits and therefore tolerance to low N soils like *A. semialata*, though no studies have investigated this. Likewise, because rainfall is so variable during the first two months (October and November) of the development of *O. australiensis* and the soils have low N content, the wild rice seedlings may have

evolved tolerance mechanisms for these co-occurring stresses. Again, this has not been investigated for *O. australiensis*.

Thus, there is considerable interest in creating crosses between *O. sativa* and *O. australiensis*, in transforming *O. sativa* with *O. australiensis* genes for improved growth and/or yield in the face of stress, and in the domestication of *O. australiensis* itself. The genetic distance between *O. sativa* and *O. australiensis* and the different timings of meiotic events makes generating crosses between these species a challenge (Shastry and Rao, 1961). Indeed, F1 hybrids of *O. sativa* and *O. australiensis*, which were made in the hopes of transferring resistance to brown plant hopper (BPH) and bacterial blight (BB), were male sterile (Multani *et al.*, 1994), which has been observed in independent experiments (Abbasi, 1999). Colchicine-derived autotetraploid *O. sativa* lines (AAAA) have been used as the female parent in crosses with diploid *O. australiensis* (EE) as the male parent (Multani *et al.*, 1994). Backcrossing the resulting triploid (AAE) progeny, after embryo rescue, with diploid *O. sativa* (AA) yielded diploid (AE) and aneuploid (AAE) offspring. Resistance to BPH and one strain of pathogen causing BB was only observed in 0.67% and 0.17% of the progeny of the diploid crosses, respectively. Further, seed set in *O. sativa* and *O. australiensis* hybrids is low (0.5-3.8%; Sitch, 1990), which is undesirable.

However, resistance to BPH can be achieved through genetic transformations using resistance genes from *O. australiensis* (Ji *et al.*, 2016; Kumar *et al.*, 2020) thereby avoiding complications from making hybrids. This provides hope that similar transformations can be made for improving tolerance to abiotic stress. There are few documented examples of transformation events utilising *O. australiensis* genes in *O. sativa* that improve tolerance to abiotic stresses. Transformation of *O. sativa* with *O.*

australiensis RCA has already proven useful for bolstering the yield of *O. sativa* when grown at 45°C (Scafaro *et al.*, 2018). Similarly, wheat RCA was made more thermotolerant by introducing transcripts with key mutations identified in *O. australiensis* RCA (Scafaro *et al.*, 2019). The identification of useful wild rice genes and gene variants (i.e., alleles) that can be used to improve tolerance traits in domestic rice would be aided by the existence of high-quality nuclear genome assemblies for *O. australiensis*. This would also allow for the investigation of useful promoter region sequences, which also play a role in abiotic tolerance phenotypes.

Current restrictions associated with genetically modified food crops and poor public perception may render such transformants useless beyond a proof-of-concept stage of investigation and hamper efforts to get the grain to market. As such, there have been discussions surrounding the domestication of the wild rice species native to Australia (both *O. meridionalis* and *O. australiensis*). To achieve this goal, *O. australiensis* would need to undergo substantial breeding to improve several agronomic traits including but not limited to its severe lodging and grain shattering habits. The genes governing these phenotypes in domestic rice are known (e.g., Ishikawa *et al.*, 2022), but if a breeding program for the improvement of *O. australiensis* is to be established breeders will need to develop markers specific to *O. australiensis* given the substantial genetic distance between the *Oryza* species. Such efforts would be made easier if a genome of high quality exists. Also, while proof-of-concept experiments with transformed *O. australiensis* plants (e.g., *O. australiensis* lines with the *sh4* locus mutation for reduced shattering) would be required to validate gene targets in this species there is currently no known way to reliably induce callus formation nor regenerate transformed *O. australiensis* plantlets (Shimizu-Sato *et al.*, 2020).

Even though there has been international interest in *O. australiensis* for decades, there is a paucity of research into the extent of abiotic stresses that this species is capable of tolerating, and the mechanisms underlying the reported and/or hypothesised tolerances are poorly understood. Furthermore, there has been little-to-no progress on developing tissue culture protocols for *O. australiensis*, nor the assembly of a contiguous nuclear genome for *O. australiensis* since the inception of the International Oryza Map Alignment Project (IOMAP; Jacquemin *et al.*, 2013).

The Role of Wild Species' Genomes in Improving Crop Stress Tolerance

In order to take full advantage of *O. australiensis* as a sources of novel genes it is important to understand the species in a genomic context (Jacquemin *et al.*, 2013). Recently, Stein *et al*. (Stein *et al.*, 2018) compiled reference genomes of 13 *Oryza* species. This study considered ten species/accessions of rice that belong to the AA genome subclade of *Oryza*, one BB genome species, and one FF genome species. In doing so, Stein *et al.* (Stein *et al.*, 2018) identified new genetic information that may be of use in the stabilisation of rice yield. The identification of novel genetic variation within the AA subclade alone makes the prospect of allele mining in other, more diverse, genomic clades desirable. For example, *O. australiensis* is the sole member of the EE genome subclade and has a large genome (~965 Mbp) compared to domestic rice (~430 Mbp; Jacquemin *et al.*, 2013). As discussed above, given its distribution in extreme environments, populations of this species have evolved tolerances to various abiotic stresses. For example, an accession of *O. australiensis* from the Keep River region has been reported to be thermotolerant (Scafaro *et al.*, 2016), however in comparison to other *O. australiensis* accessions from salt marsh areas, the Keep River genotype had low tolerance to salt (Yichie *et al.*, 2018). The size of the *O. australiensis* genome and occurrence of the species in conditions that would be sub-optimal for

domestic rice (Atwell *et al.*, 2014) make it desirable to fully sequence the nuclear genome(s) of the phenotypically diverse genotypes of this species. However, progress in this regard has been slow largely because of the size and complexity of the *O. australiensis* genome.

When the IOMAP called for the *O. australiensis* nuclear genome to be sequenced in 2003, short read sequencing technology was the predominant tool available (Jacquemin *et al.*, 2013). While short reads are highly accurate, they are not ideal as a base resource to assemble regions that are high in repeat content. They are also of limited use for the detection of large structural variations in genomes. Given that the difference in genome size between *O. sativa* and *O. australiensis* is largely driven by differences in long terminal repeat (LTR) retrotransposon (RT) content – the *O. sativa* genome contains ~10% LTR-RT sequence, while *O. australiensis* contains ~65% LTR-RT sequence – rather than differences in gene family expansion and/or contraction (Piegu *et al.*, 2006), it is no wonder that a highly contiguous assembly was unresolved at the time of this research study. As sequencing technology has developed, researchers now have access to third generation long-read sequencing technology (e.g., Oxford Nanopore Technology (ONT), and Pacific Biosciences (PacBio)). Long reads allow for genome assemblers to bridge the gaps between highly repetitive genomic regions, such as those high in LTR-RT elements, and thus increase the contiguity and quality of an assembly. A long-read assembly of *O. australiensis* is yet to be produced.

Chloroplasts contain their own (not-autonomous) genomes. Recently, it was found that the chloroplast genomes of the *Oryza* genus have played major roles in the diversification of the genus by facilitating adaptation to ecological niches (Gao *et al.*,

2019). Therefore, if we are to understand *O. australiensis*, or any species, in a genomic context we must also work to understand the contribution of the chloroplast genome to its multi-stress tolerance. As with the nuclear genome, this will involve mining diverse intraspecific accessions with different tolerances to stresses for chloroplastic genome variation. Unlike the nuclear genome, however, a reference chloroplast genome assembly exists for *O. australiensis* (Wu and Ge, 2016). While chloroplast genomes are relatively small (~120-160 kbp) a substantial portion (~30%) is made up of two large, inverted repeat (IR) sequences and these are nestled between large and small single copy regions (LSC and SSC regions, respectively; see, Chang *et al.*, 2021). The orientation of the LSC and SSC regions can change and so are described as existing in a 'flip-flop' configuration (Stein *et al.*, 1986; Kim *et al.*, 2019). As discussed above, repeat content and structural variations can make assemblies with short reads challenging. So, there is a case to be made for the use of long reads in the assembly of chloroplast genomes as well as nuclear genomes (Wang *et al.*, 2018; Wang and Lanfear, 2019). Since modern long-read sequencing technologies allow even small research groups to collect vast swathes of genomic data, the challenge now becomes deriving information from genomic data from diverse accessions. This will also be made easier by the availability of high quality *de novo* chloroplast genome assemblies.

Amazingly, both the nuclear and chloroplast (and mitochondrial) genome can be assembled from long reads collected from a single sample, if the sample is derived from photosynthetically active tissue. Plant cells contain many thousands of chloroplasts, each with their own genome copies (Palmer and Stein, 1986; Wolfe *et al.*, 1987; Morley and Nielsen, 2016; Dobrogojski *et al.*, 2020). When a read library is prepared, it will contain reads of nuclear and chloroplast origin. This provides researchers with the opportunity to add value to their sequencing projects if the

chloroplast reads are isolated from the total read pool. There are currently a number of tools available for the assembly of chloroplast genomes (Freudenthal *et al.*, 2020). However, many, though not all, rely on the use of short reads, and all present significant barriers to entry for researchers with limited training in bioinformatics. The few available tools making use of long reads are often difficult to install and run to completion, which may limit their use particularly amongst those who do not specialise in bioinformatics. This is something that needs to be addressed if inter- and intraspecific chloroplast genomic diversity is to be interrogated.

Concluding Remarks

There is a clear need to stabilise or improve rice crop yields in the coming decades. However, these efforts will be hampered by the challenges imposed by a changing climate (e.g., heat wave events, changing water regimes) and a growing world population (e.g., the need to produce more food with fewer inputs). While the AA genome clade that domestic rice species (*O. sativa* and *O. glaberrima*) occupy has been used for exploring tolerance traits, the genetic diversity of non-AA genomes offers a unique opportunity to explore responses to highly variable and extreme environments. So, rice breeders and breeding programs will need to extend their search for tolerance and/or resource use efficiency traits to previously unexplored germplasm. Part of this search will need to include wild *Oryza* species, which are adapted to several ecological niches around the globe and will require detailed understanding of tolerances and their mechanisms. Amongst the *Oryza* species, *O. australiensis* presents as an interesting candidate with tolerance to multiple agriculturally relevant abiotic stresses. Heat and salt tolerance has been established in this species, with substantial phenotypic diversity occurring between populations. Our understanding of how *O. australiensis* tolerates heat stress and whether this

tolerance extends to long term heat events (i.e., up to several weeks) is poor. Further, drought and low soil N tolerance has also been hypothesised given the native range of this species, though these hypotheses have not been explored properly. Thus, there remains much we do not understand about the stresses *O. australiensis* is able to tolerate, and the extent of its stress tolerance. Once the extent and mechanisms of its tolerance are determined, it will be possible to uncover genes associated with these traits. Efforts such as this will be aided greatly by the existence of a high-quality long-read based nuclear genome assembly. The genome of *O. australiensis* has been a challenge to assemble due to the high repeat content of its genome. Thus, sequencing efforts since 2003 have failed to produce a highly contiguous, high-quality assembly. This is something that can be tackled with modern day long-read sequencing technologies, and will facilitate the discovery of genes and, importantly, genomic variation (e.g., promoter variation, and structural variation) associated with stress tolerance. Variation in chloroplast genome sequences is also known to contribute to the adaptation of *Oryza* species to diverse environments, and so a sound understanding of chloroplast variants will also be required to understand stress tolerance mechanisms. This will require the availability of easy to install, easy to use tools for the rapid assembly of chloroplast genomes *de novo* that take into consideration the skill level of novice bioinformaticians. However, thus far no such tools exist, and this may greatly slow progress in understanding the role of the chloroplast in stress tolerance.

Thesis Aims

This thesis aims to provide a deeper understanding of the extent and mechanisms of abiotic stress tolerances of *O. australiensis*, and to provide genomic resources that will

remove some of the hurdles faced by plant biologists and breeders in taking advantage of this wild rice relative. Specifically, there were four main aims:

1. Establish a deeper understanding of the mechanisms that govern heat tolerance in *O. australiensis*, and determine whether *O. australiensis* could tolerate long term heat stress (Chapter 3),
2. Address the hypotheses of drought and low N tolerance in *O. australiensis* and explore how it and *O. sativa* (domestic rice) differed in response when both stresses were imposed concurrently (Chapter 4),
3. Generate the first long-read reference-quality draft nuclear genome assembly for *O. australiensis* and use it to explore the diversity of *O. australiensis* samples from different populations (Chapter 5), and
4. Generate an easy to install, easy to use tool for the assembly of chloroplast genomes from long-read data, accessible to even novice bioinformaticians (Chapter 6).

References

Abbasi FM. 1999. Cytogenetic and molecular analyses of Oryza sativa x O. australiensis Domin. and O. sativa L. x O. brachyantha A. Chev. et Roehr derivatives.

Ahmadi N, Audebert A, Bennett MJ, et al. 2014. The roots of future rice harvests. Rice **7**, 29.

Ajmera I, Henry A, Radanielson AM, Klein SP, Ianevski A, Bennett MJ, Band LR, Lynch JP. 2022. Integrated root phenotypes for improved rice performance under low nitrogen availability. Plant, Cell & Environment **45**, 805–822.

Alvarez-Ponce D, Sabater-Muñoz B, Toft C, Ruiz-González MX, Fares MA. 2016. Essentiality Is a Strong Determinant of Protein Rates of Evolution during Mutation Accumulation Experiments in Escherichia coli. Genome Biology and Evolution **8**, 2914–2927.

Andersson I, Backlund A. 2008. Structure and function of Rubisco. Plant Physiology and Biochemistry **46**, 275–291.

Arai-Sanoh Y, Takai T, Yoshinaga S, Nakano H, Kojima M, Sakakibara H, Kondo M, Uga Y. 2014. Deep rooting conferred by DEEPER ROOTING 1 enhances rice yield in paddy fields. Scientific Reports **4**, 5563.

Ashraf M, Harris PJC. 2013. Photosynthesis under stressful environments: An overview. Photosynthetica **51**, 163–190.

Atwell BJ, Wang H, Scafaro AP. 2014. Could abiotic stress tolerance in wild relatives of rice be used to improve Oryza sativa? Plant Science **215–216**, 48–58.

Awika HO, Mishra AK, Gill H, DiPiazza J, Avila CA, Joshi V. 2021. Selection of nitrogen responsive root architectural traits in spinach using machine learning and genetic correlations. Scientific Reports **11**, 9536.

Belbin L, Wallis E, Hobern D, Zerger A. 2021. The Atlas of Living Australia: History, current state and future directions. Biodiversity Data Journal **9**, e65023.

Bibi I, Ali Q, Niazi NK, Kanwal N, Ahsan M, Ali F. 2015. Multivariate Analysis of Grain Yield and Its Attributing Traits in Different Maize Hybrids Grown under Heat and Drought Stress. Scientifica **2015**, 1–6.

Bita CE, Gerats T. 2013. Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. Frontiers in Plant Science **4**, 1–18.

Bui EN, Henderson BL. 2013. C:N:P stoichiometry in Australian soils with respect to vegetation and environmental factors. Plant and Soil **373**, 553–568.

Caine RS, Yin X, Sloan J, et al. 2019. Rice with reduced stomatal density conserves water and has improved drought tolerance under future climate conditions. New Phytologist **221**, 371–384.

Carmo-Silva E, Scales JC, Madgwick PJ, Parry MAJ. 2015. Optimizing Rubisco and its regulation for greater resource use efficiency. Plant, Cell and Environment **38**, 1817–1832.

Casartelli A, Riewe D, Hubberten HM, Altmann T, Hoefgen R, Heuer S. 2018. Exploring traditional aus-type rice for metabolites conferring drought tolerance. Rice **11**.

Chang H, Zhang L, Xie H, Liu J, Xi Z, Xu X. 2021. The Conservation of Chloroplast Genome Structure and Improved Resolution of Infrafamilial Relationships of Crassulaceae. Frontiers in Plant Science **12**.

Chatuverdi GS, Ram PC, Singh AK, Ram P, Ingram KT, Singh BB, Singh RK, Singh VP. 1996. Carbohydrate status of rainfed lowland rices in relation to submergence, drought and shade tolerance. Physiology of Stress Tolerance In Rice. Lucknow, India: Int. Rice Res. Inst., 103–122.

Chen J, Mueller V. 2018. Coastal climate change, soil salinity and human migration in Bangladesh. Nature Climate Change **8**, 981–985.

Chen C, Norton GJ, Price AH. 2020. Genome-Wide Association Mapping for Salt Tolerance of Rice Seedlings Grown in Hydroponic and Soil Systems Using the Bengal and Assam Aus Panel. Frontiers in Plant Science **11**.

Costa MVJ Da, Ramegowda Y, Ramegowda V, Karaba NN, Sreeman SM, Udayakumar M. 2021. Combined Drought and Heat Stress in Rice: Responses, Phenotyping and Strategies to Improve Tolerance. Rice Science **28**, 233–242.

Crafts-brandner SJ, Salvucci ME. 2000. Rubisco activase constrains the photosynthetic potential of leaves at high temperature and CO2.

Debnath S, Mishra A, Mailapalli DR, Raghuwanshi NS, Sridhar V. 2020. Assessment of rice yield gap under a changing climate in India. Journal of Water and Climate Change **12**, 1245–1267.

Degen GE, Worrall D, Carmo-Silva E. 2020. An isoleucine residue acts as a thermal and regulatory switch in wheat Rubisco activase. The Plant Journal **103**, 742–751.

Dempewolf H, Guarino L. 2015. Reaching back through the domestication bottleneck: Tapping wild plant biodiversity for crop improvement. Acta Horticulturae **1101**, 165–168.

Dixon JC. 1996. Soils of the Weaber Plain, East Kimberley, Western Australia. Perth, Western Australia: Department of Primary Industries and Regional Development.

Dobrogojski J, Adamiec M, Luciński R. 2020. The chloroplast genome: a review. Acta Physiologiae Plantarum **42**, 98.

Dusenge ME, Duarte AG, Way DA. 2019. Plant carbon metabolism and climate change: elevated CO2 and temperature impacts on photosynthesis, photorespiration and respiration. New Phytologist **221**, 32–49.

Ens E, Hutley LB, Rossiter-Rachor NA, Douglas MM, Setterfield SA. 2015. Resource-use efficiency explains grassy weed invasion in a low-resource savanna in north Australia. Frontiers in Plant Science **6**.

Feng F, Xu X, Du X, Tong H, Luo L, Mei H. 2012. Assessment of drought resistance among wild rice accessions using a protocol based on single-tiller propagation and PVC-tube cultivation. Australian Journal of Crop Science **6**, 1204–1211.

Freudenthal JA, Pfaff S, Terhoeven N, Korte A, Ankenbrand MJ, Förster F. 2020. A systematic comparison of chloroplast genome assembly tools. Genome Biology **21**, 254.

Frich P, Alexander L V., Della-Marta P, Gleason B, Haylock M, Tank Klein AMG, Peterson T. 2002. Observed coherent changes in climatic extremes during the second half of the twentieth century. Climate Research **19**, 193–212.

Fukayama H, Mizumoto A, Ueguchi C, Katsunuma J, Morita R, Sasayama D, Hatanaka T, Azuma T. 2018. Expression level of Rubisco activase negatively correlates with Rubisco content in transgenic rice. Photosynthesis Research **137**, 465–474.

Fukayama H, Ueguchi C, Nishikawa K, Katoh N, Ishikawa C, Masumoto C, Hatanaka T, Misoo S. 2012. Overexpression of Rubisco activase decreases the photosynthetic CO2 assimilation rate by reducing Rubisco content in rice leaves. Plant and Cell Physiology **53**, 976–986.

Gaikwad KB, Singh N, Kaur P, Rani S, Babu H P, Singh K. 2021. Deployment of wild relatives for genetic improvement in rice (Oryza sativa). Plant Breeding **140**, 23–52.

GAO KUN, CHEN F, YUAN L, ZHANG F, MI G. 2015. A comprehensive analysis of root morphological changes and nitrogen allocation in maize in response to low nitrogen stress. Plant, Cell & Environment **38**, 740–750.

Gao L-Z, Liu Y-L, Zhang D, et al. 2019. Evolution of Oryza chloroplast genomes promoted adaptation to diverse ecological habitats. Communications Biology **2**, 1–13.

Gouda PK, Kumar Varma CM, Saikumar S, Kiran B, Shenoy V, Shashidhar HE. 2012. Direct Selection for Grain Yield under Moisture Stress in Oryza sativa cv. IR58025B × Oryza meridionalis Population. Crop Science **52**, 644–653.

Gutteridge S, Gatenby A. 1995. Rubisco Synthesis, Assembly, Mechanism, and Regulation. The Plant Cell **7**, 809–819.

Hacker RB, Toole ID, Melville GJ, Hacker RB, Toole ID, Melville GJ. 2011. Effects of nitrogen and phosphorus on vegetation dynamics of a degraded native grassland in semi-arid south-eastern Australia. The Rangeland Journal **33**, 87–97.

Hamzelou S, Kamath KS, Masoomi-Aladizgeh F, Johnsen MM, Atwell BJ, Haynes PA. 2020. Wild and Cultivated Species of Rice Have Distinctive Proteomic Responses to Drought. International Journal of Molecular Sciences **21**.

Henderson JN, Hazra S, Dunkle AM, Salvucci ME, Wachter RM. 2013. Biophysical characterization of higher plant Rubisco activase. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics **1834**, 87–97.

Henry RJ. 2019. Australian Wild Rice Populations: A Key Resource for Global Food Security. Frontiers in Plant Science **10**, 1354.

Henry RJ, Rice N, Waters DLE, Kasem S, Ishikawa R, Hao Y, Dillon S, Crayn D, Wing R, Vaughan D. 2010. Australian Oryza: utility and conservation. Rice **3**, 235–241.

Horváth I, Glatz A, Nakamoto H, Mishkind ML, Munnik T, Saidi Y, Goloubinoff P, Harwood JL, Vigh L. 2012. Heat shock response in photosynthetic organisms: membrane and lipid connections. Progress in Lipid Research **51**, 208–220.

Hsu C-H, Chiang AWT, Hwang M-J, Liao B-Y. 2016. Proteins with Highly Evolvable Domain Architectures Are Nonessential but Highly Retained. Molecular Biology and Evolution **33**, 1219–1230.

Hu Y, Wu Q, Peng Z, et al. 2017. Silencing of OsGRXS17 in rice improves drought stress tolerance by modulating ROS accumulation and stomatal closure. Scientific reports **7**, 15950–15950.

Hyten DL, Song Q, Zhu Y, Choi I-Y, Nelson RL, Costa JM, Specht JE, Shoemaker RC, Cregan PB. 2006. Impacts of genetic bottlenecks on soybean genome diversity. Proceedings of the National Academy of Sciences **103**, 16666–16671.

IPCC. 2019. Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (PR Shukla, J Skea, E Calvo Beundia, et al., Eds.).

IPCC. 2022. Climate Change 2022: Impacts, Adaptation, and Vulnerability. Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, .

Ishikawa R, Castillo CC, Htun TM, et al. 2022. A stepwise route to domesticate rice by controlling seed shattering and panicle shape. Proceedings of the National Academy of Sciences **119**, e2121692119.

Ishikawa C, Hatanaka T, Misoo S, Miyake C, Fukayama H. 2011. Functional Incorporation of Sorghum Small Subunit Increases the Catalytic Turnover Rate of Rubisco in Transgenic Rice. Plant Physiology **156**, 1603–1611.

Jacquemin J, Bhatia D, Singh K, Wing RA. 2013. The International Oryza Map Alignment Project: development of a genus-wide comparative genomics platform to help solve the 9 billion-people question. Current Opinion in Plant Biology **16**, 147–156.

Jagadish SVK, Way DA, Sharkey TD. 2021. Plant heat stress: Concepts directing future research. Plant Cell and Environment **44**, 1992–2005.

Ji H, Kim S-R, Kim Y-H, et al. 2016. Map-based Cloning and Characterization of the BPH18 Gene from Wild Rice Conferring Resistance to Brown Planthopper (BPH) Insect Pest. Scientific Reports **6**, 34376.

Ju C, Buresh RJ, Wang Z, Zhang H, Liu L, Yang J, Zhang J. 2015. Root and shoot traits for rice varieties with higher grain yield and higher nitrogen use efficiency at lower nitrogen rates application. Field Crops Research **175**, 47–55.

Kim Y, Chung YS, Lee E, Tripathi P, Heo S, Kim K-H. 2020. Root Response to Drought Stress in Rice (Oryza sativa L.). International Journal of Molecular Sciences **21**, 1513.

Kim D, Lee J, Choi JW, Yang JH, Hwang I-K, Yoon HS. 2019. Flip-flop organization in the chloroplast genome of Capsosiphon fulvescens (Ulvophyceae, Chlorophyta). Journal of Phycology **55**, 214–223.

Klein SP, Schneider HM, Perkins AC, Brown KM, Lynch JP. 2020. Multiple Integrated Root Phenotypes Are Associated with Improved Drought Tolerance. Plant Physiology **183**, 1011–1025.

Kolb PF, Robberecht R. 1996. High temperature and drought stress effects on survival of Pinus ponderosa seedlings. Tree Physiology **16**, 665–672.

Kong W, Zhang C, Zhang S, Qiang Y, Zhang Y, Zhong H, Li Y. 2021. Uncovering the Novel QTLs and Candidate Genes of Salt Tolerance in Rice with Linkage Mapping, RTM-GWAS, and RNA-seq. Rice **14**, 93.

Kosaki T, Wasano K, Juo ASR. 1989. Multivariate statistical analysis of yield-determining factors. Soil Science and Plant Nutrition **35**, 597–607.

Kumar K, Kaur P, Kishore A, Vikal Y, Singh K, Neelam K. 2020. Recent advances in genomics-assisted breeding of brown planthopper (Nilaparvata lugens) resistance in rice (Oryza sativa). Plant Breeding **139**, 1052–1066.

Lauteri M, Haworth M, Serraj R, Monteverdi MC, Centritto M. 2014. Photosynthetic Diffusional Constraints Affect Yield in Drought Stressed Rice Cultivars during Flowering. PLOS ONE **9**, e109054.

Liao Q, Chebotarov D, Islam MS, et al. 2022. Aus rice root architecture variation contributing to grain yield under drought suggests a key role of nodal root diameter class. Plant, Cell & Environment **45**, 854–870.

Liu W, Yin T, Zhao Y, Wang X, Wang K, Shen Y, Ding Y, Tang S. 2021. Effects of High Temperature on Rice Grain Development and Quality Formation Based on Proteomics Comparative Analysis Under Field Warming . Frontiers in Plant Science **12**.

Londo J, Yu-Chung C, Kuo-Hsiang H, Tzen-Yuh C, Schaal B. 2006. Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. Proceedings of the National Academy of Sciences **103**, 9578–9583.

Lynch JP. 2019. Root phenotypes for improved nutrient capture: an underexploited opportunity for global agriculture. New Phytologist **223**, 548–564.

Lynch JP. 2022. Harnessing root architecture to address global challenges. The Plant Journal **109**, 415–431.

Ma Y, Dias MC, Freitas H. 2020. Drought and Salinity Stress Responses and Microbe-Induced Tolerance in Plants. Frontiers in Plant Science **11**.

Mathan J, Singh A, Jathar V, Ranjan A. 2021a. High photosynthesis rate in two wild rice species is driven by leaf anatomy mediating high Rubisco activity and electron transport rate. Journal of Experimental Botany.

Mathan J, Singh A, Ranjan A. 2021b. Sucrose transport and metabolism control carbon partitioning between stem and grain in rice. Journal of Experimental Botany **72**, 4355–4372.

Maurino VG, Peterhansel C. 2010. Photorespiration: Current status and approaches for metabolic engineering. Current Opinion in Plant Biology **13**, 249–256.

Morita K, Hatanaka T, Misoo S, Fukayama H. 2014. Unusual Small Subunit That Is Not Expressed in Photosynthetic Cells Alters the Catalytic Properties of Rubisco in Rice. Plant Physiology **164**, 69–79.

Morley SA, Nielsen BL. 2016. Chloroplast DNA Copy Number Changes during Plant Development in Organelle DNA Polymerase Mutants. Frontiers in Plant Science **7**, 57.

Mueller-Cajar O. 2017. The Diverse AAA+ Machines that Repair Inhibited Rubisco Active Sites. Frontiers in Molecular Biosciences **4**.

Mueller-Cajar O, Stotz M, Bracher A. 2014. Maintaining photosynthetic CO2 fixation via protein remodelling: The Rubisco activases. Photosynthesis Research **119**, 191–201.

Multani DS, Jena KK, Brar DS, de los Reyes BG, Angeles ER, Khush GS. 1994. Development of monosomic alien addition lines and introgression of genes from Oryza australiensis Domin. to cultivated rice O. sativa L. Theoretical and Applied Genetics **88**, 102–109.

Nguyen HTT, Das Bhowmik S, Long H, Cheng Y, Mundree S, Hoang LTM. 2021. Rapid Accumulation of Proline Enhances Salinity Tolerance in Australian Wild Rice Oryza australiensis Domin. Plants **10**, 2044.

Nunn WM. 1965. Kimberley Research Station progress report, 1964 : introduction. Journal of the Department of Agriculture, Western Australia **6**.

Oo KS, Krishnan SG, Vinod KK, et al. 2021. Molecular Breeding for Improving Productivity of Oryza sativa L. cv. Pusa 44 under Reproductive Stage Drought Stress through Introgression of a Major QTL, qDTY12.1. Genes **12**, 967.

Palmer JD, Stein DB. 1986. Conservation of chloroplast genome structure among vascular plants. Current Genetics **10**, 823–833.

Pandey S, Bhandari H. 2009. Drought: economic costs and research implications. Drought Frontiers in Rice. WORLD SCIENTIFIC, 3–17.

Parry MAJ, Andralojc PJ, Mitchell RAC, Madgwick PJ, Keys AJ. 2003. Manipulation of Rubisco: The amount, activity, function and regulation. Journal of Experimental Botany **54**, 1321–1333.

Parry MAJ, Andralojc PJ, Scales JC, Salvucci ME, Carmo-Silva AE, Alonso H, Whitney SM. 2013. Rubisco activity and regulation as targets for crop improvement. Journal of Experimental Botany **64**, 717–730.

Parry M a. J, Madgwick PJ, Carvalho JFC, Andralojc PJ. 2006. Prospects for increasing photosynthesis by overcoming the limitations of Rubisco. The Journal of Agricultural Science **145**, 31–43.

Peng S, Huang J, Sheehy JE, Laza RC, Visperas RM, Zhong X, Centeno GS, Khush GS, Cassman KG. 2004. Rice yields decline with higher night temperature from global warming. Proceedings of the National Academy of Sciences of the United States of America **101**, 9971–9975.

Piegu B, Guyot R, Picault N, et al. 2006. Doubling genome size without polyploidization: Dynamics of retrotransposition-driven genomic expansions in Oryza australiensis, a wild relative of rice. Genome Research **16**, 1262–1269.

Portis AR. 2003. Rubisco activase – Rubisco's catalytic chaperone. Photosynthesis Research **75**, 11–27.

Powell N, Ji X, Ravash R, Edlington J, Dolferus R. 2012. Yield stability for cereals in a changing climate. Functional Plant Biology **39**, 539–552.

Prins A, Orr DJ, Andralojc PJ, Reynolds MP, Carmo-Silva E, Parry MAJ. 2016. Rubisco catalytic properties of wild and domesticated relatives provide scope for improving wheat photosynthesis. Journal of Experimental Botany **67**, 1827–1838.

Roberts TL, Ross WJ, Norman RJ, Slaton NA, Wilson Jr. CE. 2011. Predicting Nitrogen Fertilizer Needs for Rice in Arkansas Using Alkaline Hydrolyzable-Nitrogen. Soil Science Society of America Journal **75**, 1161–1171.

Salvucci ME, Crafts-Brandner SJ. 2004a. Mechanism for deactivation of Rubisco under moderate heat stress. Physiologia Plantarum **122**, 513–519.

Salvucci ME, Crafts-Brandner SJ. 2004b. Inhibition of photosynthesis by heat stress: the activation state of Rubisco as a limiting factor in photosynthesis. Physiologia Plantarum **120**, 179–186.

Salvucci ME, Crafts-Brandner SJ. 2004c. Relationship between the heat tolerance of photosynthesis and the thermal stability of rubisco activase in plants from contrasting thermal environments. Plant Physiology **134**, 1460–1470.

Salvucci ME, Osteryoung KW, Crafts-Brandner SJ, Vierling E. 2001. Exceptional sensitivity of Rubisco activase to thermal denaturation in vitro and in vivo. Plant Physiology **127**, 1053–1064.

Sandhu N, Raman KA, Torres RO, Audebert A, Dardou A, Kumar A, Henry A. 2016. Rice Root Architectural Plasticity Traits and Genetic Regions for Adaptability to Variable Cultivation and Stress Conditions . Plant Physiology **171**, 2562–2576.

Sarkar RK, Chakraborty K, Chattopadhyay K, Ray S, Panda D, Ismail AM. 2019. Chapter 13 - Responses of Rice to Individual and Combined Stresses of Flooding and Salinity. In: Hasanuzzaman M,, In: Fujita M,, In: Nahar K,, In: Biswas JKBT-A in RR for AST, eds. Woodhead Publishing, 281–297.

Scafaro AP, Atwell BJ, Muylaert S, Reusel BV, Ruiz GA, Rie JV, Gallé A. 2018. A thermotolerant variant of Rubisco activase from a wild relative improves growth and seed yield in rice under heat stress. Frontiers in Plant Science **9**, 1–11.

Scafaro AP, Bautsoens N, den Boer B, Van Rie J, Galle A. 2019. A conserved sequence from heat-adapted species improves Rubisco activase thermostability in wheat. Plant Physiology **181**, 43–54.

Scafaro AP, Gallé A, Van Rie J, Carmo-Silva E, Salvucci ME, Atwell BJ. 2016. Heat tolerance in a wild Oryza species is attributed to maintenance of Rubisco activation by a thermally stable Rubisco activase ortholog. New Phytologist **211**, 899–911.

Scafaro AP, Von Caemmerer S, Evans JR, Atwell BJ. 2011. Temperature response of mesophyll conductance in cultivated and wild Oryza species with contrasting mesophyll cell wall thickness. Plant, Cell & Environment **34**, 1999–2008.

Scafaro AP, Yamori W, Carmo-Silva AE, Salvucci ME, von Caemmerer S, Atwell BJ. 2012. Rubisco activity is associated with photosynthetic thermotolerance in a wild rice (Oryza meridionalis). Physiologia Plantarum **146**, 99–109.

Shahid SA, Zaman M, Heng L. 2018. Soil Salinity: Historical Perspectives and a World Overview of the Problem BT - Guideline for Salinity Assessment, Mitigation and Adaptation Using Nuclear and Related Techniques. In: Zaman M,, In: Shahid SA,, In: Heng L, eds. Cham: Springer International Publishing, 43–53.

Sharkey TD, Zhang R. 2010. High Temperature Effects on Electron and Proton Circuits of Photosynthesis. Journal of Integrative Plant Biology **52**, 712–722.

Shastry SVS, Rao DRR. 1961. Timing imbalance in the meiosis of the F1 hybrid Oryza sativa × O. australiensis. Genetics Research **2**, 373–383.

Shimizu-Sato S, Tsuda K, Nosaka-Takahashi M, Suzuki T, Ono S, Ta KN, Yoshida Y, Nonomura K-I, Sato Y. 2020. Agrobacterium-Mediated Genetic Transformation of Wild Oryza Species Using Immature Embryos. Rice **13**, 33.

Sitch LA. 1990. Incompatibility Barriers Operating in Crosses of Oryza sativa with Related Species and Genera. In: Gustafson JP, ed. Stadler Genetics Symposia Series. Gene Manipulation in Plant Improvement II: 19th Stadler Genetics Symposium. Boston, MA: Springer US, 77–93.

Staswick P. 1997. The occurrence and gene expression of vegetative storage proteins and a Rubisco Complex Protein in several perennial soybean species1. Journal of Experimental Botany **48**, 2031–2036.

Stein DB, Palmer JD, Thompson WF. 1986. Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus Osmunda. Current Genetics **10**, 835–841.

Stein JC, Yu Y, Copetti D, et al. 2018. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. Nature Genetics **50**, 285–296.

Studer RA, Christin P-A, Williams MA, Orengo CA. 2014. Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. Proceedings of the National Academy of Sciences **111**, 2223–2228.

Subudhi PK, Garcia RS, Coronejo S, Tapia R. 2020. Comparative Transcriptomics of Rice Genotypes with Contrasting Responses to Nitrogen Stress Reveals Genes Influencing Nitrogen Uptake through the Regulation of Root Architecture. International Journal of Molecular Sciences **21**, 5759.

Taylor JW. 2015. Australian wild oryza species: growth, canopy structure and atmospheric CO2 effects.

Tcherkez GGB, Farquhar GD, Andrews TJ. 2006. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. Proceedings of the National Academy of Sciences of the United States of America **103**, 7246–7251.

Tegeder M, Masclaux-Daubresse C. 2018. Source and sink mechanisms of nitrogen transport and use. New Phytologist **217**, 35–53.

Tian T, Chen L, Ai Y, He H. 2022. Selection of Candidate Genes Conferring Blast Resistance and Heat Tolerance in Rice through Integration of Meta-QTLs and RNA-Seq. Genes **13**, 224.

Tikapunya T, Zou W, Yu W, Powell PO, Fox GP, Furtado A, Henry RJ, Gilbert RG. 2017. Molecular structures and properties of starches of Australian wild rice. Carbohydrate polymers **172**, 213–222.

Tilman D, Balzer C, Hill J, Befort BL. 2011. Global food demand and the sustainable intensification of agriculture. Proceedings of the National Academy of Sciences **108**, 20260–20264.

Tripathi A, Tripathi DK, Chauhan DK, Kumar N, Singh GS. 2016. Paradigms of climate change impacts on some major food sources of the world: A review on current knowledge and future prospects. Agriculture, Ecosystems and Environment **216**, 356–373.

Uga Y, Sugimoto K, Ogawa S, et al. 2013. Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. Nature Genetics **45**, 1097–1102.

Wachter RM, Henderson JN. 2015. Photosynthesis: Rubisco rescue. Nature Plants **1**, 1–2.

Wahab MMS, Akkareddy S, Shanthi P, Latha P. 2020. Identification of differentially expressed genes under heat stress conditions in rice (Oryza sativa L.). Molecular Biology Reports **47**, 1935–1948.

Walker BJ, VanLoocke A, Bernacchi CJ, Ort DR. 2016. The costs of photorespiration to food production now and in the future. Annual Review of Plant Biology **67**, 107–129.

Wang W, Lanfear R. 2019. Long-Reads Reveal That the Chloroplast Genome Exists in Two Distinct Versions in Most Plants. Genome Biology and Evolution **11**, 3372–3381.

Wang W, Schalamun M, Morales-Suarez A, Kainer D, Schwessinger B, Lanfear R. 2018. Assembly of chloroplast genomes with long- and short-read data: a comparison of approaches using Eucalyptus pauciflora as a test case. BMC Genomics **19**, 977.

Wang C-H, Zheng X-M, Xu Q, Yuan X-P, Huang L, Zhou H-F, Wei X-H, Ge S. 2014. Genetic diversity and classification of Oryza sativa with emphasis on Chinese rice germplasm. Heredity **112**, 489–496.

Wei H, Liu J, Wang Y, Huang N, Zhang X, Wang L, Zhang J, Tu J, Zhong X. 2013. A Dominant Major Locus in Chromosome 9 of Rice (Oryza sativa L.) Confers Tolerance to 48 °C High Temperature at Seedling Stage . Journal of Heredity **104**, 287–294.

Whitney SM, Kane HJ, Houtz RL, Sharwood RE. 2009. Rubisco Oligomers Composed of Linked Small and Large Subunits Assemble in Tobacco Plastids and Have Higher Affinities for CO2 and O2. Plant Physiology **149**, 1887–1895.

Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proceedings of the National Academy of Sciences **84**, 9054–9058.

Wu Z, Ge S. 2016. The whole chloroplast genome of wild rice (Oryza australiensis). Mitochondrial DNA. Part A, DNA mapping, sequencing, and analysis **27**, 1062–1063.

Xiong Q, Hu J, Wei H, Zhang H, Zhu J. 2021. Relationship between Plant Roots, Rhizosphere Microorganisms, and Nitrogen and Its Special Focus on Rice. Agriculture **11**.

Yang JT, Schneider HM, Brown KM, Lynch JP. 2019. Genotypic variation and nitrogen stress effects on root anatomy in maize are node specific. Journal of Experimental Botany **70**, 5311–5325.

Yichie Y, Brien C, Berger B, Roberts TH, Atwell BJ. 2018. Salinity tolerance in Australian wild Oryza species varies widely and matches that observed in O. sativa. Rice **11**, 1–14.

Yichie Y, Hasan MT, Tobias PA, Pascovici D, Goold HD, Van Sluyter SC, Roberts TH, Atwell BJ. 2019. Salt-treated roots of Oryza australiensis seedlings are enriched with proteins involved in energetics and transport. Proteomics **19**, 1900175–1900175.

Yuan S, Stuart AM, Laborte AG, et al. 2022. Southeast Asia must narrow down the yield gap to continue to be a major rice bowl. Nature Food **3**, 217–226.

Zhang S, Huang G, Zhang Y, et al. 2022. Sustained productivity and agronomic potential of perennial rice. Nature Sustainability, 1–11.

Zhao C, Liu B, Piao S, et al. 2017. Temperature increase reduces global yields of major crops in four independent estimates. Proceedings of the National Academy of Sciences **114**, 9326–9331.

Figures

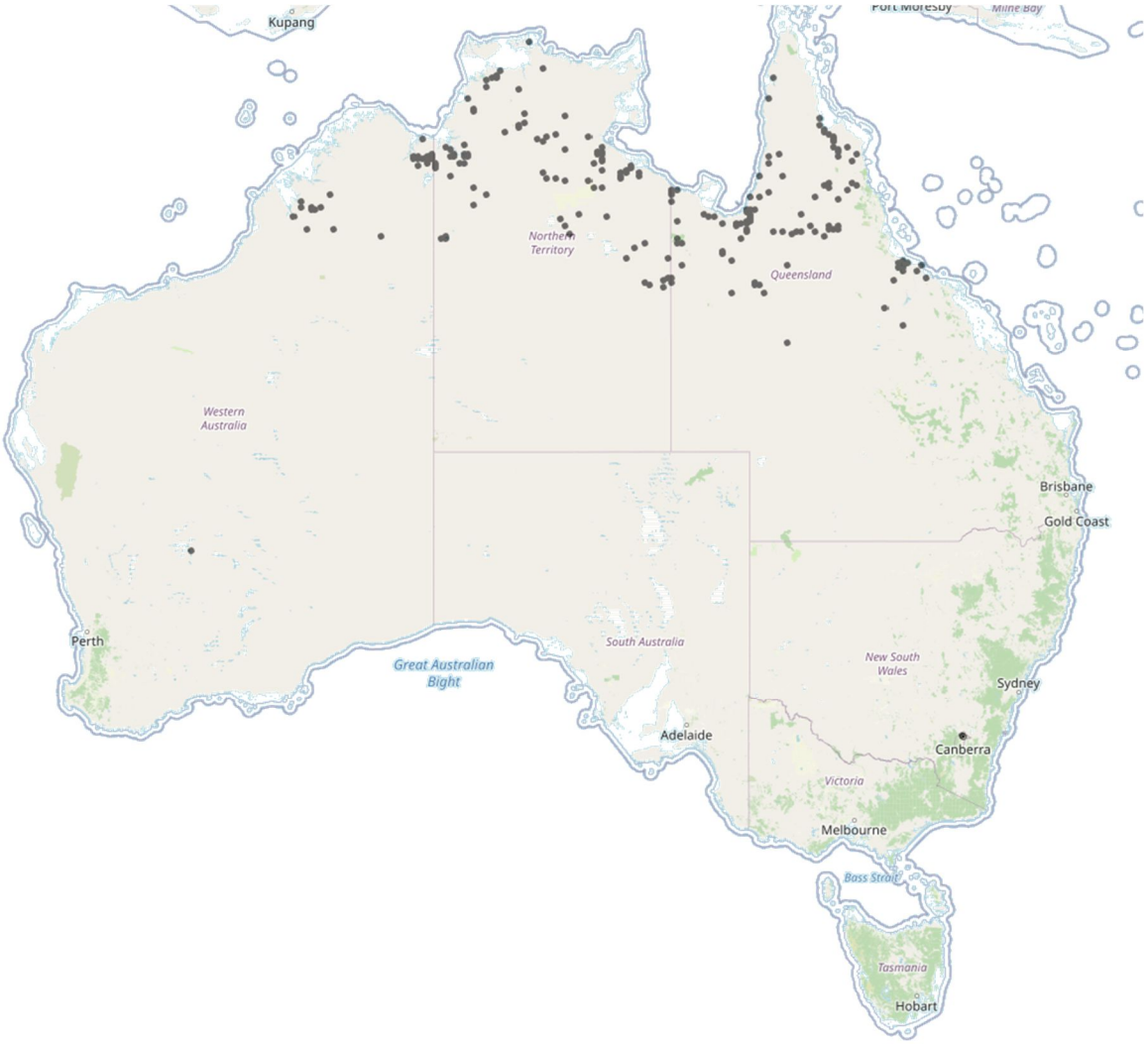


Figure 2.1. Distribution of *Oryza australiensis* (black dots) in Australia. This map was obtained from The Atlas of Living Australia (Belbin *et al.*, 2021). The occurrence data in southern central Western Australia and Canberra are herbarium records and so should be ignored.

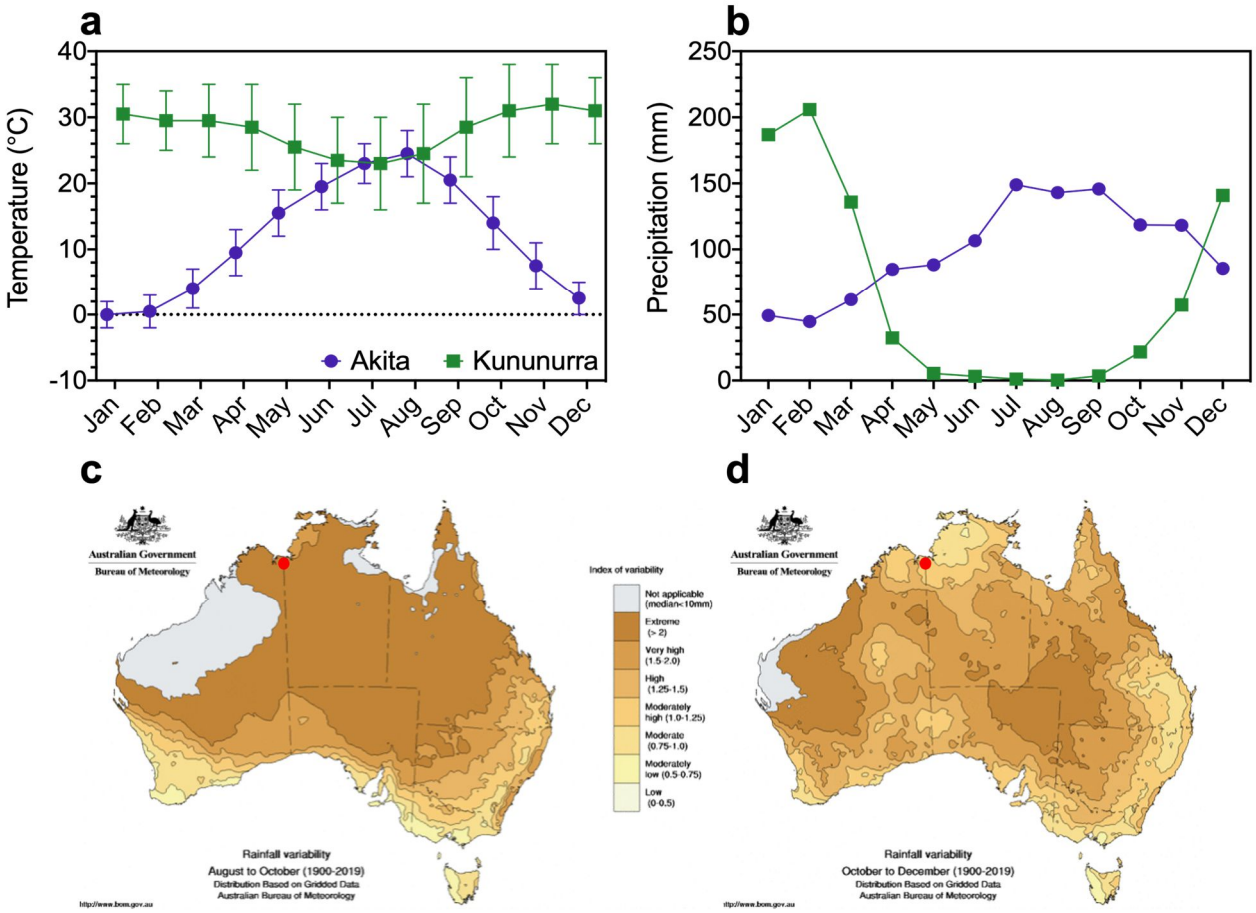
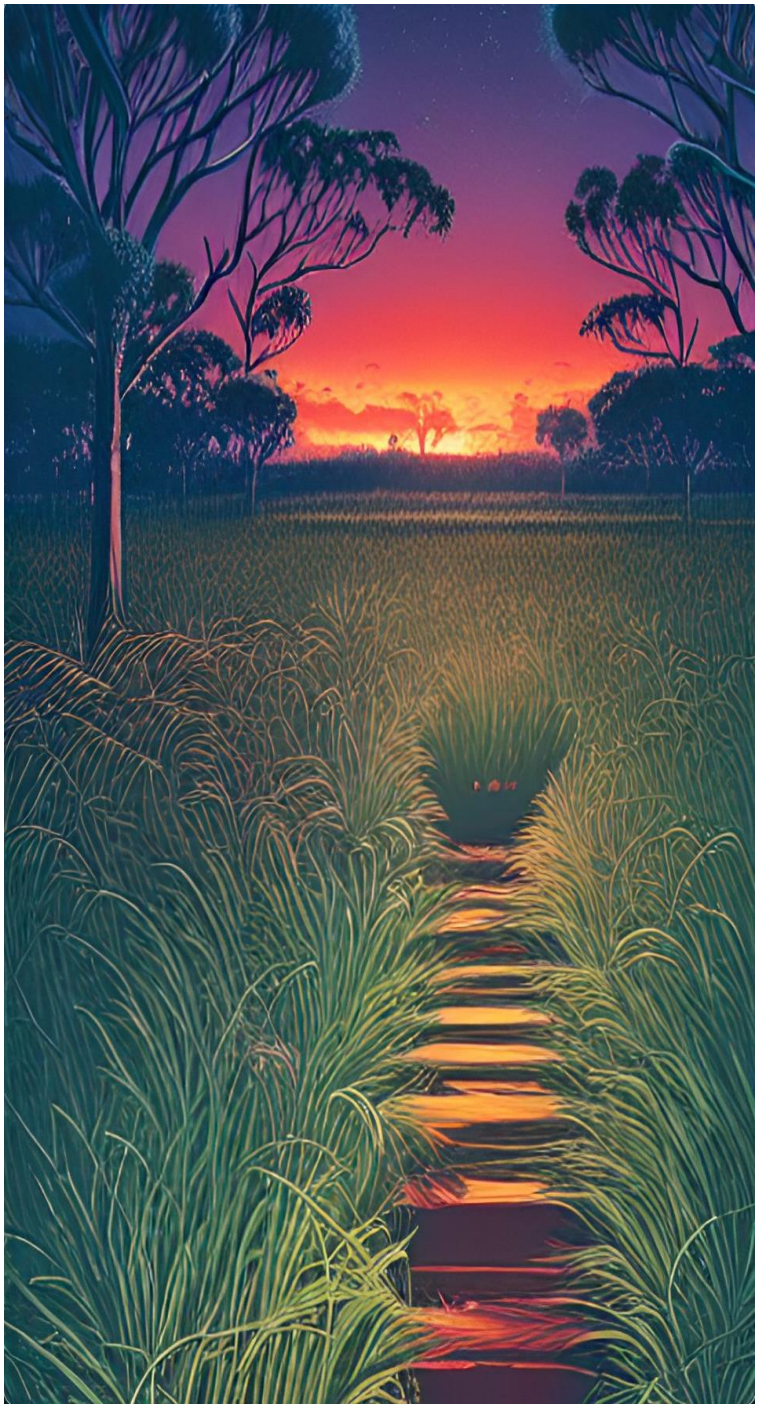


Figure 2.2. Temperature (a) and precipitation (b-d) patterns for Akita (Japan) and Kununurra (Australia). (c) shows the variability in rainfall for Australia between August (the dry season) and October (start of the wet season, when *O. australiensis* is thought to germinate). (d) shows the variability in rainfall for Australia between October and December. The 'error bars' on the temperature plot indicate monthly highs (top of the error bar) and monthly lows (bottom of the error bar). Data were obtained from WeatherSpark, whom obtained their data from the Akita Airport (AA) weather station and the East Kimberley Regional Airport (EKRA) weather station for Akita and Kununurra, respectively. The data from AA are the mean values from 1973 to 2022 for each month, while those from EKRA are the mean values from 1990 to 2022. While these data represent means, no measure of variability (e.g., standard deviation) could be obtained from WeatherSpark. The growing season for *O. sativa* in Akita is from April (transplanting in May) to September. While the phenology of *O. australiensis* is not known exactly, it is known that each plant survives the dry season (May – September) as rhizomes, so the growing season must be between October and April.

Chapter 3

Photosynthetic Traits of Australian Wild Rice (*Oryza australiensis*) Confer Tolerance to Extreme Daytime Temperatures



Wombo Prompt

"Oryza *australiensis* grassland heatwave Australia"

Wombo Theme

"Comic"

Statement of Authorship

Title of Paper	Photosynthetic Traits of Australian Wild Rice (*Oryza australiensis*) Confer Tolerance to Extreme Daytime Temperatures
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Phillips, A.L., Scafaro, A.P. & Atwell, B.J. Photosynthetic traits of Australian wild rice (*Oryza australiensis*) confer tolerance to extreme daytime temperatures. *Plant Mol Biol* (2022). https://doi.org/10.1007/s11103-021-01210-3

Principal Author

Name of Principal Author (Candidate)	Aaron L. Phillips (ALP)		
Contribution to the Paper	ALP contributed to the study conception and design. Material preparation, data collection and analysis were performed by ALP. Additional data analyses were suggested and performed by APS and ALP. The first draft of the manuscript was written by ALP.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	1st September 2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr. Andrew Scafaro		
Contribution to the Paper	Suggested and helped perform additional analyses (RACiR curves). Helped edit the manuscript.		
Signature		Date	8th September 2022

Name of Co-Author	Prof. Brian Atwell		
Contribution to the Paper	Project conceptualisation and design. Manuscript preparation.		
Signature		Date	9th September 2022

Please cut and paste additional co-author panels here as required.

Journal:
Plant Molecular Biology

Published:
2022, https://doi.org/10.1007/s11103-021-01210-3

Title:
Photosynthetic traits of Australian wild rice (*Oryza australiensis*) confer tolerance to extreme daytime temperatures

Authors: Aaron L. Phillips[1, 3, 4], Andrew P. Scafaro[2], Brian J. Atwell[3*]

Author Affiliations:
[1]University of Adelaide, Waite Research Institute and School of Agriculture, Food, and Wine, Adelaide, South Australia, Australia; [2]The Australian National University, ARC Centre of Excellence in Plant Energy Biology, Research School of Biology, Canberra, Australian Capital Territory, Australia; [3]Macquarie University, Department of Biological Sciences, Sydney, New South Wales, Australia; [4]ARC Centre of Excellence in Plant Energy Biology, School of Agriculture, Food, and Wine, The University of Adelaide, Adelaide, South Australia, Australia

Declarations:

Funding
Not Applicable

Conflicts of interest/Competing interests
Not Applicable

Availability of data and material
All data available upon request to the corresponding author.

Code availability
All code and data available at:
https://github.com/aaronphillips7493/Oa_heat_paper_2021

Author contributions
Aaron Phillips and Brian Atwell contributed to the study conception and design. Material preparation, data collection and analysis were performed by Aaron Phillips. Additional data analyses were suggested and performed by Andrew Scafaro. The first draft of the manuscript was written by Aaron Phillips and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Abstract

Above ~32°C, impaired photosynthesis compromises the productivity of rice. We compared leaf tissues from heat-tolerant wild rice (*Oryza australiensis*) with temperate-adapted *O*. *sativa* after sustained exposure to heat, as well as diurnal heat shock. Leaf elongation and shoot biomass in *O*. *australiensis* were unimpaired at 45°C, and soluble sugar concentrations trebled during 10 h of a 45°C shock treatment. By contrast, 45°C slowed growth strongly in *O. sativa*. Chloroplastic CO_2 concentrations eliminated CO_2 supply to chloroplasts as the basis of differential heat tolerance. This directed our attention to carboxylation and the abundance of the heat-sensitive chaperone Rubisco activase (Rca) in each species. Surprisingly, *O*. *australiensis* leaves at 45°C had 50% *less* Rca per unit Rubisco, even though CO_2 assimilation was *faster* than at 30°C. By contrast, Rca per unit Rubisco doubled in *O. sativa* at 45°C while CO_2 assimilation was slower, reflecting its inferior Rca thermostability. Plants grown at 45°C were simultaneously exposed to 700 ppm CO_2 to enhance the CO_2 supply to Rubisco. Growth at 45°C responded to CO_2 enrichment in *O. australiensis* but not *O. sativa*, reflecting more robust carboxylation capacity and thermal tolerance in the wild rice relative.

Key Words:

Photosynthesis, *Oryza sativa*, *Oryza australiensis*, Heat tolerance, Rubisco, Rubisco activase

Introduction

Climate models predict increasing mean maximum temperatures globally, along with more frequent and intense heatwaves (IPCC 2019). Over the past century, the number of days of extreme heat in northern Australia has risen (CSIRO 2018). As average daily temperatures rise and spasmodic heatwaves become more frequent in agricultural systems, plant productivity declines (Dhir 2018), especially in cropping regions at low latitudes (e.g. India, Africa and Australia; Mahlstein et al. 2011; Jagadish et al. 2012). For example, wheat yield is expected to fall by 6% for every 1^{o}C increase in average daily temperature (Asseng et al. 2015), with losses of up to 10% per 1°C possible in other crops (Dhir 2018). Similarly, in rice, grain yield declines by 10% for every 1°C increase in mean temperature during the dry season (Peng et al. 2004). Heat stress therefore threatens food security (Tilman et al. 2011; Fischer et al. 2014). Despite this knowledge, there is limited research into the impact that long-term heat exposure has on plant physiology, with most studies focussing on transient exposure to high temperatures.

Heat stress affects whole-plant productivity through its impact on carbon metabolism. Extended periods of above-optimal temperatures suppress net carbon gain by impairing photosynthesis relative to dark respiration, reducing whole-plant growth and yield (Wahid et al. 2007; Campbell et al. 2007). Typically, respiration increases with rising temperatures, while photosynthesis declines because it is particularly sensitive to heat (Tjoelker 2018). For example, reduced growth of sugarcane at high temperature has been ascribed to lower rates of carbon assimilation and faster respiration (Ebrahim et al. 1998; Wahid et al. 2007; Gomathi et al. 2013). Similarly, the growth rates of domesticated rice (*Oryza sativa*) (Scafaro et al. 2016) and cotton (*Gossypium hirsutum*) (Reddy et al. 1992) leaves were

reduced once air temperatures rose above about 35°C. While respiration acclimates to changes in ambient temperature (Atkin and Tjoelker 2003), acclimation of photosynthesis to heat is less widely accepted (Chi et al. 2013; Benomar et al. 2018), although newly formed rice leaves (cv. IR64) appeared to acclimate to 40^{o}C days (Rashid et al. 2020). Nonetheless, the limits of thermotolerance may soon be reached, even in warm-climate species (e.g. rice), necessitating the identification of traits that can be used to increase tolerance to high temperature (Atwell et al. 2014; Fischer et al. 2014). *A priori*, the best candidate species are crop congeners collected from the hottest biomes. Hence, this study investigates an endemic Australian wild species of rice (*O*. *australiensis*) that has evolved in northern Australia. Scafaro et al (2016) reported faster instantaneous leaf elongation rates linked to photosynthetic biochemistry during short bursts of heat shock in this species. The mechanistic basis of tolerance to *sustained* high daytime temperatures throughout vegetative development in this heat-tolerant species remains unknown and is addressed in this paper.

Plants adapted to savannahs and deserts have evolved a range of mechanisms to tolerate extreme heat, ranging from development and specialised leaf morphology and physiology (e.g. stomatal responses) (Wright et al. 2017) through to the accumulation of polymorphisms in heat-prone enzymes (e.g. for increased stability at high temperature) and changes in gene expression. This is well-documented in the case of heat shock proteins and their transcription factors (Scharf et al. 2012). Another mechanism for adaptation to heat (> 30°C) is a more thermally stable Rubisco activase (Rca), a AAA^{+} protein that hydrolyses ATP while removing inhibitory sugar phosphates from the active sites of Rubisco, facilitating CO_2 fixation (Portis 2003). Species that express high-temperature-stable isoforms of Rca might minimise impairment of photosynthesis during heat (Feller et al. 1998; Salvucci and

Crafts-Brandner 2004; Scafaro et al. 2019). Regulation of *in vivo* Rca activity and abundance is also likely to be important, for example increasing the expression of endogenous Rubisco and Rca in tandem rescues heat-sensitive *O. sativa* from the deleterious effects of heat on photosynthesis and biomass accumulation (Qu et al. 2021). Despite this, the amount of endogenous Rca protein has not been explored as a pathway to photosynthetic thermotolerance in domestic and wild rice species adapted to a range of thermal regimes.

To understand the effects of extreme temperatures on carbon assimilation and utilisation, we raised atmospheric CO_2 concentrations and measured growth and photosynthetic rates of heat-stressed (long- and short-term stress) *O. sativa* (domestic) and *O. australiensis* (wild) plants. We hypothesised that *O. australiensis* should withstand high daytime temperatures through a more robust photosynthetic metabolism. To assess the heat-resilience of photosynthesis, we used A:C_i curves, point measurements of assimilation rates and leaf elongation rates (LERs). Temperature effects on respiration, photorespiration and electron transport capacity were also characterised. Soluble sugar status of leaves and biomass accumulation in shoots and roots over four weeks are reported to establish the relationship between steady-state sugar levels and growth. Finally, we used quantitative proteomics to determine whether there were temperature-dependent changes in total Rca abundance between *O. australiensis* and *O. sativa*. This study presents empirical evidence in support of the superior heat tolerance in *O. australiensis*. We provide further support for photosynthesis being more robust in the wild species under heat and ascribe tolerance to a species-specific response of Rca expression. We also rule out other biochemical/biophysical processes such as respiration and electron transport rate as factors that explain the thermotolerance of *O. australiensis*.

Methods

Plant material

Oryza sativa ssp. *japonica* cv. Amaroo and the wild tropical species *O. australiensis* (sourced from Keep River, in the northernmost region of Western Australia; 15°58′S, 129°03′E) were germinated at 36°C overnight and sown in 1.65-L pots. The soil, a fine-textured krasnozem (sourced locally from Robertson, NSW, Australia), was kept moist at all times by keeping the pots in trays of water. Plants were watered once a week with a commercial water-soluble fertiliser at a concentration of 1 g L^{-1} (AQUASOL, Yates, Australia). Seedling leaves were sprayed once with dilute, freshly prepared ferrous sulphate solution to prevent iron deficiency.

Heat-shock Experiments

In *heat-shock* experiments, plants (n = 3) were grown for six weeks in a glasshouse under a day/night temperature regime of 30/22°C, with a midday photosynthetic photon flux density (PPFD) of 1106 ± 13 µmol quanta m^{-2} s^{-1} and a 12-h light period. Unless specified, all measurements and samples were taken over the course of one day. Plants were transferred to large growth chambers (model no. PG. 15. 18.9.TD+C [rt].3X1000.2X400.R; Thermoline Scientific Equipment, NSW, Australia) on the day of measurement. Two hours prior to the photoperiod, chambers were set to 30°C. At the beginning of the photoperiod one cabinet was set to 45°C (a second cabinet was maintained at 30°C to act as a control). Unless specified otherwise, the relative humidity in all growth chambers was maintained between 60-70% during the photoperiod.

Shoot and root dry biomass in heat-acclimated plants

Seeds of *O*. *sativa* and *O*. *australiensis* were germinated and maintained in glasshouse conditions as above for one week. After 7 d, *O. sativa* and *O. australiensis* seedlings (n = 20) were harvested for baseline measurements of shoot and root fresh and dry mass. Seedlings were exposed to 25, 35 or 45°C for four weeks in growth chambers and harvested weekly (n = 6).

To explore the effect of CO_2 fertilisation on *O. australiensis* and *O. sativa*, six-week-old plants (n_{30oC} = 16; n_{45oC} = 4) were maintained in 700 ppm CO_2, and 30 or 45°C during the light period. Shoot and root dry mass were determined when the plants were 12 weeks old. CO_2 levels were set and maintained using a custom-made CO_2 solenoid (The Canary Company, Lane Cove, Australia).

Leaf elongation rates - short- and long-term response to heat

Seedling growth was tracked through measurements of LER using a HR4000 Linear Variable Displacement Transducer (LVDT) with data logged every 3 min by the program VuGrowth ver. 1.0 (Applied Measurement, Oakleigh, Vic). Three experiments were conducted for the measurement of LERs.

In the first experiment (heat shock), six-week-old plants (n = 3) at the four- to five-leaf stage were moved to a growth cabinet (containing the LVDT apparatus) set to the conditions described above. The youngest, fully expanded leaf of each plant was measured for 12 h and the same eight plants (four per species) were used throughout the measurement period. Heat shock (45°C) was imposed as described above. Leaf growth was measured for 8 h at 45°C, and for 2 h at 30°C. LERs were calculated using increase in leaf lengths for each 1-h period. At 1, 5 and 10 h into the photoperiod leaves from these plants were used for point measurements of photosynthesis, sugar and Rca analysis.

In the second experiment, plants (*n* = 5 - 8) grown at 25, 35, or 45°C for four weeks were moved sequentially into the LVDT chamber. Measurement of LER commenced one hour into the photoperiod. Leaf elongation was measured for 8 h (8:00 - 16:00) and LERs were calculated as the rate of elongation over that time.

In the third experiment, the CO_2 level in the LVDT chamber was increased to 700 ppm. Plants (*n* = 4) of *O. australiensis* and *O. sativa*, which had been grown at 30 or 45°C and 700 ppm CO_2 for six weeks, were moved (independently) to the high-CO_2 LVDT chamber for LER measurements and measured under their respective growth conditions. Measurements were made as above and were replicated three times (total *n* = 12).

Sugar extraction and determination

Leaves of *O. sativa* and *O. australiensis* grown at 30°C or exposed to 45°C as a shock treatment were used for sugar analysis (see above). One fully expanded, healthy leaf (~200 mg) was excised from plants at each time point (1, 5 and 10 h into the light period; *n* = 3), snap-frozen in liquid nitrogen and stored at -80°C. Soluble sugar content was determined following the anthrone method (Yemm and Willis 1954) using glucose standards (25, 50, 75 and 100 mg glucose into 1000 mL H_2O).

Gas exchange – Rapid A:C*i Response (RACiR) curves, Laisk curves and point measurements*

Oryza sativa and *O. australiensis* used for the generation of RACiR curves were grown in glasshouses (as described above) for seven weeks. Plants (*n* = 4 - 6) were acclimated to the desired conditions in growth cabinets (Model BDR16; Conviron, Manitoba, Canada) set to 30 or 45°C and 400 ppm or 700 ppm CO_2 for two weeks prior to measurement (plants were nine weeks old when measured). Plants grown at

30°C were exposed to 45°C following their measurement in the steady state to form the heat shock group.

RACiR curves were made following the procedure of Stinziano *et al.* (2017) and Stinziano *et al.* (2019) on a LI-6800 (Li-Cor, Lincoln, USA) with modifications. Briefly, to measure CO_2 assimilation (A_n), conditions in the IRGA head were set to match the conditions of the growth chamber (block temperature of 30 or 45°C; reference CO_2 level of 400 or 700 ppm), with a flow rate of 500 µmol air s^{-1}, relative humidity of ~70%, a fan speed of 10,000 rpm, with a photosynthetic photon flux density (PPFD) of 1500 µmol quanta m^{-2} s^{-1}. For curve generation, reference CO_2 was reduced to 10 µmol mol^{-1} air and increased to 1010 µmol mol^{-1} air over a 10-min period (ramping = 100 µmol CO_2 mol^{-1} air min^{-1}). Measurements were taken by the LI-6800 every two seconds. All RACiR curves were corrected with empty chamber measurements. Leaves used for RACiR measurements were also used for Laisk curves and fluorescence measurements.

RACiR curves were initially converted from intercellular CO_2 partial pressure (C_i) to chloroplastic CO_2 partial pressure (C_c) using mesophyll conductance (g_m) and its temperature response for *O. sativa* and *O. australiensis* as reported by Scafaro *et al.* (2016). The subsequent A-C_c (where A is the CO_2 assimilation rate measured by the LI-6800) curves were analysed using the standard C_3 photosynthesis model (Farquhar et al. 1980),

$$A_c = V_{cmax} \times \frac{C_c - \Gamma^*}{C_c + K_{air}} - R_{light}$$

$$A_j = J_{max} \times \frac{C_c - \Gamma^*}{4C_c + 8\Gamma^*} - R_{light}$$

where A_c (carboxylation-limited CO_2 assimilation rate) was fit to C_c partial pressures in the linear phase (corresponding to a mean C_c value of 38 Pa) and A_j (electron-transport limited CO_2 assimilation rate) fit to C_c partial pressures above this point. V_{cmax} (maximum rate of carboxylation) and J_{max} (maximum rate of electron transport) were iteratively fit using a non-linear least-squares fit (R statistical). The Michaelis-Menten coefficient of Rubisco in air (K_{air}) and its temperature response (value at 25°C = 30.5 kPa; activation energy = 60.5 kJ mol^{-1}) came from previous *in vitro* measurements in *O. sativa* (Hermida-Carrera et al. 2016). We assumed no difference in Rubisco kinetics between the two species. Respiration in the light (R_{light}) and the CO_2 compensation point in the absence of mitochondrial respiration (Γ^*) were those measured by the Laisk method (Laisk 1977) (see below).

Laisk curves (n = 5 - 6) were generated with a LI-6800 (Li-Cor, Lincoln, USA). The curves were made with a C_i range of 50, 75, 100 and 120 ppm at three different light levels (PPFD = 100, 200 and 400 µmol quanta m^{-2} s^{-1}). The intersection of the three lines was used to estimate R_{light} and Γ^*.

Further, point measurements of gas exchange were made on *O. sativa* and *O. australiensis* using a LI-6400 (Li-Cor, Lincoln, USA) at saturating light levels (1500 µmol quanta m^{-2} s^{-1}). All measurements were made 1, 5 and 10 h into the photoperiod on plants (n = 3) acclimated to 30°C growth chambers (see above). Heat shock was imposed as described above. The CO_2 in the reference chamber was set to 400 µmol CO_2 mol^{-1} air and the block temperature was set to 30 or 45°C.

Chlorophyll fluorescence - dark and light adapted leaves

Following the RACiR and Laisk curves, the same leaves (n = 5 - 6) were used for determining chlorophyll fluorescence parameters as per instructions (see Licor Support guidelines; accessed September 2018). Leaves of interest were wrapped in

aluminium foil and the lights in the growth cabinets were switched off. The plants were kept in the dark for 15 h (5:00 pm - 8:00 am) before dark-adapted fluorescence measurements were made. Foil was removed from each leaf in the dark and the leaf was inserted into the chamber of a LI-6800 fitted with a fluorescence head. Minimal fluorescence (F_o) and maximal fluorescence (F_m) were measured with a rectangular flash of actinic light. Variable fluorescence (F_v; estimated from F_o and F_m) was used to calculate the maximal efficiency of photosystem II (PSII; F_v/F_m).

Dark-adapted leaves were then exposed to 1500 µmol quanta m^{-2} s^{-1} for 1 h to induce light acclimation. A Multiphase Flash™ (MPF) of actinic light was used in combination with a 'dark pulse' flash (i.e. a pulse of infrared light) to estimate F_s and F_m'. MPF is a method for more accurately determining F_m' in light-adapted leaves, while the dark pulse allows for the calculation of F_o' (the minimal fluorescence in the dark of a light-adapted leaf; Loriaux et al 2013; Avenson & Saathoof 2018). These parameters were used to calculate the photochemical yield of PSII (Φ_{PSII}) and the efficiency of PSII energy harvesting in the 'open' (oxidised) state (F_v'/F_m'). Electron transport rates (ETRs) were derived from these values by the LI-6800.

Quantifying relative abundance of RbcL and Rca by tandem mass spectrometry (MS/MS)

Oryza sativa and *O. australiensis* leaves (~200 mg; n = 3) sampled 1, 5 and 10 h into the photoperiod (see above) were ground to a fine powder in liquid nitrogen. Leaf powder was washed in ice-cold 10% TCA/0.07% DTT in acetone and incubated at -20°C for 1 h. Following centrifugation at 4000 × g for 45 min, supernatant was discarded and the pellet was washed with 0.07% DTT in acetone and stored at -20°C overnight. Samples were centrifuged at 4000 × g for 45 min at 5°C and the supernatant was discarded. Following resuspension in 0.07% DTT in acetone,

samples were incubated at -20°C for 1 h. These washes were repeated twice. Pellets were air-dried, suspended in extraction buffer (250 mM TEAB/1.5% SDS), and stored at -20°C overnight. Pellets (kept on ice) were homogenised six times with a Precellys24 tissue homogeniser (Bertin Instruments, Montigny-le-Bretonneux, France) set to 6500 rpm for 20 s. Homogenised samples were centrifuged at 4000 × *g* for 1 h at 10°C. Total protein concentration of the supernatant was determined using a Direct Detect infra-red spectrometer (Merck Millipore, Darmstadt, Germany).

100 µL of 250 mM TEAB was added to 20 µg of sample. DTT was added to a final concentration of 10 mM. This mixture was incubated at 60°C for 1 h. IAA was added to a final concentration of 20 mM, and samples were incubated at room temperature for 1 h in the dark. Trypsin (0.1 µg μL^{-1}) was added in a 1:50 (trypsin:sample) ratio and incubated overnight at 37°C. Digested samples were dried in a SpeediVac and resuspended in 250 mM TEAB. Pierce Detergent Removal spin columns (Thermo Scientific, IL, USA) were used to remove residual SDS. Samples were dried again in a SpeediVac and stored at -20°C. Prior to MS/MS analysis, samples were resuspended in loading buffer (2% acetonitrile (ACN); 0.1% formic acid).

Each sample (10 µl containing 3 µg of peptides) was injected onto a peptide trap (Bruker peptide Captrap) for pre-concentration and desalted with 0.1% formic acid, 2% ACN, at 10 µL min^{-1} for 5 min. The peptide trap was then switched into line with the analytical column. Peptides were eluted from the column using linear solvent gradients, with steps, from mobile phase A (0.1% formic acid): mobile phase B (99.9% ACN/0.1% formic acid) (98:2) to (90:10) for 10 min, then to (65:35), at 600 nL min^{-1} over a 78-min period. After peptide elution, the column was cleaned with 95% buffer B for 10 min and then equilibrated with 98% buffer A for 20 min before the next

sample injection. The reverse phase nanoLC eluent was subject to positive ion nanoflow electrospray analysis in an Information Dependent Acquisition (IDA) mode.

In the IDA mode, a TOF-MS survey scan was acquired (m/z 350 - 1500, 0.25 s), with the ten most intense multiply charged ions (2+ - 5+; counts > 150) in the survey scan sequentially subjected to MS/MS analysis. MS/MS spectra were accumulated for 50 ms in the mass range m/z 100 – 1500 with rolling collision energy.

Data analysis

Data analysis was done in R (R Core Team 2017). All data, unless specified otherwise, were analysed by two- or three-way ANOVAs and Tukey HSD or Duncan's Multiple Range tests. Figures were produced using R and GraphPad Prism (version 9).

The data collection and processing for mass spectrometry was conducted by the Australian Proteomics Analysis Facility (APAF; Macquarie University, Sydney, Australia). The LC-MS/MS data of the IDA runs of each sample were searched using ProteinPilot (v4.2; AB Sciex) in 'thorough' mode. MGF files (peak list) were exported, submitted to Mascot (Matrix Science, UK) and searched against NCBInr *O. sativa* (rice) database for Rubisco activase (Rca). Relative abundances for Rca and the large subunit of Rubisco (RbcL) were determined across samples using emPAI values, which were normalised against emPAI values of the total protein pools in *O. sativa* and *O. australiensis* (see Ishihama et al. (2005) and Zhu et al. (2010) for details on the emPAI method).

Results

Sequential destructive harvests

There was a significant three-way interaction between harvest time, temperature and species in shoot dry weight ($p < 0.0001$; Fig 3.1). Follow-up two-way ANOVAs for each harvest time revealed that shoot dry weights of the two species first diverged after 35 d of exposure to 45°C ($p < 0.0001$). *Oryza sativa* had significantly reduced shoot growth at 45°C compared with plants growing at 25 and 35°C. By contrast, *O. australiensis* shoots were significantly lighter at 25 than at 35 and 45°C. Similar trends were seen in the root dry weight (Supplementary Figure 3.8), demonstrating slowest growth of *O. australiensis* was consistently at 25°C whereas in *O. sativa*, plants grew slowest at 45°C. The impact of the 45°C treatment on the growth of *O. sativa* shoots became severe after 28 d. Similarly, differences in root mass were only detected in both species following 21 d exposure to treatment conditions.

Leaf elongation rates - short term vs. long term response to temperature

A two-way repeated-measures ANOVA was conducted to determine the effect of heat shock on LERs in *O. sativa* and *O. australiensis* over a 17-h period (Fig 3.2A). There was a significant interaction between species and time of day ($p < 0.01$), indicating that the two species responded to heat differentially throughout the day. For example, LER at 11:00 was ~3.17 mm h^{-1} (*O. sativa*) and ~3.76 mm h^{-1} (*O. australiensis*); not significantly different. However, after 8 h of heat treatment (15:00), there was a significant ~28% reduction in LER in *O. sativa*, while a non-significant increase was registered in *O. australiensis* (~4% increase). LERs of the two species remained significantly different from this point until the end of the photoperiod.

After long-term exposure to 35 and 45°C, LERs in *O. australiensis* but not *O. sativa* increased ($p < 0.01$; Fig 3.2B). This interaction was driven by a significant increase in LER for *O. australiensis* when grown at 35 or 45°C compared with 25°C ($p < 0.01$) but this was not seen for *O. sativa*.

Soluble sugars

A significant three-way interaction showed that the effect of temperature on leaf sugar levels was different for the two species depending on the time of day (Fig 3.3; $p = 0.036$). Follow up two-way ANOVAs showed that the interaction between species and temperature ($p = 0.03$) changed throughout the day. That is to say, the effect of temperature on sugar accumulation was different for each species, and this relationship changed throughout the day. This reflects the large increase in soluble sugar concentration in *O. australiensis* leaves at 45°C 10 h into the photoperiod ($p = 0.02$), indicating an accumulation of sugar in spite of the high temperature. In contrast, the effect of heat on leaf sugar levels was the same for both species 1 h and 5 h into the photoperiod. After 10 h at 45°C the concentration of sugars in *O. australiensis* leaves equalled that of *O. sativa* leaves.

Effect of CO_2 fertilisation on growth and LER

Contrary to the significantly faster LER of *O. australiensis* leaves when grown at 45°C in ambient (400 ppm) CO_2 conditions (Fig 3.2B), when CO_2 levels were raised to 700 ppm the LER of *O. australiensis* was identical at 30 and 45°C (Fig 3.4A). The LER of *O. sativa* was consistently ~3 mm h^{-1} at 400 ppm (Fig 3.2B) and 700 ppm CO_2 (Fig 3.4A) and was unaffected by atmospheric CO_2 level. Despite the fact that 700 ppm CO_2 could not further accelerate the rapid leaf elongation rate when *O. australiensis* grew at 45°C, shoot dry mass increased significantly at 45°C in an elevated CO_2 atmosphere ($p < 0.01$; Fig 3.4B), as was observed for biomass increments at 400 ppm (Fig 3.1). On the contrary, shoot dry mass was unaffected by heat when *O. sativa* grew at 700 ppm CO_2 (Fig 3.4B), while growth was slower at

45°C in 400 ppm CO_2 (Fig 3.1). That is, a 700 ppm CO_2 atmosphere reversed the impaired growth rates seen in *O. sativa* at 45°C in 400 ppm CO_2.

Net CO_2 assimilation; R_{light}, Γ* *and* ETR

Respiration in the light (R_{light}) was similar for *O. australiensis* and *O. sativa* (Fig 3.5A). Both species had a significantly faster R_{light} with rising temperature, and the stimulation of R_{light} was not significantly different between the species. The stimulation of R_{light} was the same irrespective of a long-term heat exposure or heat-shock, suggesting limited acclimation potential of R_{light} for rice or its wild relative. The CO_2 compensation point of leaves in the absence of respiration (*Γ**)—equivalent to the amount of CO_2 released by leaves via the process of photorespiration—increased significantly with temperature (Fig 3.5B). Again, the control 30°C values and the extent of temperature stimulation was not significantly different between the two species or between the long-term heat exposure and heat shock. Photosynthetic electron transport rate (*ETR*) determined by chlorophyll fluorescence was significantly different between *O. sativa* and *O. australiensis* irrespective of growth temperature (Fig 3.5C), with *O. sativa* having significantly faster *ETR* under all treatments. While R_{light} and *Γ** did not perform differently with 700 ppm CO_2 enrichment, the *ETR* of both species declined somewhat with heat shock, though perhaps more so for *O. australiensis*.

Net CO_2 assimilation; RACiR curves

To further characterise the response of photosynthesis in *O. sativa* and *O. australiensis* to temperature, RACiR curves were made and analysed using the standard C_3 photosynthesis model (Fig 3.6, Table 3.1). CO_2 carboxylation was the prevailing limitation on A_n at 400 ppm for both species. Consistent with growth and

leaf sugar accumulation, at an ambient CO_2 of 400 ppm *O. australiensis* had faster and *O. sativa* slower A_n after acclimation to 45°C, relative to 30°C (Fig 3.6). The beneficial response of A_n at 45°C relative to 30°C for *O. australiensis* at 400 ppm (Fig 3.6) can in part be attributed to overcoming the inherently lower g_s, C_i and subsequent C_c concentrations in *O. australiensis* at 30°C (Table 3.1). Interestingly, comparisons of long-term 45°C heat exposure versus heat-shocked plants revealed reduced photosynthetic capacity of *O. australiensis* in heat-shocked plants only, while similar reductions in capacity were observed in *O. sativa* irrespective of exposure time (Table 3.1). This may suggest potential for acclimation of CO_2 assimilation to heat in the wild rice but not in *O. sativa*.

In plants acclimated to a [CO_2] of 700 ppm, overall photosynthetic capacity was significantly reduced for *O. australiensis* at 45°C. This was demonstrated by a lower A_n and substantially reduced V_{cmax} and J_{max} at 45°C and 700 ppm CO_2 relative to 400 ppm (Fig 3.6; Table 3.1). In contrast, *O. sativa* had an increase in A_n and a more limited suppression of V_{cmax} and J_{max} when grown at 45°C and exposed to 700 ppm. For example, at 45°C and 700 ppm the V_{cmax} of *O. australiensis* was 66% of the rate at 400 ppm, while for *O. sativa*, V_{cmax} remained at 93% of the rate at 400 ppm. A greater downregulation of photosynthetic capacity under heat and high CO_2 was therefore seen in *O. australiensis* compared with *O. sativa*. Significant suppression of photosynthetic capacity of *O. australiensis* exposed to high CO_2 and heat was accompanied by greater photosynthetic efficiency under these conditions, and therefore little change in A_n was observed; this was particularly evident in *O. australiensis* because it had a higher ratio of A_n/V_{cmax} (Table 3.1).

Table 3.1. Photosynthetic characteristics of wild and domestic rice grown under variable temperature and CO_2 regimes, derived from RACiR curves. The shock group

was made up of 30°C plants exposed to 45°C for 2 d. All values are means (± SE), except for A_n/V_{cmax} (%). Means that do not share a letter are significantly different (as determined by Duncan's Multiple Range Test).

Temperature (°C)	30				45				45 (Shock)			
[CO_2] (ppm)	400		700		400		700		400		700	
Species	*O. sat*	*O. aus*	*O. sat*	*O. aus*	*O. sat*	*O. aus*	*O. sat*	*O. aus*	*O. sat*	*O. aus*	*O. sat*	*O. aus*
V_{Cmax} (µmol CO_2 m^{-2} s^{-1})	99.4 ± 7.1 ab	90.2 ± 3.4 ab	102.1 ± 11.4 ab	78.6 ± 1 a	185.2 ± 10.1 c	172.4 ± 18.4 c	167.3 ± 16.5 c	125.2 ± 8.4 b	180.6 ± 14 c	125.3 ± 15.4 b	167.6 ± 6.3 c	111.6 ± 16 ab
A_n/V_{Cmax} (%)	24.5	20.2	34.5	34.8	10.6	12.5	19	18.8	12.2	12.5	17.8	18.4
J_{max} (µmol m^{-2} s^{-1})	179.2 ± 11 abcde	164.1 ± 5.8 abc	186.2 ± 19.3 abcdef	145.7 ± 3.0 ab	226 ± 6.9 ef	215.3 ± 22.6 def	194.2 ± 16.9 bcdef	145.2 ± 11.3 ab	217.8 ± 16.9 f	165 ± 16.8 abcd	204.7 ± 17.8 cdef	132 ± 16.6 a
C_c (Pa)	26.6 ± 0.9 ab	24.3 ± 0.5 a	44.6 ± 2.7 c	46.1 ± 2 cd	27.0 ± 0.8 ab	29.9 ± 0.3 b	51.4 ± 1.3 e	51.3 ± 1.6 e	28.5 ± 0.2 ab	30.5 ± 0.6 b	47.9 ± 1.6 cde	49.8 ± 2.8 de
g_m (µmol CO_2 m^{-2} s^{-1} Pa^{-1})	10.9 ± 0.1 bcd	5.5 ± 0.07 a	11.2 ± 0.11 bcd	5.4 ± 0.06 a	22.1 ± 0.46 f	10.5 ± 0.26 bc	21.6 ± 0.12 f	11.3 ± 0.32 cd	20.6 ± 0.41 e	10.2 ± 0.26 b	22.4 ± 0.58 f	11.7 ± 0.49 d
C_i (Pa)	28.9 ± 0.8 a	27.9 ± 0.7 a	47.6 ± 2.5 b	51.1 ± 2 bc	27.9 ± 0.9 a	31.9 ± 0.4 a	52.9 ± 1.2 c	53.4 ± 1.7 c	29.6 ± 0.2 a	32 ± 0.6 a	49.2 ± 1.7 bc	51.6 ± 2.9 bc
Leaf Temperature (°C)	28.8 ± 0.2 a	29.5 ± 0.2 a	29.6 ± 0.2 a	29.6 ± 0.2 a	41.5 ± 0.4 d	38.9 ± 1.1 bc	41.1 ± 0.1 cd	41.6 ± 0.5 d	40.2 ± 0.4 bc	39.4 ± 0.9 b	41.7 ± 0.5 d	42.2 ± 0.7 d
g_s (mol H_2O m^{-2} s^{-1})	0.64 ± 0.08 abc	0.49 ± 0.1 abc	0.34 ± 0.04 a	0.4 ± 0.05 ab	0.67 ± 0.08 bc	0.77 ± 0.08 cd	0.6 ± 0.003 abc	0.53 ± 0.11 abc	0.84 ± 0.15 d	0.79 ± 0.09 cd	0.45 ± 0.08 ab	0.39 ± 0.11 ab
n	5	6	5	4	6	5	6	6	6	5	5	4

A_n, CO_2 assimilation rate measured on LI-6800; V_{Cmax}, maximum velocity of Rubisco; J_{max}, maximum velocity of electron transport; C_C, pressure of CO_2 at the chloroplast; g_m, mesophyll conductance; C_i, pressure of intercellular CO_2; g_s, stomatal conductance to water; *n*, number of replicate RACiR curves per group.

Relative abundance of Rca as a percentage of Rubisco

At 45°C the abundance of Rca as a proportion of RbcL was significantly reduced in *O. australiensis*. In spite of this, A_n increased in *O. australiensis*. The opposite was observed for *O. sativa* – at 45°C the abundance of Rca as a proportion of RbcL was significantly increased, while photosynthetic rates declined.

The effect of temperature on the relative abundance of Rca as a proportion of Rubisco for *O. sativa* and *O. australiensis* was determined (Fig 3.7). There was a significant interaction between species and temperature ($p < 0.001$). At 30°C, the concentration of Rca relative to Rubisco in *O. australiensis* leaves was ~2.2 times greater than it was at 45°C. Conversely, in *O. sativa* leaves the concentration of Rca relative to Rubisco at 30°C was half that at 45°C (Fig 3.7a). Differences in Rca abundance relative to RbcL were not strictly due to differences in the abundance of

RbcL but were driven more by changes in the abundance of Rca (Fig 3.7b & c). Despite the relative abundance of Rca decreasing at 45°C in *O. australiensis*, steady-state A_n appeared to increase (measured on the same leaves as those sampled for protein; Fig 3.7a). By contrast, the increase in relative Rca abundance in *O. sativa* was associated with decreased A_n with heat.

A two-way interaction between temperature and time of day ($p < 0.01$) shows that the abundance of Rca relative to RbcL increased following exposure to 45°C for 5 h, though no other significant differences were detected (Supplementary Figure 3.9; Tukey HSD). Similarly, a two-way interaction between species and time of day ($p < 0.001$) shows that Rca:RbcL increased 5 h into the photoperiod in *O. sativa*, and this ratio was greater than that seen in *O. australiensis* under the same conditions (Tukey HSD).

Discussion

This study explores two closely related grasses in which we have previously established starkly contrasting degrees of thermotolerance (Scafaro et al. 2011; Scafaro et al. 2016). The tolerant species, *Oryza australiensis*, is an extremophile distributed across the hot northern savannah of Australia, while *O. sativa* ssp. *japonica* is largely cultivated in the humid sub-tropics where daily maxima rarely exceed 35°C. However, the mechanisms that enable wild rice to continue growing, even during sustained periods of high temperature, are still not known. This paper links photosynthetic processes in *O. australiensis* with its sustained thermotolerance during early vegetative growth.

When grown at high temperature (45°C) and ambient CO_2 (400 ppm), *O. australiensis* accumulated more biomass than *O. sativa*. Furthermore, *O.*

australiensis was susceptible to low temperature (25°C) while *O. sativa* suffered severe growth penalties when grown at high temperature. Ali et al. (2019) showed reductions in growth of the same *O. sativa* cultivar (Amaroo) when grown at high temperature for a short time during grain filling. Furthermore, Perdomo et al. (2015) show reduced biomass accumulation in *O. sativa* (cv. Bomba) when grown at high temperature. Taken together these results underline the superior thermotolerance of *O. australiensis* to sustained heat, relative to *O. sativa*. Interestingly, 45°C days only slowed biomass gain in *O. sativa* after three weeks, even though leaves elongated more slowly within a few hours of 45°C being imposed. Previous studies have highlighted the impact of heat stress on the photosynthetic and respiratory rates and subsequent biomass accumulation of *O. sativa* (Jagadish et al. 2015; Shi et al. 2016; Zhang et al. 2018; Karwa et al. 2020). While these studies only consider the impact of heat over a relatively short time (max. 20 days), and usually during or after anthesis with a focus on grain yield, they reveal alterations to rice carbon metabolism initiated by high temperatures. In this context, the results of the present study indicate that vegetative growth in *O. sativa* is impeded at sustained high temperature by its cumulative impact on the whole-plant carbon budget. We ascribe this to reduced carbohydrate availability and mobilisation from reserves rather than a shift in partitioning of biomass between roots and shoots because root-to-shoot ratios were relatively stable over time (Supplementary Figure 3.10).

LER data showed that leaves of both species elongated rapidly 1 - 2 h after transfer to 45°C, which suggests that heat did not directly impair the mechanics of cell expansion, for example by disrupting the cytoskeleton (Smertenko et al. 1997) or membrane integrity (Horváth et al. 2012) but rather it acted through secondary impacts of heat on metabolism. The species contrast was first seen 4 h after plants were exposed to 45°C, consistent with a growing deficit of carbohydrates in *O. sativa*

and subsequent slowing of leaf growth over the diurnal cycle of heat. This is in accord with the finding that the impact of heat on grain filling in *O. sativa* becomes more severe with prolonged exposure to heat (Shi et al. 2016). On the other hand, LER in *O. australiensis* was essentially maintained above 3 mm h^{-1} for 10 h at 45°C, while rates returned to 1 - 1.5 mm h^{-1} in both species at night. These findings corroborate those of Scafaro *et al.* (2010) and Scafaro *et al.* (2016), who showed reductions in LER of *O. sativa* following 4 h of exposure to 45°C of up to 50%, while Scafaro *et al.* (2016) report no significant reduction in the LER of *O. australiensis* after 4 h at 45°C, which was linked to its photosynthetic thermotolerance. That *O. australiensis* has consistently demonstrated superior growth to *O. sativa* in the face of high temperature suggests that it has access to the resources (namely carbon resources either generated via photosynthesis or mobilised from stored carbohydrates) that are required to support such growth.

We found that after 10 h at 45°C, sugar concentrations accumulated threefold in *O. australiensis* and doubled in *O. sativa* (Fig 3.3). These data contrast with the steady sugar levels recorded in rice leaves exposed to longer periods at less extreme temperatures up to 40ºC (Rashid et al. 2020). Soluble sugars are thought to play a role in various abiotic stress tolerances by acting as signalling molecules, by conferring antioxidant properties (Harsh et al. 2016), by acting as compatible solutes (Kaplan and Guy 2004) and stabilising protein structure (Lee and Timasheff 1981). Given the role that soluble sugars play in protecting plants against thermally induced damage, we speculate that high soluble sugar levels in both rice species contributed to thermotolerance at the molecular level. In addition, the species-specific accumulation of sugars in response to heat is interesting as a resource for growth and photosynthetic rates. Steady-state soluble sugar concentrations in leaves are the consequence of assimilation rates, starch degradation, and sugar utilisation (Stitt and

Zeeman 2012). Therefore, trebling of leaf soluble sugars in *O. australiensis* should be seen in the context of the faster growth at 45°C compared with *O. sativa*; leaf growth in *O. australiensis* continued unabated at 45°C while *O. sativa* leaves elongated significantly slower (Fig 3.2A & B). This makes the marked accumulation of soluble sugars in leaves of *O. australiensis* remarkable because it occurred in spite of a greater demand for carbohydrates that would be required to sustain steady growth. The maintenance of *O. sativa* leaf sugar levels at 45°C may have come at a cost to starch reserves, considering photosynthesis was impaired under these hotter growth conditions in the domesticated rice (Fig 3.6, 7). Increasing soluble sugar content in plant tissues is often reported alongside reductions in starch reserves as stored carbohydrates are mobilised, requiring the expression of a suite of genes involved in starch degradation and monosaccharide transport (Kaplan et al. 2006). For example, Yichie et al. (2019) reported increases in expression of a starch synthase in a salt-susceptible population of wild rice, while a tolerant population of the same species had increased expression of monosaccharide transporters. The rapid slowing of biomass accumulation in *O. sativa* after 28 days at 45°C (Fig 3.1) may in fact be due to starch reserves reaching a critically low level after sustained depletion.

Contrasting sugar levels in the two species and previous observations that photosynthesis is vulnerable to heat-induced damage at 45°C (Salvucci et al. 2001; Sage et al. 2008; Scafaro et al. 2010; Yamori et al. 2014; Scafaro et al. 2016; Busch and Sage 2017) led to an inquiry into photosynthetic metabolism in the two species under heat shock. Unsurprisingly, peak A_n recorded (as instantaneous measurements) at 30°C were 50% higher in the domesticated cultivar of *O. sativa*, while at 45°C rates in *O. australiensis* and *O. sativa* were approximately equal (Fig 3.7). This finding emphasises the highly significant opposing effects of heat on assimilation in the two species (i.e. a 30% increase in *O. australiensis* and a 20%

decrease in *O. sativa* when exposed to 45°C). Similarly, IR64 (an *O. sativa* cultivar) had reduced photosynthesis when exposed to a high temperature shock (Rashid et al. 2020). Furthermore, *O. sativa* (cv. Bomba) had reduced daily average assimilation rates when grown at high temperature (Perdomo et al. 2015). Importantly, the changes in instantaneous measures of A_n during heat shock could not be explained by differences in stomatal conductance (g_s), intracellular CO_2 (C_i), leaf temperature nor vapour pressure deficit (Supplementary Figure 3.11). This finding is supported by a preliminary study we performed, where significant changes in VPD could not account for differences in assimilation rates because, despite the response of g_s to high temperature, intercellular CO_2 availability was the same for both species (Supplementary Figure 3.12). Differences in leaf anatomical properties between rice and its wild relatives do exist, and these differences seem to influence CO_2 conductance properties between the air and chloroplasts (Scafaro et al. 2011; Giuliani et al. 2013). However, the results presented here suggest that temperature dependent changes in A_n are dictated by changes in the efficiency of the underlying photosynthetic biochemistry.

RACiR curves are a novel technique that have been used successfully to rapidly acquire gas exchange data for the estimation of V_{cmax} and J_{max} while simultaneously overcoming the issues of stomatal attenuation and changes in enzyme kinetics that are characteristic of the traditional A:C_i curve (Stinziano et al. 2017; Coursolle et al. 2019; Stinziano et al. 2019; Vincent et al. 2020). RACiR curves, Laisk curves, and point measures of chlorophyll fluorescence showed that differences in assimilation between wild and domestic rice during long- or short-term heat exposure could not be explained by differences in photorespiration, respiration or *ETR* (Fig 3.5). That is, the thermotolerance of assimilation in *O. australiensis*

cannot be due to a species-specific response of photorespiration, respiration, or *ETR* to heat.

Rates of photorespiration increase with increasing temperature because the specificity of Rubisco to CO_2 declines relative to O_2 under heat, and the abundance of CO_2 relative to O_2 is reduced due to differences in the solubility of these two gases at high temperature (Walker et al. 2016; Dusenge et al. 2019). Given that rates of photorespiration are driven in part by the affinity of Rubisco for CO_2, we conclude that the lack of species differences observed here indicate that the wild and domestic rice Rubisco isoforms must have similar affinities for their substrates. This is in accordance with recent work done on *Arabidopsis* that showed no genetic basis for variation in photorespiration between ecotypes (Tomeo and Rosenthal 2018).

Similarly, respiration increases with increasing temperature (Dusenge et al. 2019). Here, we showed similar increases in respiration at high temperature in both species. Rashid et al. (2020) showed that rice (IR64) had limited respiratory acclimation over seven days of exposure to 40^{o}C, mirroring the results obtained here. Interestingly, *O. australiensis* respiration showed greater potential to acclimate to sustained high temperature at 400 ppm CO_2 than *O. sativa*. Thus, while we cannot attribute differences in assimilation to the observed differences in respiration between the species at high temperature, *O. australiensis* may provide useful germplasm to explore variation in respiratory response under heat. Finally, reduced A_n in *O. sativa* at 45°C (both long- and short-term) comes about even though concentrations of CO_2 at the chloroplast remained stable or increase. Taken together, the results presented thus far provide evidence for the notion of biochemical limitations at the chloroplast during episodes of high temperature in *O. sativa* but not *O. australiensis*. Interestingly, the RACiR curves showed some degree of acclimation of assimilation

to both long-term heat (*cf*. short-term) and elevated ambient CO_2 in *O. australiensis* though not in *O. sativa* (Fig 6). The downregulation of photosynthetic capacity in *O. australiensis* when grown at 700 ppm CO_2 suggests that the wild species is capable of meeting sink demands even when reducing its investment in photosynthetic machinery. This is seen in its most extreme form in *O. australiensis* plants grown at 700 ppm CO_2 *and* sustained high temperature, which may indicate a metabolism that is adapted to operate most efficiently in hot environments. On the other hand, *O. sativa* appears to maintain investment in photosynthetic machinery at high CO_2, resulting in much faster A_n with CO_2 enrichment. Under a similar 700 ppm CO_2 enrichment study, *O. sativa* significantly increased A_n while the wild relative *Oryza meridionalis* had a limited response (Rahman et al. 2018). The A_n of wild relatives thus seems to be less responsive to CO_2 fertilisation relative to domestic rice. We speculate that *O. australiensis* reduced photosynthetic capacity with CO_2 enrichment because of limited sink demand, whereas *O. sativa* maintained photosynthetic capacity due to artificial selection increasing demand for assimilates needed for grain filling. Supporting this postulate, enriched CO_2 has consistently been shown to increase grain yield in commercial rice (Madan et al. 2012), demonstrating that improved crop performance by engineering faster assimilation will likely be effective only in cases where demand for assimilates (sink strength) has not already reached saturation.

The CO_2 response curves suggested Rubisco carboxylation limitations existed in *both species* at 45°C and 400 ppm CO_2, and thus a direct effect of heat on CO_2 fixation by Rubisco is strongly implicated as the reason for differences in photosynthetic thermotolerance between the species. Prior evidence (Scafaro et al. 2016) suggests that the properties of Rca form part of the explanation for heat tolerance in *O. australiensis*. Indeed, the photosynthetic heat tolerance of *O. sativa*

can be improved with overexpression of *O. australiensis* or maize Rca (Yamori et al. 2012; Scafaro et al. 2018). The activation of Rubisco by Rca depends on the ratio of the two enzymes and the susceptibility of Rca to heat (Perdomo et al. 2017). Rca abundance can respond dynamically to heat stress in a timeframe of hours to days through transcriptional and post-transcriptional regulation (Degen et al. 2021). In relation to the ratio of Rubisco to Rca, Fukayama *et al.* (2012) and Fukayama *et al.* (2018) showed that rice overexpressing functional Rca had reduced abundance of Rubisco, while an Rca knock-out line had increased Rubisco abundance. Suganami et al. (2018) showed that as the abundance of Rubisco increases, cells disinvest in Rca, and *vice versa*, indicating that this feedback operates in both directions. Further, a recent report in rice demonstrated that overexpression of both Rca and Rubisco leads to improvement in assimilation, particularly under heat stress (Qu et al. 2021). Given that Rca functions in the activation of Rubisco (Portis 2003), these previous results imply that photosynthetic capacity is tightly regulated and set by the number of active Rubisco sites. As we observed changes in Rca abundance, but no significant changes in Rubisco abundance with heat, Rca was regulating photosynthetic capacity. Contrary changes in total Rca abundance between the species when grown at 45°C complement the inherent thermostability of the Rca isoforms in each species. Specifically, the dominant Rca β isoform found in *O. australiensis* is known to have greater thermostability than its *O. sativa* ortholog (Scafaro et al. 2016). Thus, it may be the case here that at 30°C *O. australiensis* Rca was operating below its thermal optimum. An increase in abundance of Rca presumably compensated for the loss in efficiency at this sub-optimal temperature (Fig 3.7). This limitation was presumably lifted at higher temperatures, allowing for increased enzyme function and a decrease in Rca abundance. Interestingly, reduced investment in Rca because of increased Rca activity may enhance plant growth over

long periods of heat because after Rubisco, Rca is the second highest consumer of ATP in leaves (Li et al. 2017). Furthermore, the savings in ATP costs associated with reduced Rca may contribute to the greater respiratory acclimation potential of hot grown *O. australiensis,* as we explore above. We suggest that the opposite applies in *O. sativa*, whose Rca may have been operating outside its thermal optimum at 45°C, likewise explaining the increased accumulation of the enzyme during exposure to heat and reduced accumulation under optimal conditions (30°C). When grown at high temperatures, the stability of Rca may therefore explain the reduced Rca abundance, greater photosynthetic performance, and ultimately faster growth in the heat-adapted *O. australiensis*. This is supported by our earlier evidence that transgenic rice expressing the thermotolerant ortholog of Rca from *O. australiensis* had increased growth rates and seed yield at 45°C (Scafaro et al. 2018). We further demonstrate in this study that neither photosynthetic electron transport, respiratory capacity nor photorespiratory performance could explain the susceptibility of *O. sativa* to extreme heat. Furthermore, we found no evidence for differences in Rubisco's affinity for CO_2 in the wild rice, which would have provided partial explanation for improved assimilation at high temperature. Rubisco activase as a driver of heat tolerance in wild rice therefore seems highly likely (Perdomo et al. 2017). Whether this mechanistic basis of heat tolerance applies to other important crops should be a priority of future research.

References

Ali F, Waters DLE, Ovenden B, Bundock P, Raymond CA, Rose TJ (2019) Australian rice varieties vary in grain yield response to heat stress during reproductive and grain filling stages. J Agron Crop Sci 205:179–187. https://doi.org/10.1111/jac.12312

Asseng S, Ewert F, Martre P, Rötter RP, Lobell DB, Cammarano D, Kimball BA, Ottman MJ, Wall GW, White JW, Reynolds MP, Alderman PD, Prasad PV V, Aggarwal PK, Anothai J, Basso B, Biernath C, Challinor AJ, De Sanctis G, Doltra J, Fereres E, Garcia-Vila M, Gayler S, Hoogenboom G, Hunt LA, Izaurralde RC, Jabloun M, Jones CD, Kersebaum KC, Koehler A-K, Müller C, Naresh Kumar S, Nendel C, O'Leary G, Olesen JE, Palosuo T, Priesack E, Eyshi Rezaei E, Ruane AC, Semenov MA, Shcherbak I, Stöckle C, Stratonovitch P, Streck T, Supit I, Tao F, Thorburn PJ, Waha K, Wang E, Wallach D, Wolf J, Zhao Z, Zhu Y (2015) Rising temperatures reduce global wheat production. Nat Clim Chang 5:143–147. https://doi.org/10.1038/nclimate2470

Atkin OK, Tjoelker MG (2003) Thermal acclimation and the dynamic response of plant respiration to temperature. Trends Plant Sci 8:343–351. https://doi.org/https://doi.org/10.1016/S1360-1385(03)00136-5

Atwell BJ, Wang H, Scafaro AP (2014) Could abiotic stress tolerance in wild relatives of rice be used to improve *Oryza sativa*? Plant Sci 215–216:48–58. https://doi.org/10.1016/j.plantsci.2013.10.007

Avenson TJ, Saathoff AJ (2018) Sub-saturating multiphase flash irradiances to estimate maximum fluorescence yield. Methods Mol Bio 1770:105-120. https://doi.org/10.1007/978-1-4939-7786-4_7

Benomar L, Lamhamedi MS, Pepin S, Rainville A, Lambert M-C, Margolis HA, Bousquet J, Beaulieu J (2018) Thermal acclimation of photosynthesis and

respiration of southern and northern white spruce seed sources tested along a regional climatic gradient indicates limited potential to cope with temperature warming. Ann Bot 121:443–457. https://doi.org/10.1093/aob/mcx174

Busch FA, Sage RF (2017) The sensitivity of photosynthesis to O_2 and CO_2 concentration identifies strong Rubisco control above the thermal optimum. New Phytol 213:1036–1051. https://doi.org/10.1111/nph.14258

Campbell C, Atkinson L, Zaragoza-Castells J, Lundmark M, Atkin O, Hurry V (2007) Acclimation of photosynthesis and respiration is asynchronous in response to changes in temperature regardless of plant functional group. New Phytol 176:375–389. https://doi.org/10.1111/j.1469-8137.2007.02183.x

Chi Y, Xu M, Shen R, Yang Q, Huang B, Wan S (2013) Acclimation of foliar respiration and photosynthesis in response to experimental warming in a temperate steppe in Northern China. PLoS One 8:e56482. https://doi.org/https://doi.org/10.1371/journal.pone.0056482

Coursolle C, Otis Prud'homme G, Lamothe M, Isabel N (2019) Measuring rapid *A–Ci* curves in boreal conifers: black spruce and balsam fir. Front. Plant Sci. 10:1276. https://doi.org/10.3389/fpls.2019.01276

CSIRO (2018) State of the Climate. https://www.csiro.au/en/Showcase/state-of-the-climate. Accessed 3 Sep 2020

Degen, GE, Orr, DJ, & Carmo-Silva, E (2021) Heat-induced changes in the abundance of wheat Rubisco activase isoforms. New Phytol. 229:1298-1311. https://doi.org/10.1111/nph.16937

Dhir B (2018) Crop productivity in changing climate. In: Lichtfouse E (ed) Sustainable Agriculture Reviews 27. Springer International Publishing, Cham, pp 213–241

Dusenge ME, Duarte AG, Way DA (2019) Plant carbon metabolism and climate change: elevated CO2 and temperature impacts on photosynthesis,

photorespiration and respiration. New Phytol 221:32–49. https://doi.org/10.1111/nph.15283

Ebrahim MK, Zingsheim O, El-Shourbagy MN, Moore PH, Komor E (1998) Growth and sugar storage in sugarcane grown at temperatures below and above optimum. J Plant Physiol 153:593–602. https://doi.org/https://doi.org/10.1016/S0176-1617(98)80209-5

Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species. Planta 149:78–90. https://doi.org/10.1007/BF00386231

Feller U, Crafts-Brandner SJ, Salvucci ME (1998) Moderately high temperatures inhibit ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) activase-mediated activation of Rubisco. Plant Physiol 116:539–546. https://doi.org/10.1104/pp.116.2.539

Fischer T, Byerlee D, Edmeades G (2014) Crop yields and global food security: will yield increase continue to feed the world? Aust Cent Int Agric Res 634

Fukayama H, Mizumoto A, Ueguchi C, Katsunuma J, Morita R, Sasayama D, Hatanaka T, Azuma T (2018) Expression level of Rubisco activase negatively correlates with Rubisco content in transgenic rice. Photosynth Res 137:465–474. https://doi.org/10.1007/s11120-018-0525-9

Fukayama H, Ueguchi C, Nishikawa K, Katoh N, Ishikawa C, Masumoto C, Hatanaka T, Misoo S (2012) Overexpression of Rubisco activase decreases the photosynthetic CO2 assimilation rate by reducing Rubisco content in rice leaves. Plant Cell Physiol 53:976–986. https://doi.org/10.1093/pcp/pcs042

Giuliani R, Koteyeva N, Voznesenskaya E, Evans MA, Cousins AB, Edwards GE (2013) Coordination of leaf photosynthesis, transpiration, and structural traits in rice and wild relatives (Genus *Oryza*). Plant Physiol 162:1632–1651.

https://doi.org/10.1104/pp.113.217497

Gomathi R, Shiyamala S, Vasantha S, Johnson DE, Janani PK (2013) Kinetics of metabolism in sugarcane (*Saccharum officinarum* L.) under heat stress. Indian J Plant Physiol 18:41–47. https://doi.org/10.1007/s40502-013-0011-5

Harsh A, Sharma YK, Joshi U, Rampuria S, Singh G, Kumar S, Sharma R (2016) Effect of short-term heat stress on total sugars, proline and some antioxidant enzymes in moth bean (*Vigna aconitifolia*). Ann Agric Sci 61:57–64. https://doi.org/https://doi.org/10.1016/j.aoas.2016.02.001

Hermida-Carrera C, Kapralov M V, Galmés J (2016) Rubisco catalytic properties and temperature response in crops. Plant Physiol 171:pp.01846.2016. https://doi.org/10.1104/pp.16.01846

Horváth I, Glatz A, Nakamoto H, Mishkind ML, Munnik T, Saidi Y, Goloubinoff P, Harwood JL, Vigh L (2012) Heat shock response in photosynthetic organisms: membrane and lipid connections. Prog Lipid Res 51:208–220. https://doi.org/https://doi.org/10.1016/j.plipres.2012.02.002

IPCC (2019) Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems

Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 4:1265–1272 . https://doi.org/https://doi.org/10.1074/mcp.M500061-MCP200

Jagadish SVK, Murty MVR, Quick WP (2015) Rice responses to rising temperatures – challenges, perspectives and future directions. Plant Cell Environ 38:1686–1698. https://doi.org/10.1111/pce.12430

Jagadish SVK, Septiningsih EM, Kohli A, Thomson MJ, Ye C, Redoña E, Kumar A, Gregorio GB, Wassmann R, Ismail AM, Singh RK (2012) Genetic advances in adapting rice to a rapidly changing climate. J Agron Crop Sci 198:360–373. https://doi.org/10.1111/j.1439-037X.2012.00525.x

Kaplan F, Guy CL (2004) β-amylase induction and the protective role of maltose during temperature shock. Plant Physiol 135:1674–1684. https://doi.org/10.1104/pp.104.040808

Kaplan F, Sung DY, Guy CL (2006) Roles of β-amylase and starch breakdown during temperatures stress. Physiol Plant 126:120–128. https://doi.org/10.1111/j.1399-3054.2006.00604.x

Karwa S, Bahuguna RN, Chaturvedi AK, Maurya S, Arya SS, Chinnusamy V, Pal M (2020) Phenotyping and characterization of heat stress tolerance at reproductive stage in rice (*Oryza sativa* L.). Acta Physiol Plant 42:29. https://doi.org/10.1007/s11738-020-3016-5

Laisk AK (1977) Kinetics of photosynthesis and photorespiration of C3 in plants. Nauka, Moscow

Lee JC, Timasheff SN (1981) The stabilization of proteins by sucrose. J Biol Chem 256:7193–7201

Li L, Nelson CJ, Trösch J, Castleden I, Huang S, Millar AH (2017) Protein degradation rate in *Arabidopsis thaliana* leaf growth and development. Plant Cell 29:207-228. https://doi.org/10.1105/tpc.16.00768

Loriaux SD, Avenson TJ, Welles JM, McDermitt DK, Eckles RD, Rienshe B, Genty B (2013) Closing in on maximum yield of chlorophyll fluorescence using a single multiphase flash of sub-saturating intensity. Plant Cell Environ 36: 1755-1770. https://doi.org/10.1111/pce.12115

Madan P, Jagadish SVK, Craufurd PQ, Fitzgerald M, Lafarge T, Wheeler TR (2012)

Effect of elevated CO_2 and high temperature on seed-set and grain quality of rice. J Exp Bot 63:3843–3852. https://doi.org/10.1093/jxb/ers077

Mahlstein I, Knutti R, Solomon S, Portmann RW (2011) Early onset of significant local warming in low latitude countries. Environ Res Lett 6:034009. https://doi.org/10.1088/1748-9326/6/3/034009

Peng S, Huang J, Sheehy JE, Laza RC, Visperas RM, Zhong X, Centeno GS, Khush GS, Cassman KG (2004) Rice yields decline with higher night temperature from global warming. Proc Natl Acad Sci U S A 101:9971–9975. https://doi.org/10.1073/pnas.0403720101

Perdomo JA, Conesa M, Medrano H, Ribas-Carbó M, Galmés J (2015) Effects of long-term individual and combined water and temperature stress on the growth of rice, wheat and maize: Relationship with morphological and physiological acclimation. Physiol Plant 155:149–165. https://doi.org/10.1111/ppl.12303

Perdomo JA, Capó-Bauçà S, Carmo-Silva E, Galmés J (2017) Rubisco and Rubisco activase play an important role in the biochemical limitations of photosynthesis in rice, wheat, and maize under high temperature and water deficit. Front. Plant Sci. 8:490. https://doi.org/10.3389/fpls.2017.00490

Portis AR (2003) Rubisco activase – Rubisco's catalytic chaperone. Photosynth Res 75:11–27. https://doi.org/10.1023/A:1022458108678

Qu Y, Sakoda K, Fukayama H, Kondo E, Suzuki Y, Makino A, Terashima I, Yamori W (2021) Overexpression of both Rubisco and Rubisco activase rescues rice photosynthesis and biomass under heat stress. Plant Cell Environ 44:2308–2320. https://doi.org/https://doi.org/10.1111/pce.14051

R Core Team (2017) R: A language and environment for statistical computing

Rahman S, Duursma RA, Muktadir MA, Roberts TH, Atwell BJ (2018) Leaf canopy architecture determines light interception and carbon gain in wild and

domesticated *Oryza* species. Environ Exp Bot 155:672–680. https://doi.org/https://doi.org/10.1016/j.envexpbot.2018.08.008

Rashid FAA, Crisp PA, Zhang Y, Berkowitz O, Pogson BJ, Day DA, Masle J, Dewar RC, Whelan J, Atkin OK, Scafaro AP (2020) Molecular and physiological responses during thermal acclimation of leaf photosynthesis and respiration in rice. Plant Cell Environ 43:594–610. https://doi.org/10.1111/pce.13706

Reddy KR, Reddy VR, Hodges HF (1992) Temperature effects on early season cotton growth and development. Agron J 84:229. https://doi.org/10.2134/agronj1992.00021962008400020021x

Sage RF, Way DA, Kubien DS (2008) Rubisco, Rubisco activase, and global climate change. J Exp Bot 59:1581–1595. https://doi.org/10.1093/jxb/ern053

Salvucci ME, Crafts-Brandner SJ (2004) Relationship between the heat tolerance of photosynthesis and the thermal stability of rubisco activase in plants from contrasting thermal environments. Plant Physiol 134:1460–1470. https://doi.org/10.1104/pp.103.038323

Salvucci ME, Osteryoung KW, Crafts-Brandner SJ, Vierling E (2001) Exceptional sensitivity of Rubisco activase to thermal denaturation in vitro and in vivo. Plant Physiol 127:1053–1064. https://doi.org/10.1104/pp.010357

Scafaro AP, Atwell BJ, Muylaert S, Reusel B Van, Ruiz GA, Rie J Van, Gallé A (2018) A thermotolerant variant of Rubisco activase from a wild relative improves growth and seed yield in rice under heat stress. Front Plant Sci 9:1–11. https://doi.org/10.3389/fpls.2018.01663

Scafaro AP, Bautsoens N, den Boer B, Van Rie J, Galle A (2019) A conserved sequence from heat-adapted species improves Rubisco activase thermostability in wheat. Plant Physiol 181: 43–54. https://doi.org/10.1104/pp.19.00425

Scafaro AP, Gallé A, Van Rie J, Carmo-Silva E, Salvucci ME, Atwell BJ (2016) Heat

tolerance in a wild *Oryza* species is attributed to maintenance of Rubisco activation by a thermally stable Rubisco activase ortholog. New Phytol 211:899–911. https://doi.org/10.1111/nph.13963

Scafaro AP, Haynes P, Atwell B (2010) Physiological and molecular changes in *Oryza meridionalis* Ng., a heat-tolerant species of wild rice. J Exp Bot 61:191–202. https://doi.org/https://doi.org/10.1093/jxb/erp294

Scafaro AP, Von Caemmerer S, Evans JR, Atwell BJ (2011) Temperature response of mesophyll conductance in cultivated and wild *Oryza* species with contrasting mesophyll cell wall thickness. Plant Cell Environ 34:1999–2008. https://doi.org/10.1111/j.1365-3040.2011.02398.x

Scharf KD, Berberich T, Ebersberger I, Nover L (2012) The plant heat stress transcription factor (Hsf) family: Structure, function and evolution. Biochim Biophys Acta - Gene Regul Mech 1819:104–119. https://doi.org/10.1016/j.bbagrm.2011.10.002

Shi P, Zhu Y, Tang L, Chen J, Sun T, Cao W, Tian Y (2016) Differential effects of temperature and duration of heat stress during anthesis and grain filling stages in rice. Environ Exp Bot 132:28–41. https://doi.org/https://doi.org/10.1016/j.envexpbot.2016.08.006

Smertenko A, Dráber P, Viklický V, Opatrný Z (1997) Heat stress affects the organization of microtubules and cell division in *Nicotiana tabacum* cells. Plant Cell Environ 20:1534–1542. https://doi.org/10.1046/j.1365-3040.1997.d01-44.x

Stinziano JR, McDermitt DK, Lynch DJ, Saathoff AJ, Morgan PB, Hanson DT (2019) The rapid *A/Ci* response: a guide to best practices. New Phytol 221:625–627. https://doi.org/10.1111/nph.15383

Stinziano JR, Morgan PB, Lynch DJ, Saathoff AJ, McDermitt DK, Hanson DT (2017) The rapid *A–Ci* response: photosynthesis in the phenomic era. Plant Cell

Environ 40:1256–1262. https://doi.org/10.1111/pce.12911

Stitt M, Zeeman SC (2012) Starch turnover: pathways, regulation and role in growth. Curr Opin Plant Biol 15:282–292 . https://doi.org/https://doi.org/10.1016/j.pbi.2012.03.016

Suganami M, Suzuki Y, Sato T, Makino A (2018) Relationship between Rubisco activase and Rubisco contents in transgenic rice plants with overproduced or decreased Rubisco content. Soil Sci Plant Nutr 64:352–359. https://doi.org/10.1080/00380768.2018.1433455

Tilman D, Balzer C, Hill J, Befort BL (2011) Global food demand and the sustainable intensification of agriculture. Proc Natl Acad Sci 108:20260–20264. https://doi.org/10.1073/pnas.1116437108

Tjoelker MG (2018) The role of thermal acclimation of plant respiration under climate warming: Putting the brakes on a runaway train? Plant Cell Environ 41:501–503. https://doi.org/10.1111/pce.13126

Tomeo NJ, Rosenthal DM (2018) Photorespiration differs among Arabidopsis thaliana ecotypes and is correlated with photosynthesis. J Exp Bot 69:5191–5204. https://doi.org/10.1093/jxb/ery274

Vincent C, Pierre MO, Stinziano JR (2020) Racing against stomatal attenuation: rapid CO_2 response curves more reliably estimate photosynthetic capacity than steady state curves in a low conductance species. bioRxiv. https://doi.org/10.1101/2020.08.28.270785

Wahid A, Gelani S, Ashraf M, Foolad MR (2007) Heat tolerance in plants: An overview. Environ Exp Bot 61:199–223. https://doi.org/https://doi.org/10.1016/j.envexpbot.2007.05.011

Walker BJ, VanLoocke A, Bernacchi CJ, Ort DR (2016) The costs of photorespiration to food production now and in the future. Annu Rev Plant Biol 67:107–129.

https://doi.org/10.1146/annurev-arplant-043015-111709

Wright IJ, Dong N, Maire V, Prentice IC, Westoby M, Díaz S, Gallagher R V, Jacobs BF, Kooyman R, Law EA, Leishman MR, Niinemets Ü, Reich PB, Sack L, Villar R, Wang H, Wilf P (2017) Global climatic drivers of leaf size. Science 357:917–921. https://doi.org/10.1126/science.aal4760

Yamori, W, Masumoto, C, Fukayama, H, & Makino, A (2012) Rubisco activase is a key regulator of non-steady-state photosynthesis at any leaf temperature and, to a lesser extent, of steady-state photosynthesis at high temperature. Plant J. 71:871-880. https://doi.org/10.1111/j.1365-313x.2012.05041.x

Yamori W, Hikosaka K, Way DA (2014) Temperature response of photosynthesis in C3, C4, and CAM plants: temperature acclimation and temperature adaptation. Photosynth Res 119:101–117. https://doi.org/10.1007/s11120-013-9874-6

Yemm EW, Willis AJ (1954) The estimation of carbohydrates in plant extracts by anthrone. Biochem J 57:508–514. https://doi.org/10.1042/bj0570508

Yichie Y, Hasan MT, Tobias PA, Pascovici D, Goold HD, Van Sluyter SC, Roberts TH, Atwell BJ (2019) Salt-treated roots of *Oryza australiensis* seedlings are enriched with proteins involved in energetics and transport. Proteomics 19:1900175. https://doi.org/10.1002/pmic.201900175

Zhang CX, Feng BH, Chen TT, Fu WM, Li HB, Li GY, Jin QY, Tao LX, Fu GF (2018) Heat stress-reduced kernel weight in rice at anthesis is associated with impaired source-sink relationship and sugars allocation. Environ Exp Bot 155:718–733. https://doi.org/10.1016/j.envexpbot.2018.08.021

Zhu W, Smith JW, Huang C-M (2010) Mass spectrometry-based label-free quantitative proteomics. J Biomed Biotechnol 2010:840518. https://doi.org/10.1155/2010/840518

Figures & Captions

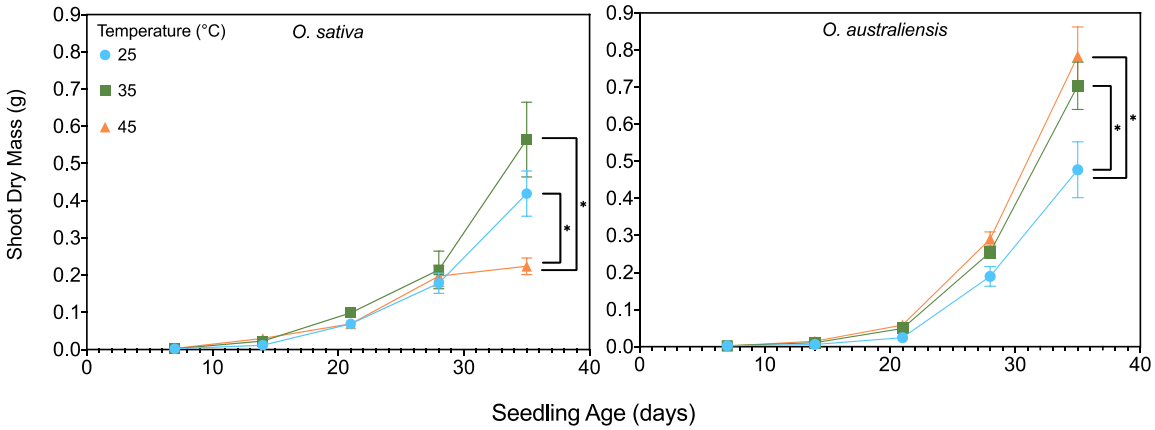


Figure 3.1. Shoot dry mass of *O. australiensis* and *O. sativa* plants grown at 25, 35, and 45˚C from one week of age for four weeks (35 d). Values are means ± SE; $n_{age=7}$ = 20; $n_{age=14-35}$ = 6. Asterisks denote final harvest means that are significantly different within each species as determined by Tukey HSD.

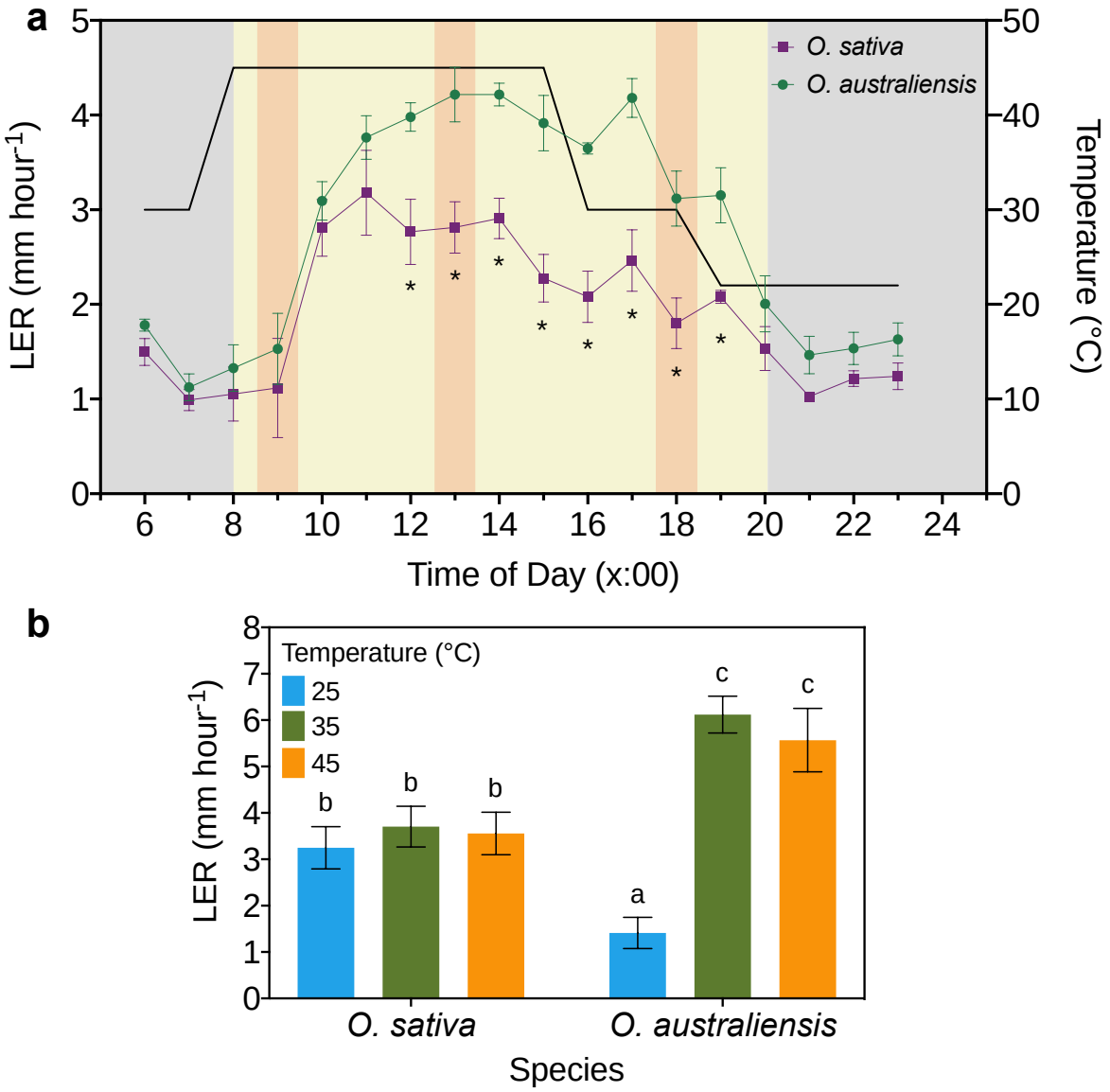


Figure 3.2. Effect of different growth temperatures on the elongation of wild and domestic rice leaves. (A) Mean (*n* = 3; ± SE) leaf elongation rates (LERs) of *O. sativa* and *O. australiensis* plants exposed to 45°C for 8 h (heat 'shock'). Plants were grown at 30°C daytime temperature for six weeks prior to the shock event. The same plants were measured across the sampling period. The solid black line shows the temperature of the growth cabinet. Grey and yellow shading indicates the photoperiod (lights off and on, respectively). Red shading indicates sample times (9:00, 13:00 and 18:00) for tissue collection for sugar and Rca analyses. Asterisks are used to indicate when *O. australiensis* and *O. sativa* LERs are significantly different. (B) Mean (± SE) daily LERs of *O. sativa* and *O. australiensis* plants grown at 25, 35 or 45°C for 5 weeks (*n* for each group is reported in Table 3.1). In a second experiment, plants were measured in morning (8:00 - 12:00) and afternoon (12:00 - 16:00) groups to account for time-of-day effects. LERs were pooled across morning and afternoon groups because no difference was detected between the groups. Means that do not share a letter are significantly different (Duncan's Multiple Range test).

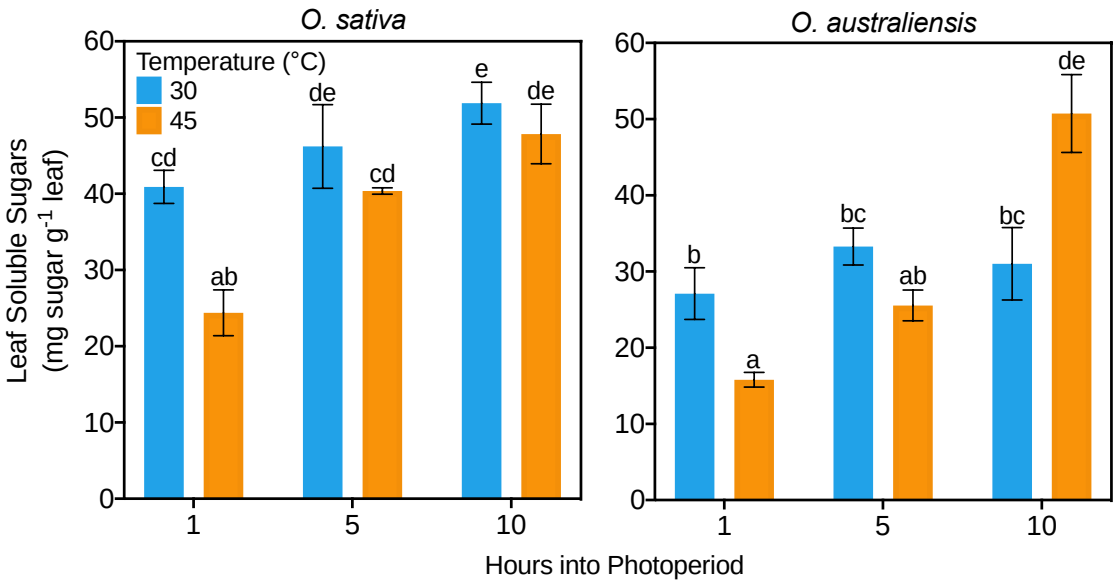


Figure 3.3. The effect of high temperature (shock) on soluble sugar production in the leaves of *O. sativa* and *O. australiensis* plants 1 h (9:00), 5 h (13:00) and 10 h (18:00) into the photoperiod (see Fig 3.1). Plants acclimated to 30°C were exposed to 45°C for the entire photoperiod to induce a heat shock. Values are means ± SE of three pot replicates (different to LVDT replicates). Means that do not share a letter are significantly different (Duncan's Multiple Range test).

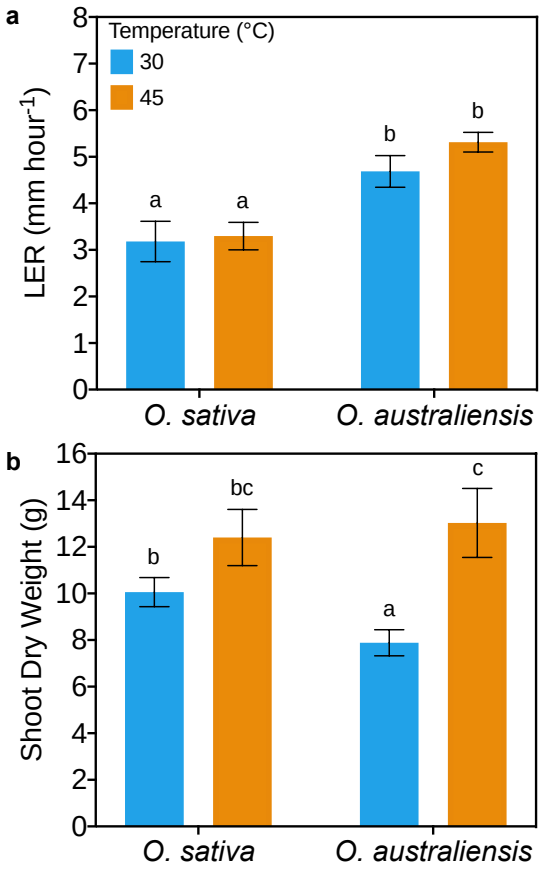


Figure 3.4. Leaf elongation rates (LERs) and shoot dry weights of *O. australiensis* and *O. sativa* plants grown at 30 and 45°C under elevated (700 ppm) atmospheric CO_2 concentrations. (A) LERs are the average elongation rates over 4 h of exposure. This experiment was repeated three times for each group. Values are means ± SE; *n* = 12. (B) Plants (six weeks old at commencement of experiment) were grown for a further six weeks at the indicated conditions before destructive harvesting. Values are means ± SE; n_{30C} = 16, n_{45C} = 4. Means that do not share a letter are significantly different (Duncan's Multiple Range test).

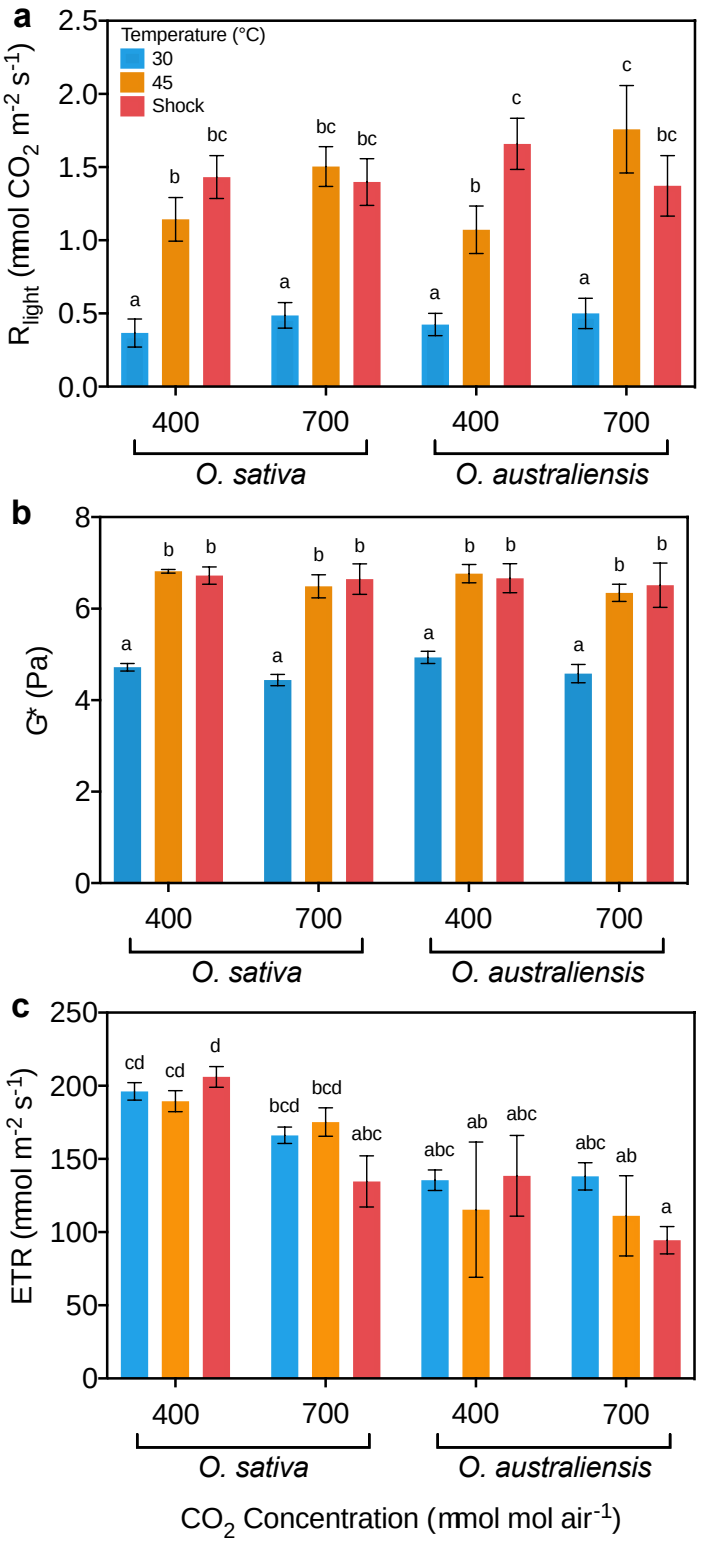


Figure 3.5. Effect of long- and short-term heat exposure on (A) respiration in the light (R_{light}), (B) CO_2 compensation point in the absence of mitochondrial respiration (Γ^*), and (C) electron transport rates (*ETR*) in *O. australiensis* and *O. sativa*. Seven-week-old plants were grown at 30 or 45°C and 400 ppm or 700 ppm for two weeks before measurement. The shock group was established by re-measuring the 30°C plants after one day of exposure to 45°C. Values are means ± SE ($n \geq 5$). These measurements were made using a LI-6800. Means that do not share a letter are significantly different (Duncan's Multiple Range test).

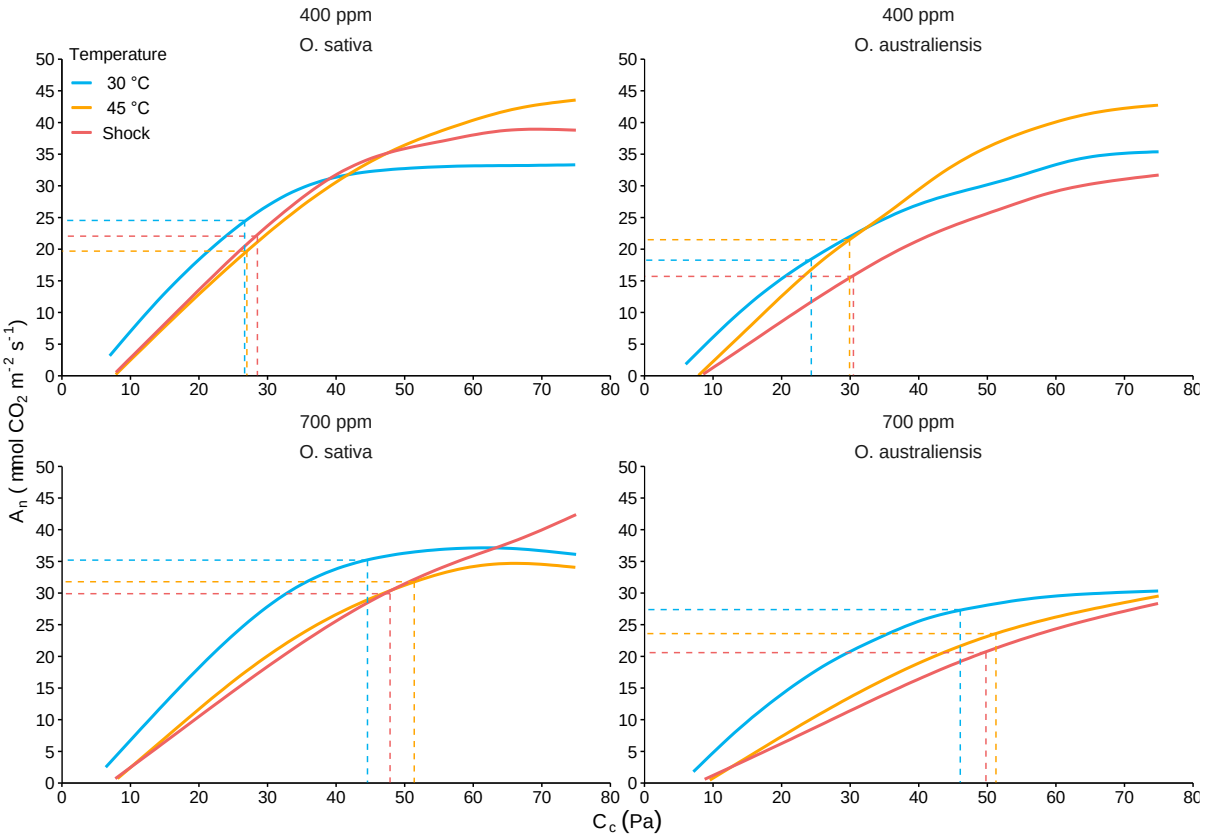


Figure 3.6. Mean (*n* for each group is reported in Table 3.1) RACiR curves for *O. sativa* and *O. australiensis* grown under variable temperature and CO_2 regimes. Seven-week-old plants were grown at 30 or 45°C and 400 ppm or 700 ppm CO_2 for two weeks in growth cabinets before measurement. The “Shock” group was established by raising the 30°C growth cabinet temperature to 45°C for 2 d. C_i values were converted to C_c values. Shading around each response curve shows the 95% Confidence Interval for each group. Dotted lines show the mean C_c values of the plants at their respective growth CO_2 concentrations and corresponding modelled A_n rates. Note: inlayed lines do not represent steady-state assimilation rates and thus only provide an indication CO2 assimilation rate response to the specified growth conditions. RACiR curves were generated using a LI-6800.

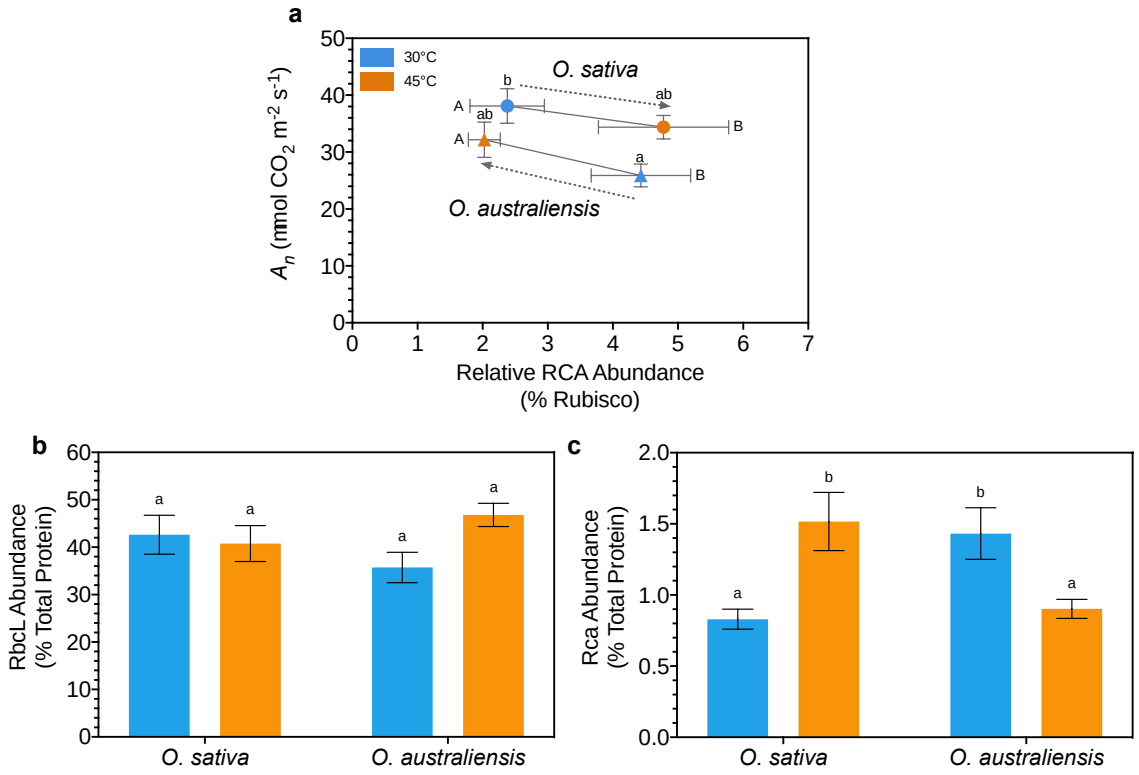
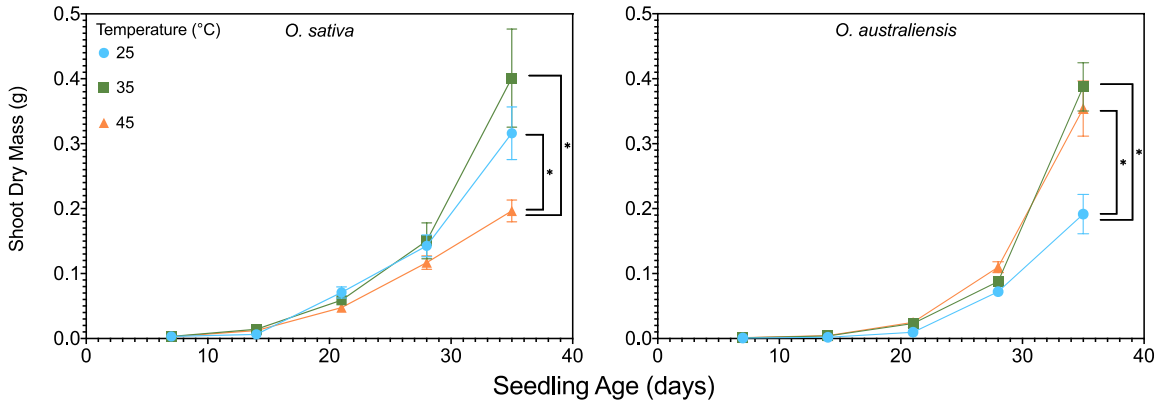
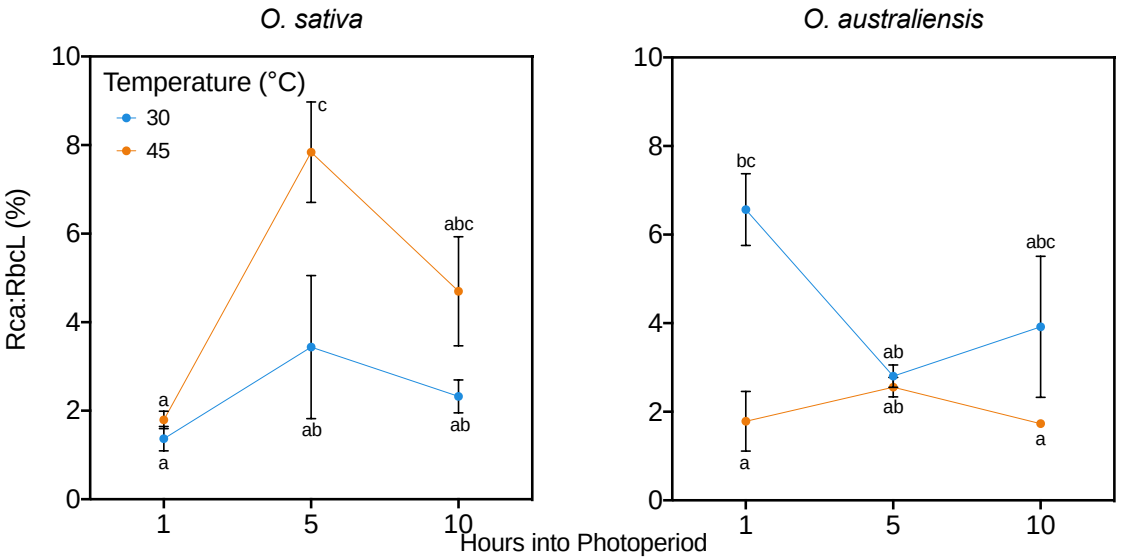


Figure 3.7. Effect of heat on CO_2 assimilation and the abundance of RbcL and Rca in *O. sativa* and *O. australiensis.* (A) Relationship between the abundance of Rca relative to Rubisco (Rca emPAI/RbcL emPAI) and net CO_2 assimilation (A_n) in *O. sativa* and *O. australiensis* leaves exposed to 30 and 45°C (shock). Arrowheads emphasise the relationship between relative Rca abundance and photosynthetic rates for each species. A_n was estimated using a LI-6400. Capital letters show differences in Rca relative abundance; lowercase letters show differences in assimilation rates. (B) Abundance of RbcL relative to total protein emPAI (RbcL emPAI/total emPAI). (C) Abundance of Rca relative to total protein emPAI (Rca emPAI/total emPAI). Values are means ± SE of nine pot replicates (time of day replicates were pooled). Means that do not share a letter are significantly different (Duncan's Multiple Range test). See Figure 3.2 for sampling information.

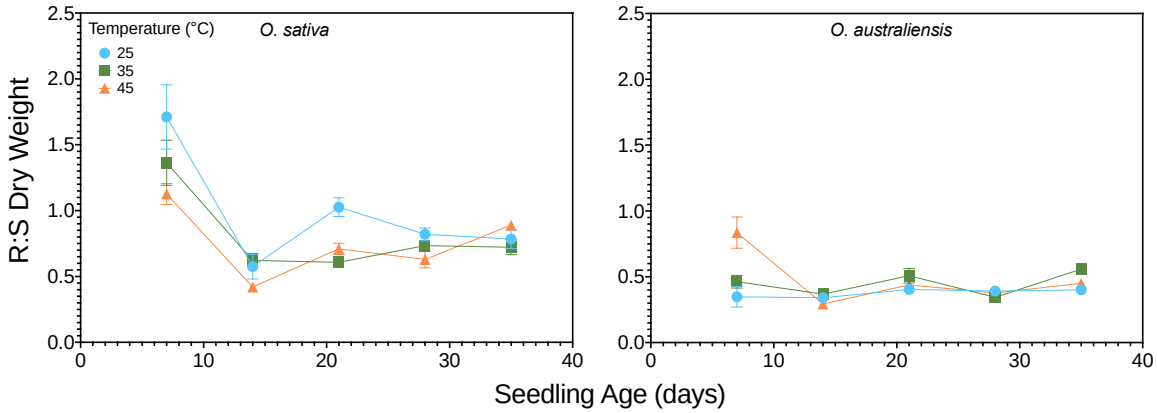
Supplementary Figures



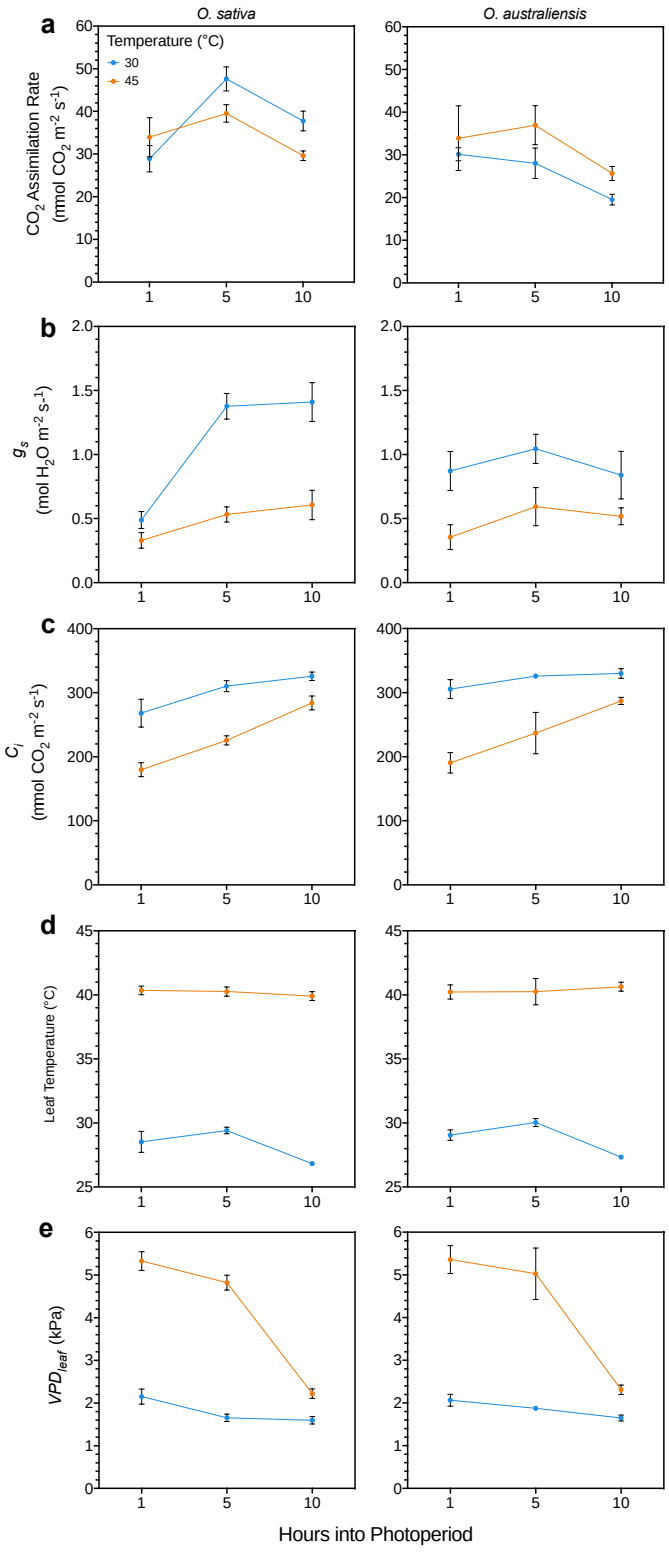
Supplementary Figure 3.8. Root dry weight of *O. australiensis* and *O. sativa* grown at 25, 35 and 45°C.



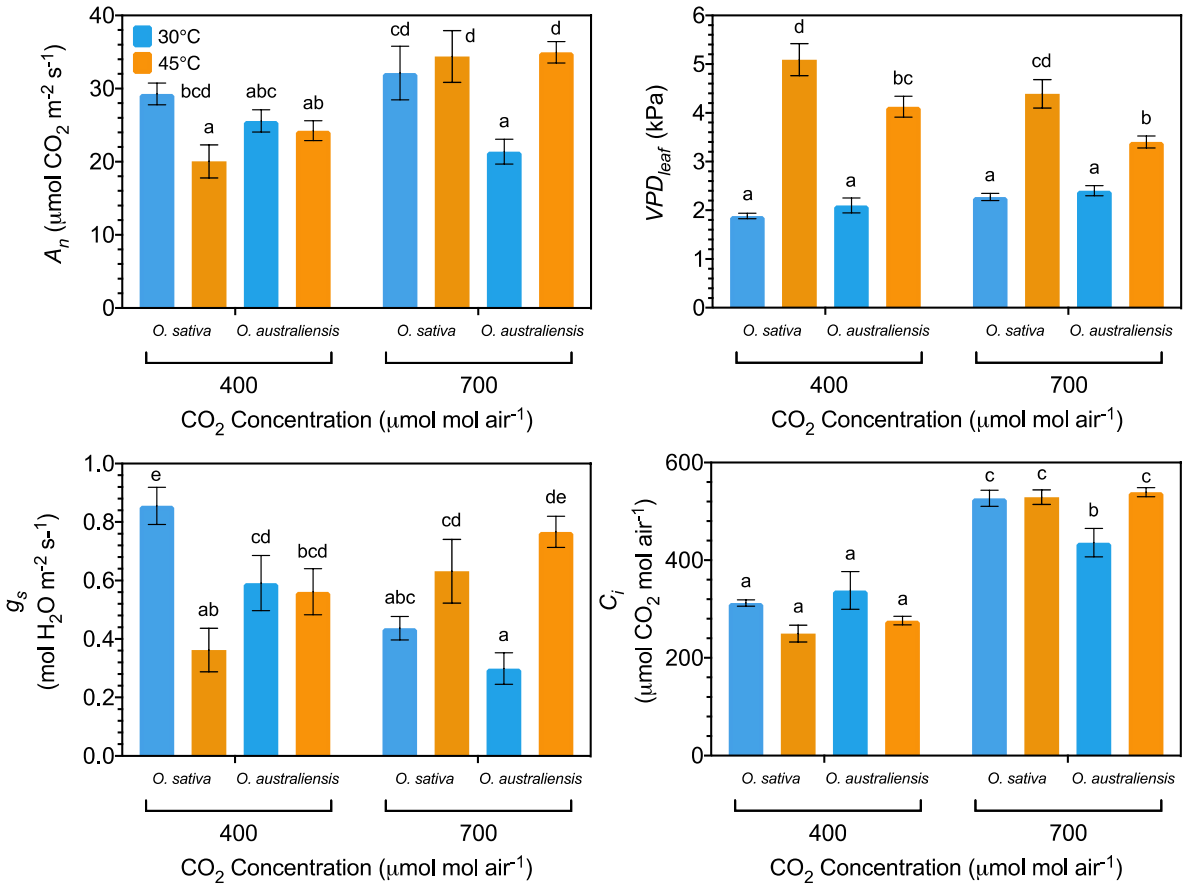
Supplementary Figure 3.9. Rca:RbcL in temperature-treated *O. australiensis* and *O. sativa* leaves throughout the photoperiod.



Supplementary Figure 3.10. Root to shoot ratio (dry weights) of *O. australiensis* and *O. sativa* grown at 25, 35 and 45°C.



Supplementary Figure 3.11. (a) CO2 assimilation rates, (b) Stomatal conductance, (c) intercellular leaf CO_2 concentration, (d) leaf temperature, and (e) vapour pressure deficit of *O. australiensis* and *O. sativa* exposed to 30°C (blue line) or 45°C (orange line).



Supplementary Figure 3.12. CO_2 assimilation, vapour pressure deficit, stomatal conductance, and intercellular CO_2 concentration of *O. australiensis* and *O. sativa* plants grown at 30°C or 45°C and 400 or 700 ppm CO_2 (n = 6).

Plant Molecular Biology
https://doi.org/10.1007/s11103-021-01210-3

Photosynthetic traits of Australian wild rice (*Oryza australiensis*) confer tolerance to extreme daytime temperatures

Aaron L. Phillips[1,3,4] · **Andrew P. Scafaro**[2] · **Brian J. Atwell**[3]

Received: 5 July 2021 / Accepted: 27 October 2021
© Crown 2022

Abstract

Key message **A wild relative of rice from the Australian savannah was compared with cultivated rice, revealing thermotolerance in growth and photosynthetic processes and a more robust carbon economy in extreme heat.**

Abstract Above ~ 32 °C, impaired photosynthesis compromises the productivity of rice. We compared leaf tissues from heat-tolerant wild rice (*Oryza australiensis*) with temperate-adapted *O. sativa* after sustained exposure to heat, as well as diurnal heat shock. Leaf elongation and shoot biomass in *O. australiensis* were unimpaired at 45 °C, and soluble sugar concentrations trebled during 10 h of a 45 °C shock treatment. By contrast, 45 °C slowed growth strongly in *O. sativa*. Chloroplastic CO_2 concentrations eliminated CO_2 supply to chloroplasts as the basis of differential heat tolerance. This directed our attention to carboxylation and the abundance of the heat-sensitive chaperone Rubisco activase (Rca) in each species. Surprisingly, *O. australiensis* leaves at 45 °C had 50% *less* Rca per unit Rubisco, even though CO_2 assimilation was *faster* than at 30 °C. By contrast, Rca per unit Rubisco doubled in *O. sativa* at 45 °C while CO_2 assimilation was slower, reflecting its inferior Rca thermostability. Plants grown at 45 °C were simultaneously exposed to 700 ppm CO_2 to enhance the CO_2 supply to Rubisco. Growth at 45 °C responded to CO_2 enrichment in *O. australiensis* but not *O. sativa*, reflecting more robust carboxylation capacity and thermal tolerance in the wild rice relative.

Keywords Photosynthesis · *Oryza sativa* · *Oryza australiensis* · Heat tolerance · Rubisco · Rubisco activase

Introduction

Climate models predict increasing mean maximum temperatures globally, along with more frequent and intense heatwaves (IPCC 2019). Over the past century, the number of days of extreme heat in northern Australia has risen (CSIRO 2018). As average daily temperatures rise and spasmodic heatwaves become more frequent in agricultural systems, plant productivity declines (Dhir 2018), especially in cropping regions at low latitudes (e.g. India, Africa and Australia; Mahlstein et al. 2011; Jagadish et al. 2012). For example, wheat yield is expected to fall by 6% for every 1 °C increase in average daily temperature (Asseng et al. 2015), with losses of up to 10% per 1 °C possible in other crops (Dhir 2018). Similarly, in rice, grain yield declines by 10% for every 1 °C increase in mean temperature during the dry season (Peng et al. 2004). Heat stress therefore threatens food security (Tilman et al. 2011; Fischer et al. 2014). Despite this knowledge, there is limited research into the impact that long-term heat exposure has on plant physiology, with most studies focussing on transient exposure to high temperatures.

Heat stress affects whole-plant productivity through its impact on carbon metabolism. Extended periods of above-optimal temperatures suppress net carbon gain by impairing photosynthesis relative to dark respiration, reducing whole-plant growth and yield (Wahid et al. 2007; Campbell et al. 2007). Typically, respiration increases with rising temperatures, while photosynthesis declines because

✉ Brian J. Atwell
brian.atwell@mq.edu.au

1 Waite Research Institute and School of Agriculture, Food, and Wine, University of Adelaide, Adelaide, SA, Australia

2 ARC Centre of Excellence in Plant Energy Biology, Research School of Biology, The Australian National University, Canberra, ACT, Australia

3 Department of Biological Sciences, Macquarie University, Sydney, NSW, Australia

4 ARC Centre of Excellence in Plant Energy Biology, School of Agriculture, Food, and Wine, The University of Adelaide, Adelaide, SA, Australia

Published online: 08 January 2022

it is particularly sensitive to heat (Tjoelker 2018). For example, reduced growth of sugarcane at high temperature has been ascribed to lower rates of carbon assimilation and faster respiration (Ebrahim et al. 1998; Wahid et al. 2007; Gomathi et al. 2013). Similarly, the growth rates of domesticated rice (*Oryza sativa*) (Scafaro et al. 2016) and cotton (*Gossypium hirsutum*) (Reddy et al. 1992) leaves were reduced once air temperatures rose above about 35 °C. While respiration acclimates to changes in ambient temperature (Atkin and Tjoelker 2003), acclimation of photosynthesis to heat is less widely accepted (Chi et al. 2013; Benomar et al. 2018), although newly formed rice leaves (cv. IR64) appeared to acclimate to 40 °C days (Rashid et al. 2020). Nonetheless, the limits of thermotolerance may soon be reached, even in warm-climate species (e.g. rice), necessitating the identification of traits that can be used to increase tolerance to high temperature (Atwell et al. 2014; Fischer et al. 2014). A priori, the best candidate species are crop congeners collected from the hottest biomes. Hence, this study investigates an endemic Australian wild species of rice (*O. australiensis*) that has evolved in northern Australia. Scafaro et al. (2016) reported faster instantaneous leaf elongation rates linked to photosynthetic biochemistry during short bursts of heat shock in this species. The mechanistic basis of tolerance to *sustained* high daytime temperatures throughout vegetative development in this heat-tolerant species remains unknown and is addressed in this paper.

Plants adapted to savannahs and deserts have evolved a range of mechanisms to tolerate extreme heat, ranging from development and specialised leaf morphology and physiology (e.g. stomatal responses) (Wright et al. 2017) through to the accumulation of polymorphisms in heat-prone enzymes (e.g. for increased stability at high temperature) and changes in gene expression. This is well-documented in the case of heat shock proteins and their transcription factors (Scharf et al. 2012). Another mechanism for adaptation to heat (> 30 °C) is a more thermally stable Rubisco activase (Rca), a AAA^+ protein that hydrolyses ATP while removing inhibitory sugar phosphates from the active sites of Rubisco, facilitating CO_2 fixation (Portis 2003). Species that express high-temperature-stable isoforms of Rca might minimise impairment of photosynthesis during heat (Feller et al. 1998; Salvucci and Crafts-Brandner 2004; Scafaro et al. 2019). Regulation of in vivo Rca activity and abundance is also likely to be important, for example increasing the expression of endogenous Rubisco and Rca in tandem rescues heat-sensitive *O. sativa* from the deleterious effects of heat on photosynthesis and biomass accumulation (Qu et al. 2021). Despite this, the amount of endogenous Rca protein has not been explored as a pathway to photosynthetic thermotolerance in domestic and wild rice species adapted to a range of thermal regimes.

To understand the effects of extreme temperatures on carbon assimilation and utilisation, we raised atmospheric CO_2 concentrations and measured growth and photosynthetic rates of heat-stressed (long- and short-term stress) *O. sativa* (domestic) and *O. australiensis* (wild) plants. We hypothesised that *O. australiensis* should withstand high daytime temperatures through a more robust photosynthetic metabolism. To assess the heat-resilience of photosynthesis, we used A:C_i curves, point measurements of assimilation rates and leaf elongation rates (LERs). Temperature effects on respiration, photorespiration and electron transport capacity were also characterised. Soluble sugar status of leaves and biomass accumulation in shoots and roots over 4 weeks are reported to establish the relationship between steady-state sugar levels and growth. Finally, we used quantitative proteomics to determine whether there were temperature-dependent changes in total Rca abundance between *O. sativa* and *O. australiensis*. This study presents empirical evidence in support of the superior heat tolerance in *O. australiensis*. We provide further support for photosynthesis being more robust in the wild species under heat and ascribe tolerance to a species-specific response of Rca expression. We also rule out other biochemical/biophysical processes such as respiration and electron transport rate (ETR) as factors that explain the thermotolerance of *O. australiensis*.

Methods

Plant material

Oryza sativa ssp. *japonica* cv. Amaroo and the wild tropical species *O. australiensis* (sourced from Keep River, in the northernmost region of Western Australia; 15° 58′ S, 129° 03′ E) were germinated at 36 °C overnight and sown in 1.65-L pots. The soil, a fine-textured krasnozem (sourced locally from Robertson, NSW, Australia), was kept moist at all times by keeping the pots in shallow trays of water. Plants were watered once a week with a commercial water-soluble fertiliser at a concentration of 1 g L^{-1} (AQUASOL, Yates, Australia). Seedling leaves were sprayed once with dilute, freshly prepared ferrous sulphate solution to prevent iron deficiency.

Heat-shock experiments

In *heat-shock* experiments, plants ($n = 3$) were grown for 6 weeks in a glasshouse under a day/night temperature regime of 30/22 °C, with a midday photosynthetic photon flux density (PPFD) of 1106 ± 13 µmol quanta m^{-2} s^{-1} and a 12-h light period. Unless specified, all measurements and samples were taken over the course of 1 day. Plants were transferred to large growth chambers (model no. PG. 15.

18.9.TD+C [rt]0.3 × 1000.2 × 400.R; Thermoline Scientific Equipment, NSW, Australia) on the day of measurement. Two hours prior to the photoperiod, chambers were set to 30 °C. At the beginning of the photoperiod one cabinet was set to 45 °C (a second cabinet was maintained at 30 °C to act as a control). Unless specified otherwise, the relative humidity in all growth chambers was maintained between 60 and 70% during the photoperiod.

Shoot and root dry biomass in heat-acclimated plants

Seeds of *O. sativa* and *O. australiensis* were germinated and maintained in glasshouse conditions as above for 1 week. After 7 days, *O. sativa* and *O. australiensis* seedlings ($n = 20$) were harvested for baseline measurements of shoot and root fresh and dry mass. Seedlings were exposed to 25, 35 or 45 °C for 4 weeks in growth chambers and harvested weekly ($n = 6$).

To explore the effect of CO_2 fertilisation on *O. sativa* and *O. australiensis*, 6-week-old plants ($n_{30\ °C} = 16$; $n_{45\ °C} = 4$) were maintained in 700 ppm CO_2, and 30 or 45 °C during the light period. Shoot and root dry mass were determined when the plants were 12 weeks old. CO_2 levels were set and maintained using a custom-made CO_2 solenoid (The Canary Company, Lane Cove, Australia).

Leaf elongation rates: short- and long-term response to heat

Seedling growth was tracked through measurements of LER using a HR4000 Linear Variable Displacement Transducer (LVDT) with data logged every 3 min by the program VuGrowth ver. 1.0 (Applied Measurement, Oakleigh, Vic). Three experiments were conducted for the measurement of LERs.

In the first experiment (heat shock), 6-week-old plants ($n = 3$) at the four- to five-leaf stage were moved to a growth cabinet (containing the LVDT apparatus) set to the conditions described above. The youngest, fully expanded leaf of each plant was measured for 12 h and the same six plants (three per species) were used throughout the measurement period. Heat shock (45 °C) was imposed as described above. Leaf growth was measured for 8 h at 45 °C, and for 2 h at 30 °C. LERs were calculated using increase in leaf lengths for each 1-h period. At 1, 5 and 10 h into the photoperiod leaves from these plants were used for point measurements of photosynthesis, sugar and Rca analysis.

In the second experiment, plants ($n = 5$–8) grown at 25, 35, or 45 °C for 4 weeks were moved sequentially into the LVDT chamber. Measurement of LER commenced 1 h into the photoperiod. Leaf elongation was measured for 8 h (8:00–16:00) and LERs were calculated as the rate of elongation over that time.

In the third experiment, the CO_2 level in the LVDT chamber was increased to 700 ppm. Plants ($n = 4$) of *O. sativa* and *O. australiensis*, which had been grown at 30 or 45 °C and 700 ppm CO_2 for 6 weeks, were moved (independently) to the high-CO_2 LVDT chamber for LER measurements and measured under their respective growth conditions. Measurements were made as above and were replicated three times (total $n = 12$).

Sugar extraction and determination

Leaves of *O. sativa* and *O. australiensis* grown at 30 °C or exposed to 45 °C as a shock treatment were used for sugar analysis (see above). One fully expanded, healthy leaf (~ 200 mg) was excised from plants at each time point (1, 5 and 10 h into the light period; $n = 3$), snap-frozen in liquid nitrogen and stored at − 80 °C. Soluble sugar content was determined following the anthrone method (Yemm and Willis 1954) using glucose standards (25, 50, 75 and 100 mg glucose into 1000 mL H_2O).

Gas exchange: rapid A:Ci response (RACiR) curves, Laisk curves and point measurements

Oryza sativa and *O. australiensis* used for the generation of RACiR curves were grown in glasshouses (as described above) for 7 weeks. Plants ($n = 4$–6) were acclimated to the desired conditions in growth cabinets (Model BDR16; Conviron, Manitoba, Canada) set to 30 or 45 °C and 400 ppm or 700 ppm CO_2 for 2 weeks prior to measurement (plants were 9 weeks old when measured). Plants grown at 30 °C were exposed to 45 °C following their measurement in the steady state to form the heat shock group.

RACiR curves were made following the procedure of Stinziano et al. (2017) and Stinziano et al. (2019) on a LI-6800 (Li-Cor, Lincoln, USA) with modifications. Briefly, to measure CO_2 assimilation (A_n), conditions in the IRGA head were set to match the conditions of the growth chamber (block temperature of 30 or 45 °C; reference CO_2 level of 400 or 700 ppm), with a flow rate of 500 µmol air s^{-1}, relative humidity of ~ 70%, a fan speed of 10,000 rpm, with a photosynthetic photon flux density (PPFD) of 1500 µmol quanta m^{-2} s^{-1}. For curve generation, reference CO_2 was reduced to 10 µmol mol^{-1} air and increased to 1010 µmol mol^{-1} air over a 10-min period (ramping = 100 µmol CO_2 mol^{-1} air min^{-1}). Measurements were taken by the LI-6800 every two seconds. All RACiR curves were corrected with empty chamber measurements. Leaves used for RACiR measurements were also used for Laisk curves and fluorescence measurements.

RACiR curves were initially converted from intercellular CO_2 partial pressure (C_i) to chloroplastic CO_2 partial pressure (C_c) using mesophyll conductance (g_m) and its temperature response for *O. sativa* and *O. australiensis* as reported by Scafaro et al. (2016). The subsequent *A*-C_c (where *A* is the CO_2 assimilation rate measured by the LI-6800) curves were analysed using the standard C_3 photosynthesis model (Farquhar et al. 1980),

$$A_c = V_{cmax} \times \frac{C_c - \Gamma^*}{C_c + K_{air}} - R_{light}$$

$$A_j = J_{max} \times \frac{C_c - \Gamma^*}{4C_c + 8\Gamma^*} - R_{light}$$

where A_c (carboxylation-limited CO_2 assimilation rate) was fit to C_c partial pressures in the linear phase (corresponding to a mean C_c value of 38 Pa) and A_j (electron-transport limited CO_2 assimilation rate) fit to C_c partial pressures above this point. V_{cmax} (maximum rate of carboxylation) and J_{max} (maximum rate of electron transport) were iteratively fit using a non-linear least-squares fit (R statistical). The Michaelis–Menten coefficient of Rubisco in air (K_{air}) and its temperature response (value at 25 °C = 30.5 kPa; activation energy = 60.5 kJ mol^{-1}) came from previous in vitro measurements in *O. sativa* (Hermida-Carrera et al. 2016). We assumed no difference in Rubisco kinetics between the two species. Respiration in the light (R_{light}) and the CO_2 compensation point in the absence of mitochondrial respiration (Γ^*) were those measured by the Laisk method (Laisk 1977) (see below).

Laisk curves ($n = 5$–6) were generated with a LI-6800 (Li-Cor, Lincoln, USA). The curves were made with a C_i range of 50, 75, 100 and 120 ppm at three different light levels (PPFD = 100, 200 and 400 µmol quanta m^{-2} s^{-1}). The intersection of the three lines was used to estimate R_{light} and Γ^*.

Further, point measurements of gas exchange were made on *O. sativa* and *O. australiensis* using a LI-6400 (Li-Cor, Lincoln, USA) at saturating light levels (1500 µmol quanta m^{-2} s^{-1}). All measurements were made 1, 5 and 10 h into the photoperiod on plants ($n = 3$) acclimated to 30 °C growth chambers (see above). Heat shock was imposed as described above. The CO_2 in the reference chamber was set to 400 µmol CO_2 mol^{-1} air and the block temperature was set to 30 or 45 °C.

Chlorophyll fluorescence: dark and light adapted leaves

Following the RACiR and Laisk curves, the same leaves ($n = 5$–6) were used for determining chlorophyll fluorescence parameters as per instructions (see *Licor Support guidelines*; accessed September 2018). Leaves of interest were wrapped in aluminium foil and the lights in the growth cabinets were switched off. The plants were kept in the dark for 15 h (5:00 pm–8:00 am) before dark-adapted fluorescence measurements were made. Foil was removed from each leaf in the dark and the leaf was inserted into the chamber of a LI-6800 fitted with a fluorescence head. Minimal fluorescence (F_o) and maximal fluorescence (F_m) were measured with a rectangular flash of actinic light. Variable fluorescence (F_v; estimated from F_o and F_m) was used to calculate the maximal efficiency of photosystem II (PSII; F_v/F_m).

Dark-adapted leaves were then exposed to 1500 µmol quanta m^{-2} s^{-1} for 1 h to induce light acclimation. A Multiphase Flash™ (MPF) of actinic light was used in combination with a 'dark pulse' flash (i.e. a pulse of infrared light) to estimate F_s and F_m'. MPF is a method for more accurately determining F_m' in light-adapted leaves, while the dark pulse allows for the calculation of F_o' (the minimal fluorescence in the dark of a light-adapted leaf; Loriaux et al 2013; Avenson and Saathoff 2018). These parameters were used to calculate the photochemical yield of PSII (Φ_{PSII}) and the efficiency of PSII energy harvesting in the 'open' (oxidised) state (F_v'/F_m'). ETRs were derived from these values by the LI-6800.

Quantifying relative abundance of RbcL and Rca by tandem mass spectrometry (MS/MS)

Oryza sativa and *O. australiensis* leaves (~200 mg; $n = 3$) sampled 1, 5 and 10 h into the photoperiod (see above) were ground to a fine powder in liquid nitrogen. Leaf powder was washed in ice-cold 10% TCA/0.07% DTT in acetone and incubated at -20 °C for 1 h. Following centrifugation at 4000×*g* for 45 min, supernatant was discarded and the pellet was washed with 0.07% DTT in acetone and stored at – 20 °C overnight. Samples were centrifuged at 4000×*g* for 45 min at 5 °C and the supernatant was discarded. Following resuspension in 0.07% DTT in acetone, samples were incubated at − 20 °C for 1 h. These washes were repeated twice. Pellets were air-dried, suspended in extraction buffer (250 mM TEAB/1.5% SDS), and stored at − 20 °C overnight. Pellets (kept on ice) were homogenised six times with a Precellys24 tissue homogeniser (Bertin Instruments, Montigny-le-Bretonneux, France) set to 6500 rpm for 20 s. Homogenised samples were centrifuged at 4000×*g* for 1 h at 10 °C. Total protein concentration of the supernatant was determined using a Direct Detect infra-red spectrometer (Merck Millipore, Darmstadt, Germany).

100 µL of 250 mM TEAB was added to 20 µg of sample. DTT was added to a final concentration of 10 mM. This mixture was incubated at 60 °C for 1 h. IAA was added to a final concentration of 20 mM, and samples were incubated at room temperature for 1 h in the dark. Trypsin (0.1 µg μL^{-1}) was added in a 1:50 (trypsin:sample) ratio and incubated

overnight at 37 °C. Digested samples were dried in a SpeediVac and resuspended in 250 mM TEAB. Pierce Detergent Removal spin columns (Thermo Scientific, IL, USA) were used to remove residual SDS. Samples were dried again in a SpeediVac and stored at − 20 °C. Prior to MS/MS analysis, samples were resuspended in loading buffer (2% acetonitrile (ACN); 0.1% formic acid).

Each sample (10 µL containing 3 µg of peptides) was injected onto a peptide trap (Bruker peptide Captrap) for preconcentration and desalted with 0.1% formic acid, 2% ACN, at 10 µL min^{-1} for 5 min. The peptide trap was then switched into line with the analytical column. Peptides were eluted from the column using linear solvent gradients, with steps, from mobile phase A (0.1% formic acid): mobile phase B (99.9% ACN/0.1% formic acid) (98:2) to (90:10) for 10 min, then to (65:35), at 600 nL min^{-1} over a 78-min period. After peptide elution, the column was cleaned with 95% buffer B for 10 min and then equilibrated with 98% buffer A for 20 min before the next sample injection. The reverse phase nanoLC eluent was subject to positive ion nanoflow electrospray analysis in an Information Dependent Acquisition (IDA) mode.

In the IDA mode, a TOF–MS survey scan was acquired (m/z 350–1500, 0.25 s), with the ten most intense multiply charged ions (2 + –5 + ; counts > 150) in the survey scan sequentially subjected to MS/MS analysis. MS/MS spectra were accumulated for 50 ms in the mass range m/z 100–1500 with rolling collision energy.

Data analysis

Data analysis was done in R (R Core Team 2017). All data, unless specified otherwise, were analysed by two- or three-way ANOVAs and Tukey HSD or Duncan's Multiple Range tests. Figures were produced using R and GraphPad Prism (version 9).

The data collection and processing for mass spectrometry was conducted by the Australian Proteomics Analysis Facility (APAF; Macquarie University, Sydney, Australia). The LC–MS/MS data of the IDA runs of each sample were searched using ProteinPilot (v4.2; AB Sciex) in 'thorough' mode. MGF files (peak list) were exported, submitted to Mascot (Matrix Science, UK) and searched against NCBInr *O. sativa* (rice) database for Rca. Relative abundances for Rca and the large subunit of Rubisco (RbcL) were determined across samples using emPAI values, which were normalised against emPAI values of the total protein pools in *O. sativa* and *O. australiensis* (see Ishihama et al. (2005) and Zhu et al. (2010) for details on the emPAI method).

Results

Sequential destructive harvests

There was a significant three-way interaction between harvest time, temperature and species in shoot dry weight ($p < 0.0001$; Fig. 1). Follow-up two-way ANOVAs for each harvest time revealed that shoot dry weights of the two species first diverged after 35 d of exposure to 45 °C ($p < 0.0001$). *Oryza sativa* had significantly reduced shoot growth at 45 °C compared with plants growing at 25 and 35 °C. By contrast, *O. australiensis* shoots were significantly lighter at 25 than at 35 and 45 °C. Similar trends were seen in the root dry weight (Fig. S1), demonstrating slowest growth of *O. australiensis* was consistently at 25 °C whereas

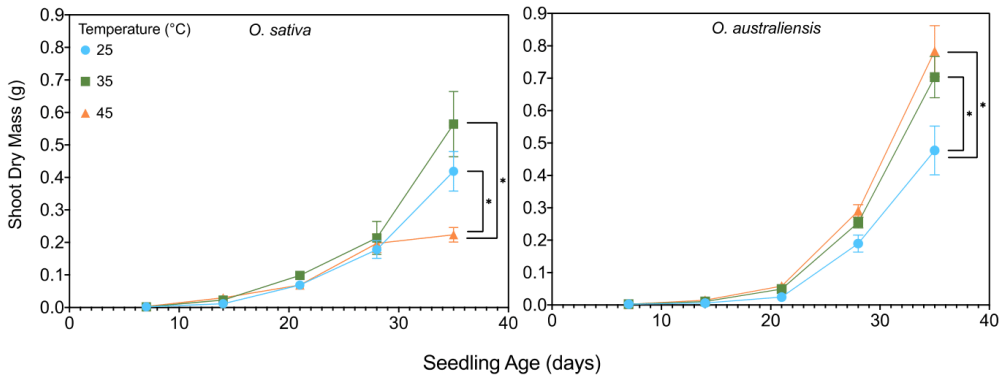


Fig. 1 Shoot dry mass of *O. sativa* and *O. australiensis* plants grown at 25, 35, and 45 °C from 1 week of age for 4 weeks (35 days). Values are means ± SE; $n_{age=7} = 20$; $n_{age=14–35} = 6$. Asterisks denote final-harvest means that are significantly different within each species as determined by Tukey HSD

in *O. sativa*, plants grew slowest at 45 °C. The impact of the 45 °C treatment on the growth of *O. sativa* shoots became severe after 28 days. Similarly, differences in root mass were only detected in both species following 21 days exposure to treatment conditions.

Leaf elongation rates: short term vs. long term response to temperature

A two-way repeated-measures ANOVA was conducted to determine the effect of heat shock on LERs in *O. sativa* and *O. australiensis* over a 17-h period (Fig. 2A). There was a significant interaction between species and time of day ($p < 0.01$), indicating that the two species responded to heat differentially throughout the day. For example, LER at 11:00 was ~ 3.17 mm h^{-1} (*O. sativa*) and ~ 3.76 mm h^{-1} (*O. australiensis*); not significantly different. However, after 8 h of heat treatment (15:00), there was a significant ~ 28% reduction in LER in *O. sativa*, while a non-significant increase was registered in *O. australiensis* (~ 4% increase). LERs of the two species remained significantly different from this point until the end of the photoperiod.

After long-term exposure to 35 and 45 °C, LERs in *O. australiensis* but not *O. sativa* increased ($p < 0.01$; Fig. 2B). This interaction was driven by a significant increase in LER for *O. australiensis* when grown at 35 or 45 °C compared with 25 °C ($p < 0.01$) but this was not seen for *O. sativa*.

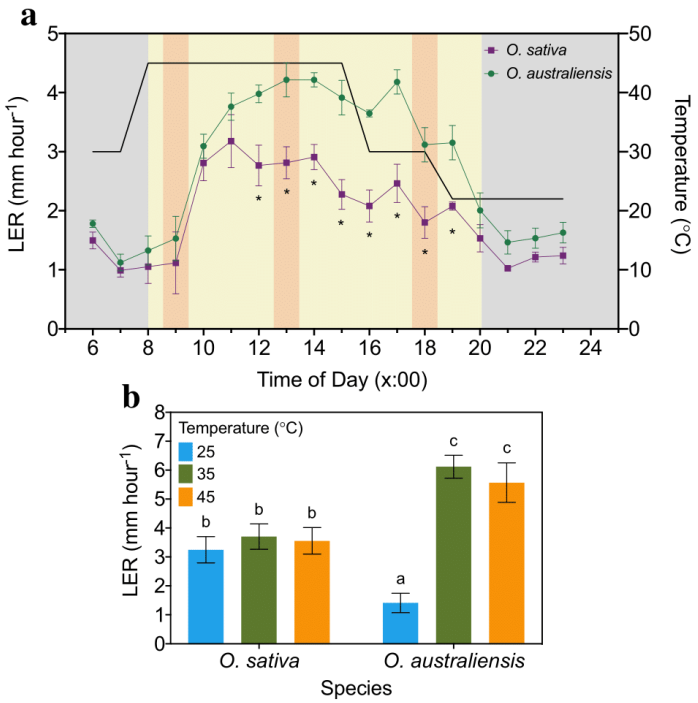


Fig. 2 Effect of different growth temperatures on the elongation rates of leaves of wild and domestic rice. **A** Mean ($n = 3; \pm$SE) leaf elongation rates (LERs) of *O. sativa* and *O. australiensis* plants exposed to 45 °C for 8 h (heat 'shock'). Plants were grown at 30 °C daytime temperature for 6 weeks prior to the shock event. The same plants were measured across the sampling period. The solid black line shows the temperature of the growth cabinet. Grey and yellow shading indicates the photoperiod (lights off and on, respectively). Red shading indicates sample times (9:00, 13:00 and 18:00) for tissue collection for sugar and Rca analyses. Asterisks are used to indicate when *O. sativa* and *O. australiensis* LERs are significantly different. **B** Mean (±SE) daily LERs of *O. sativa* and *O. australiensis* plants grown at 25, 35 or 45 °C for 5 weeks (*n* for each group is reported in Table 1). In a second experiment, plants were measured in morning (8:00–12:00) and afternoon (12:00–16:00) groups to account for time-of-day effects. LERs were pooled across morning and afternoon groups because no difference was detected between the groups. Means that do not share a letter are significantly different (Duncan's Multiple Range test)

Soluble sugars

A significant three-way interaction showed that the effect of temperature on leaf sugar levels was different for the two species depending on the time of day (Fig. 3; $p = 0.036$). Follow up two-way ANOVAs showed that the interaction between species and temperature ($p = 0.03$) changed throughout the day. That is to say, the effect of temperature on sugar accumulation was different for each species, and this relationship changed throughout the day. This reflects the large increase in soluble sugar concentration in *O. australiensis* leaves at 45 °C 10 h into the photoperiod ($p = 0.02$), indicating an accumulation of sugar in spite of the high temperature. In contrast, the effect of heat on leaf sugar levels was the same for both species 1 h and 5 h into the photoperiod. After 10 h at 45 °C the concentration of sugars in *O. australiensis* leaves equalled that of *O. sativa* leaves.

Effect of CO_2 fertilisation on growth and LER

Contrary to the significantly faster LER of *O. australiensis* leaves when grown at 45 °C in ambient (400 ppm) CO_2 conditions (Fig. 2B), when CO_2 levels were raised to 700 ppm the LER of *O. australiensis* was identical at 30 and 45 °C (Fig. 4A). The LER of *O. sativa* was consistently ~ 3 mm h^{-1} at 400 ppm (Fig. 2B) and 700 ppm CO_2 (Fig. 4A) and was unaffected by atmospheric CO_2 level. Despite 700 ppm CO_2 not further accelerating the rapid leaf elongation rate when *O. australiensis* grew at 45 °C, shoot dry mass increased significantly at 45 °C in an elevated CO_2 atmosphere ($p < 0.01$; Fig. 4B), as was observed for biomass increments at 400 ppm (Fig. 1). On the contrary, shoot dry mass was unaffected by heat when *O. sativa* grew at 700 ppm CO_2 (Fig. 4B), while growth was slower at 45 °C in 400 ppm CO_2 (Fig. 1). That is, a 700 ppm CO_2 atmosphere reversed the impaired growth rates seen in *O. sativa* at 45 °C in 400 ppm CO_2.

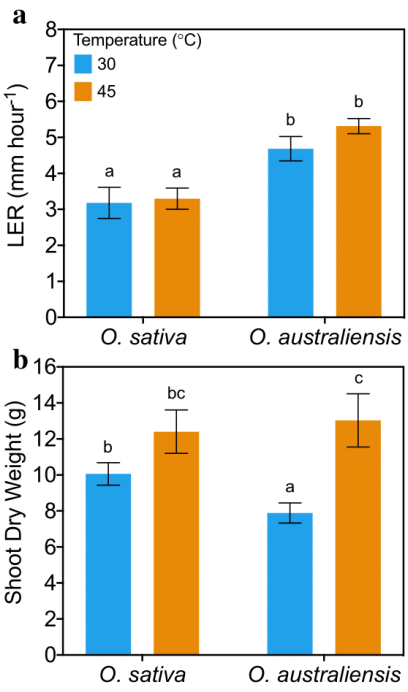


Fig. 4 Leaf elongation rates (LERs) and shoot dry weights of *O. sativa* and *O. australiensis* plants grown at 30 and 45 °C under elevated (700 ppm) atmospheric CO_2 concentrations. **A** LERs are the average elongation rates over 4 h of exposure. This experiment was repeated three times for each group. Values are means ± SE; $n = 12$. **B** Plants (6 weeks old at commencement of experiment) were grown for a further 6 weeks at the indicated conditions before destructive harvesting. Values are means ± SE; $n_{30C} = 16$, $n_{45C} = 4$. Means that do not share a letter are significantly different (Duncan's Multiple Range test)

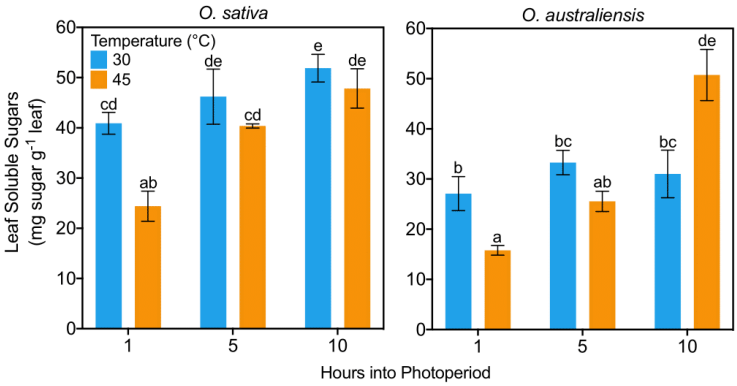


Fig. 3 The effect of high temperature (shock) on soluble sugar production in the leaves of *O. sativa* and *O. australiensis* plants 1 h (9:00), 5 h (13:00) and 10 h (18:00) into the photoperiod (see Fig. 1). Plants acclimated to 30 °C were exposed to 45 °C for the entire photoperiod to induce a heat shock. Values are means ± SE of three pot replicates (different to LVDT replicates). Means that do not share a letter are significantly different (Duncan's Multiple Range test)

Net CO_2 assimilation; R_{light}, Γ* and ETR

Respiration in the light (R_{light}) was similar for *O. sativa* and *O. australiensis* (Fig. 5A). Both species had a significantly faster R_{light} with rising temperature, and the stimulation of R_{light} was not significantly different between the species. The stimulation of R_{light} was the same irrespective of a long-term heat exposure or heat-shock, suggesting limited acclimation potential of R_{light} for rice or its wild relative. The CO_2 compensation point of leaves in the absence of respiration (Γ^*)—equivalent to the amount of CO_2 released by leaves via the process of photorespiration—increased significantly with temperature (Fig. 5B). Again, the control 30 °C values and the extent of temperature stimulation was not significantly different between the two species or between the long-term heat exposure and heat shock. Photosynthetic *ETR* determined by chlorophyll fluorescence was significantly different between *O. sativa* and *O. australiensis* irrespective of growth temperature (Fig. 5C), with *O. sativa* having significantly faster *ETR* under all treatments. While R_{light} and Γ^* did not perform differently with 700 ppm CO_2 enrichment, the *ETR* of both species declined somewhat with heat shock, though perhaps more so for *O. australiensis*.

Net CO_2 assimilation; RACiR curves

To further characterise the response of photosynthesis in *O. sativa* and *O. australiensis* to temperature, RACiR curves were made and analysed using the standard C_3 photosynthesis model (Fig. 6, Table 1). CO_2 carboxylation was the prevailing limitation on A_n at 400 ppm for both species. Consistent with growth and leaf sugar accumulation, at an ambient CO_2 of 400 ppm *O. australiensis* had faster and *O. sativa* slower A_n after acclimation to 45 °C, relative to 30 °C (Fig. 6). The beneficial response of A_n at 45 °C relative to 30 °C for *O. australiensis* at 400 ppm (Fig. 6) can in part be attributed to overcoming the inherently lower g_s, C_i and subsequent C_c concentrations in *O. australiensis* at 30 °C (Table 1). Interestingly, comparisons of long-term 45 °C heat exposure versus heat-shocked plants revealed reduced photosynthetic capacity of *O. australiensis* in heat-shocked plants only, while similar reductions in capacity were observed in *O. sativa* irrespective of exposure time (Table 1). This may suggest potential for acclimation of CO_2 assimilation to heat in the wild rice but not in *O. sativa*.

In plants acclimated to a [CO_2] of 700 ppm, overall photosynthetic capacity was significantly reduced for *O. australiensis* at 45 °C. This was demonstrated by a lower A_n and substantially reduced V_{cmax} and J_{max} at 45 °C and 700 ppm CO_2 relative to 400 ppm (Fig. 6; Table 1). In contrast, *O. sativa* had an increase in A_n and a more limited suppression of V_{cmax} and J_{max} when grown at 45 °C and exposed to 700 ppm. For example, at 45 °C

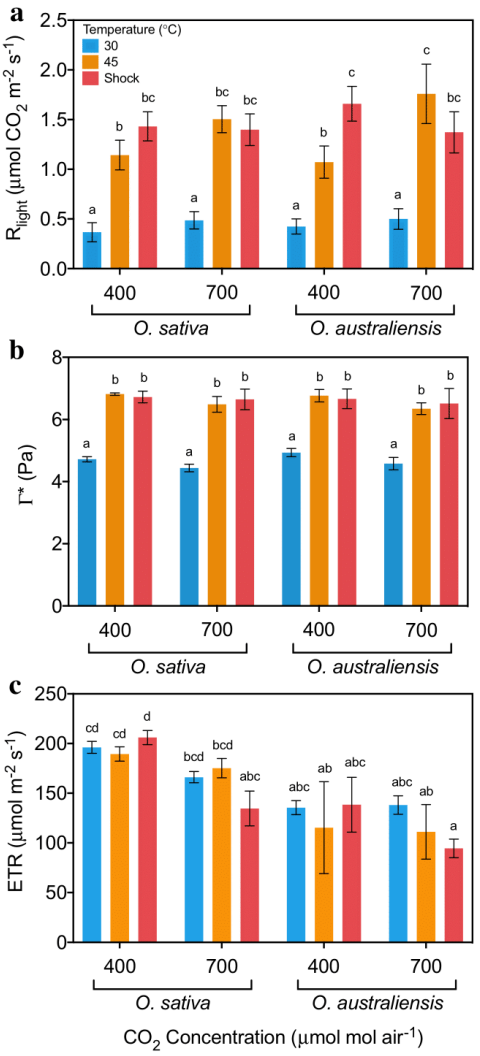


Fig. 5 Effect of long- and short-term heat exposure on **A** respiration in the light (R_{light}), **B** CO_2 compensation point in the absence of mitochondrial respiration (Γ^*), and (C) electron transport rates (*ETR*) in *O. sativa* and *O. australiensis*. Seven-week-old plants were grown at 30 or 45 °C and 400 ppm or 700 ppm for 2 weeks before measurement. The shock group was established by re-measuring the 30 °C plants after one day of exposure to 45 °C. Values are means ± SE ($n \geq 5$). These measurements were made using a LI-6800. Means that do not share a letter are significantly different (Duncan's Multiple Range test)

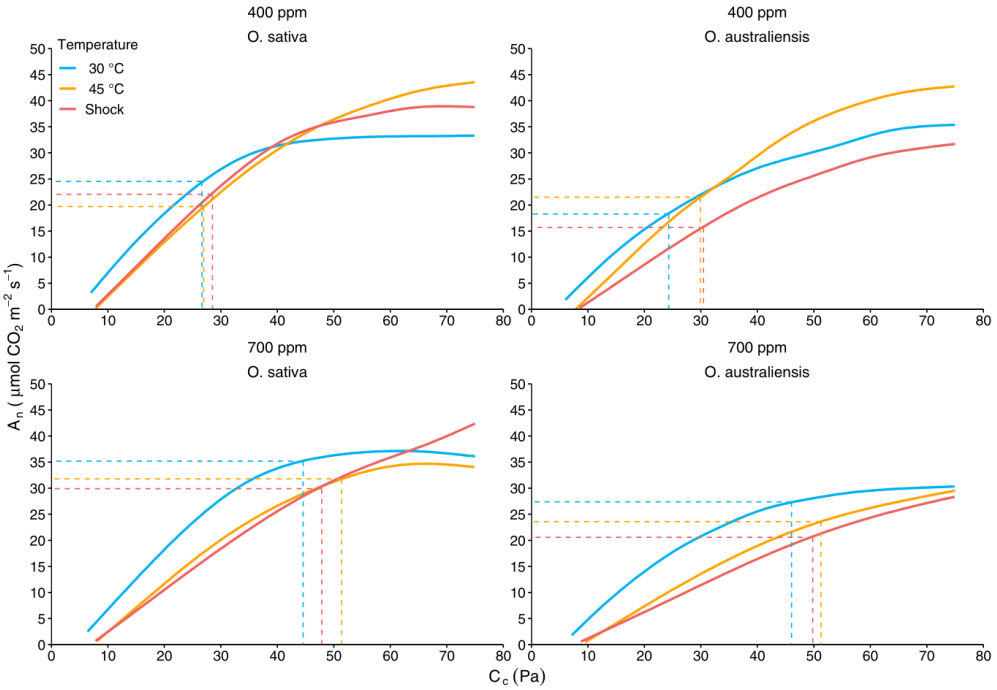


Fig. 6 Mean (*n* for each group is reported in Table 1) RACiR curves for *O. sativa* and *O. australiensis* grown under variable temperature and CO_2 regimes. Seven-week-old plants were grown at 30 or 45 °C and 400 ppm or 700 ppm CO_2 for 2 weeks in growth cabinets before measurement. The "Shock" group was established by raising the 30 °C growth cabinet temperature to 45 °C for 2 days. C_i values were converted to C_c values. Shading around each response curve shows the 95% Confidence Interval for each group. Dotted lines show the mean C_c values of the plants at their respective growth CO_2 concentrations and corresponding modelled A_n rates. Note: inlayed lines do not represent steady-state assimilation rates and thus only provide an indication of CO_2 assimilation rate response to the specified growth conditions. RACiR curves were generated using a LI-6800

and 700 ppm the V_{cmax} of *O. australiensis* was 66% of the rate at 400 ppm, while for *O. sativa*, V_{cmax} remained at 93% of the rate at 400 ppm. A greater downregulation of photosynthetic capacity under heat and high CO_2 was therefore seen in *O. australiensis* compared with *O. sativa*. Significant suppression of photosynthetic capacity of *O. australiensis* exposed to high CO_2 and heat was accompanied by greater photosynthetic efficiency under these conditions, and therefore little change in A_n was observed (Table 1).

Relative abundance of Rca as a percentage of Rubisco

The effect of temperature on the relative abundance of Rca as a proportion of Rubisco for *O. sativa* and *O. australiensis* was determined (Fig. 7). At 45 °C the abundance of Rca as a proportion of RbcL was significantly reduced in *O. australiensis*. In spite of this, A_n increased in *O. australiensis*. The opposite was observed for *O. sativa*—at 45 °C the abundance of Rca as a proportion of RbcL was significantly increased, while photosynthetic rates declined.

There was a significant interaction between species and temperature ($p < 0.001$). At 30 °C, the concentration of Rca relative to Rubisco in *O. australiensis* leaves was ~ 2.2

Table 1 Photosynthetic characteristics of wild and domestic rice grown under variable temperature and CO_2 regimes, derived from RACiR curves

Temperature (°C)	30				45				45 (Shock)			
[CO_2] (ppm)	400		700		400		700		400		700	
Species	*O. sat*	*O. aus*	*O. sat*	*O. aus*	*O. sat*	*O. aus*	*O. sat*	*O. aus*	*O. sat*	*O. aus*	*O. sat*	*O. aus*
V_{Cmax} (μmol CO_2 m^{-2} s^{-1})	99.4 ± 7.1 a b	90.2 ± 3.4 ab	102.1 ± 11.4 ab	78.6 ± 1 a	185.2 ± 10.1 c	172.4 ± 18.4 c	167.3 ± 16.5 c	125.2 ± 8.4 b	180.6 ± 14 c	125.3 ± 15.4 b	167.6 ± 6.3 c	111.6 ± 16 a b
A_n/V_{Cmax} (%)	24.5	20.2	34.5	34.8	10.6	12.5	19	18.8	12.2	12.5	17.8	18.4
J_{max} (μmol m^{-2} s^{-1})	179.2 ± 11 abcde	164.1 ± 5.8 abc	186.2 ± 19.3 abcdef	145.7 ± 3.0 ab	226 ± 6.9 ef	215.3 ± 22.6 def	194.2 ± 16.9 bcdef	145.2 ± 11.3 ab	217.8 ± 16.9 f	165 ± 16.8 abcd	204.7 ± 17.8 cdef	132 ± 16.6 a
C_c (Pa)	26.6 ± 0.9 ab	24.3 ± 0.5 a	44.6 ± 2.7 c	46.1 ± 2 cd	27.0 ± 0.8 ab	29.9 ± 0.3 b	51.4 ± 1.3 e	51.3 ± 1.6 e	28.5 ± 0.2 ab	30.5 ± 0.6 b	47.9 ± 1.6 cde	49.8 ± 2.8 de
g_m (μmol CO_2 m^{-2} s^{-1} Pa^{-1})	10.9 ± 0.1 bcd	5.5 ± 0.07 a	11.2 ± 0.11 bcd	5.4 ± 0.06 a	22.1 ± 0.46 f	10.5 ± 0.26 bc	21.6 ± 0.12 f	11.3 ± 0.32 cd	20.6 ± 0.41 e	10.2 ± 0.26 b	22.4 ± 0.58 f	11.7 ± 0.49 d
C_i (Pa)	28.9 ± 0.8 a	27.9 ± 0.7 a	47.6 ± 2.5 b	51.1 ± 2 bc	27.9 ± 0.9 a	31.9 ± 0.4 a	52.9 ± 1.2 c	53.4 ± 1.7 c	29.6 ± 0.2 a	32 ± 0.6 a	49.2 ± 1.7 bc	51.6 ± 2.9 bc
Leaf Temperature (°C)	28.8 ± 0.2 a	29.5 ± 0.2 a	29.6 ± 0.2 a	29.6 ± 0.2 a	41.5 ± 0.4 d	38.9 ± 1.1 bc	41.1 ± 0.1 cd	41.6 ± 0.5 d	40.2 ± 0.4 bc	39.4 ± 0.9 b	41.7 ± 0.5 d	42.2 ± 0.7 d
g_s (mol H_2O m^{-2} s^{-1})	0.64 ± 0.08 abc	0.49 ± 0.1 abc	0.34 ± 0.04 a	0.4 ± 0.05 ab	0.67 ± 0.08 bc	0.77 ± 0.08 cd	0.6 ± 0.003 abc	0.53 ± 0.11 abc	0.84 ± 0.15 d	0.79 ± 0.09 cd	0.45 ± 0.08 ab	0.39 ± 0.11 ab
n	5	6	5	4	6	5	6	6	6	5	5	4

The shock group was made up of 30 °C plants exposed to 45 °C for 2 days. All values are means (± SE), except for A_n/V_{cmax} (%). Means that do not share a letter are significantly different (as determined by Duncan's Multiple Range Test)

A_n CO_2 assimilation rate measured on LI-6800, V_{Cmax} maximum velocity of Rubisco, J_{max} maximum velocity of electron transport, C_C pressure of CO_2 at the chloroplast, g_m mesophyll conductance, C_i pressure of intercellular CO_2, g_s stomatal conductance to water, *n* number of replicate RACiR curves per group

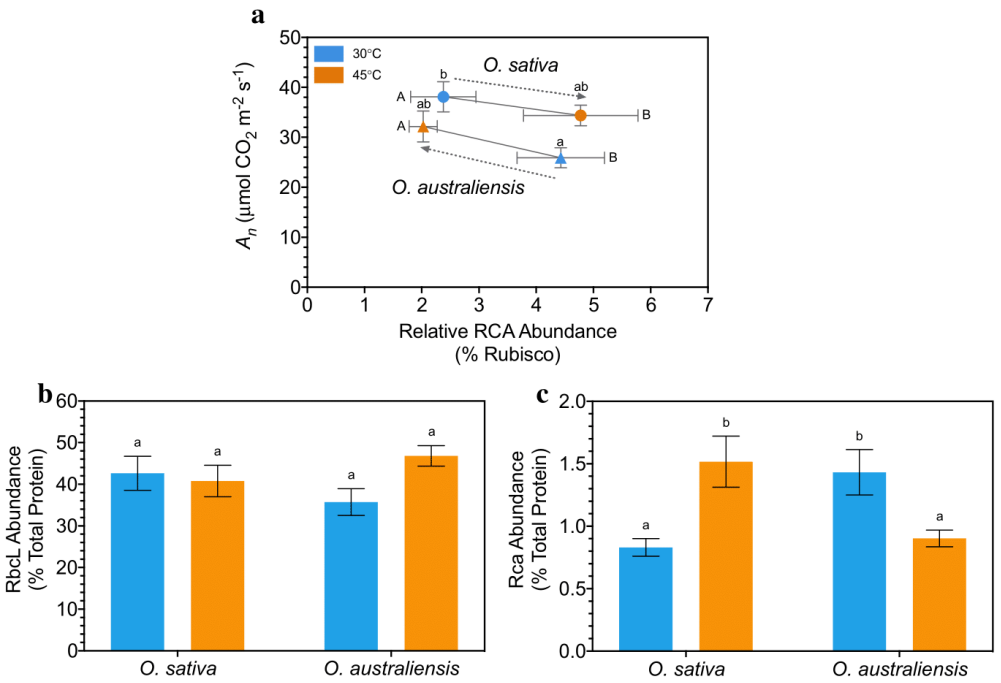


Fig. 7 Effect of heat on CO_2 assimilation and the abundance of RbcL and Rca in *O. sativa* and *O. australiensis*. **A** Relationship between the abundance of Rca relative to Rubisco (Rca emPAI/RbcL emPAI) and net CO_2 assimilation (A_n) in *O. sativa* and *O. australiensis* leaves exposed to 30 and 45 °C (shock). Arrowheads emphasise the relationship between relative Rca abundance and photosynthetic rates for each species. A_n was estimated using a LI-6400. Capital letters show differences in Rca relative abundance; lowercase letters show differences in assimilation rates. **B** Abundance of RbcL relative to total protein emPAI (RbcL emPAI/total emPAI). **C** Abundance of Rca relative to total protein emPAI (Rca emPAI/total emPAI). Values are means ± SE of nine pot replicates (time of day replicates were pooled). Means that do not share a letter are significantly different (Duncan's Multiple Range test). See Fig. 2 for sampling information

times greater than it was at 45 °C. Conversely, in *O. sativa* leaves the concentration of Rca relative to Rubisco at 30 °C was half that at 45 °C (Fig. 7a). Differences in Rca abundance relative to RbcL were not strictly due to differences in the abundance of RbcL but were driven more by changes in the abundance of Rca (Fig. 7b, c). Despite the relative abundance of Rca decreasing at 45 °C in *O. australiensis*, steady-state A_n appeared to increase (measured on the same leaves as those sampled for protein; Fig. 7a). By contrast, the increase in relative Rca abundance in *O. sativa* was associated with decreased A_n with heat.

A two-way interaction between temperature and time of day ($p < 0.01$) shows that the abundance of Rca relative to RbcL increased following exposure to 45 °C for 5 h, though no other significant differences were detected (Fig. S2; Tukey HSD). Similarly, a two-way interaction between species and time of day ($p < 0.001$) shows that Rca:RbcL increased 5 h into the photoperiod in *O. sativa*, and this ratio was greater than that seen in *O. australiensis* under the same conditions (Tukey HSD).

Discussion

This study explores two closely related grasses in which we have previously established starkly contrasting degrees of thermotolerance (Scafaro et al. 2011, 2016). The tolerant species, *Oryza australiensis*, is an extremophile distributed across the hot northern savannah of Australia, while *O. sativa* ssp. *japonica* is largely cultivated in the humid subtropics where daily maxima rarely exceed 35 °C. However, the mechanisms that enable wild rice to continue growing, even during sustained periods of high temperature, are still not known. This paper links photosynthetic processes in *O.*

australiensis with its sustained thermotolerance during early vegetative growth.

When grown at high temperature (45 °C) and ambient CO_2 (400 ppm), *O. australiensis* accumulated more biomass than *O. sativa*. Furthermore, *O. australiensis* was susceptible to low temperature (25 °C) while *O. sativa* suffered severe growth penalties when grown at high temperature. Ali et al. (2019) showed reductions in growth of the same *O. sativa* cultivar (Amaroo) when grown at high temperature for a short time during grain filling. Furthermore, Perdomo et al. (2015) show reduced biomass accumulation in *O. sativa* (cv. Bomba) when grown at high temperature. Taken together these results underline the superior thermotolerance of *O. australiensis* to sustained heat, relative to *O. sativa*. Interestingly, 45 °C days only slowed biomass gain in *O. sativa* after 3 weeks, even though leaves elongated more slowly within a few hours of 45 °C being imposed. Previous studies have highlighted the impact of heat stress on the photosynthetic and respiratory rates and subsequent biomass accumulation of *O. sativa* (Jagadish et al. 2015; Shi et al. 2016; Zhang et al. 2018; Karwa et al. 2020). While these studies only consider the impact of heat over a relatively short time (max. 20 days), and usually during or after anthesis with a focus on grain yield, they reveal alterations to rice carbon metabolism initiated by high temperatures. In this context, the results of the present study indicate that vegetative growth in *O. sativa* is impeded at sustained high temperature by its cumulative impact on the whole-plant carbon budget. We ascribe this to reduced carbohydrate availability and mobilisation from reserves rather than a shift in partitioning of biomass between roots and shoots because root-to-shoot ratios were relatively stable over time (Fig. S3).

LER data showed that leaves of both species elongated rapidly 1–2 h after transfer to 45 °C, which suggests that heat did not directly impair the mechanics of cell expansion, for example by disrupting the cytoskeleton (Smertenko et al. 1997) or membrane integrity (Horváth et al. 2012) but rather it acted through secondary impacts of heat on metabolism. The species contrast was first seen 4 h after plants were exposed to 45 °C, consistent with a growing deficit of carbohydrates in *O. sativa* and subsequent slowing of leaf growth over the diurnal cycle of heat. This is in accord with the finding that the impact of heat on grain filling in *O. sativa* becomes more severe with prolonged exposure to heat (Shi et al. 2016). On the other hand, LER in *O. australiensis* was essentially maintained above 3 mm h^{-1} for 10 h at 45 °C, while rates returned to 1—1.5 mm h^{-1} in both species at night. These findings corroborate those of Scafaro et al. (2010) and Scafaro et al. (2016), who showed reductions in LER of *O. sativa* following 4 h of exposure to 45 °C of up to 50%, while Scafaro et al. (2016) report no significant reduction in the LER of *O. australiensis* after 4 h at 45 °C, which was linked to its photosynthetic thermotolerance. That *O. australiensis* has consistently demonstrated superior growth to *O. sativa* in the face of high temperature suggests that it has access to the resources (namely carbon resources either generated via photosynthesis or mobilised from stored carbohydrates) that are required to support such growth.

We found that after 10 h at 45 °C, sugar concentrations accumulated threefold in *O. australiensis* and doubled in *O. sativa* (Fig. 3). These data contrast with the steady sugar levels recorded in rice leaves exposed to longer periods at less extreme temperatures up to 40 °C (Rashid et al. 2020). Soluble sugars are thought to play a role in various abiotic stress tolerances by acting as signalling molecules, by conferring antioxidant properties (Harsh et al. 2016), by acting as compatible solutes (Kaplan and Guy 2004) and stabilising protein structure (Lee and Timasheff 1981). Given the role that soluble sugars play in protecting plants against thermally induced damage, we speculate that high soluble sugar levels in both rice species contributed to thermotolerance at the molecular level. In addition, the species-specific accumulation of sugars in response to heat is interesting as a resource for growth and photosynthetic rates. Steady-state soluble sugar concentrations in leaves are the consequence of assimilation rates, starch degradation, and sugar utilisation (Stitt and Zeeman 2012). Therefore, trebling of leaf soluble sugars in *O. australiensis* should be seen in the context of the faster growth at 45 °C compared with *O. sativa*; leaf growth in *O. australiensis* continued unabated at 45 °C while *O. sativa* leaves elongated significantly slower (Fig. 2A, B). This makes the marked accumulation of soluble sugars in leaves of *O. australiensis* remarkable because it occurred despite a greater demand for carbohydrates that would be required to sustain steady growth. The maintenance of *O. sativa* leaf sugar levels at 45 °C may have come at a cost to starch reserves, considering photosynthesis was impaired under these hotter growth conditions in the domesticated rice (Figs. 6, 7). Increasing soluble sugar content in plant tissues is often reported alongside reductions in starch reserves as stored carbohydrates are mobilised, requiring the expression of a suite of genes involved in starch degradation and monosaccharide transport (Kaplan et al. 2006). For example, Yichie et al. (2019) reported increases in expression of a starch synthase in a salt-susceptible population of wild rice, while a tolerant population of the same species had increased expression of monosaccharide transporters. The rapid slowing of biomass accumulation in *O. sativa* after 28 days at 45 °C (Fig. 1) may in fact be due to starch reserves reaching a critically low level after sustained depletion.

Contrasting sugar levels in the two species and previous observations that photosynthesis is vulnerable to heat-induced damage at 45 °C (Salvucci et al. 2001; Sage et al. 2008; Scafaro et al. 2010, 2016; Yamori et al. 2014; Busch and Sage 2017) led to an inquiry into photosynthetic metabolism in the two species under heat shock. Unsurprisingly,

peak A_n recorded (as instantaneous measurements) at 30 °C were 50% higher in the domesticated cultivar of *O. sativa*, while at 45 °C rates in *O. sativa* and *O. australiensis* were approximately equal (Fig. 7). This finding emphasises the highly significant opposing effects of heat on assimilation in the two species (i.e. a 30% increase in *O. australiensis* and a 20% decrease in *O. sativa* when exposed to 45 °C). Similarly, IR64 (an *O. sativa* cultivar) had reduced photosynthesis when exposed to a high temperature shock (Rashid et al. 2020). Furthermore, *O. sativa* (cv. Bomba) had reduced daily average assimilation rates when grown at high temperature (Perdomo et al. 2015). Importantly, the changes in instantaneous measures of A_n during heat shock could not be explained by differences in stomatal conductance (g_s), intracellular CO_2 (C_i), leaf temperature nor vapour pressure deficit (Fig. S4). This finding is supported by a preliminary study we performed, where significant changes in VPD could not account for differences in assimilation rates because, despite the response of g_s to high temperature, intercellular CO_2 availability was the same for both species (Fig. S5). Differences in leaf anatomical properties between rice and its wild relatives do exist, and these differences seem to influence CO_2 conductance properties between the air and chloroplasts (Scafaro et al. 2011; Giuliani et al. 2013). However, the results presented here suggest that temperature-dependent changes in A_n are dictated by changes in the efficiency of the underlying photosynthetic biochemistry.

RACiR curves are a novel technique that have been used successfully to rapidly acquire gas exchange data for the estimation of V_{cmax} and J_{max} while simultaneously overcoming the issues of stomatal attenuation and changes in enzyme kinetics that are characteristic of the traditional A:C_i curve (Stinziano et al. 2017, 2019; Coursolle et al. 2019; Vincent et al. 2020). RACiR curves, Laisk curves, and point measures of chlorophyll fluorescence showed that differences in assimilation between wild and domestic rice during long- or short-term heat exposure could not be explained by differences in photorespiration, respiration or *ETR* (Fig. 5). That is, the thermotolerance of assimilation in *O. australiensis* cannot be due to a species-specific response of photorespiration, respiration, or *ETR* to heat.

Rates of photorespiration increase with increasing temperature because the specificity of Rubisco to CO_2 declines relative to O_2 under heat, and the abundance of CO_2 relative to O_2 is reduced due to differences in the solubility of these two gases at high temperature (Walker et al. 2016; Dusenge et al. 2019). Given that rates of photorespiration are driven in part by the affinity of Rubisco for CO_2, we conclude that the lack of species differences observed here indicate that the wild and domestic rice Rubisco isoforms must have similar affinities for their substrates. This is in accordance with recent work on *Arabidopsis* that showed no genetic basis for variation in photorespiration between ecotypes (Tomeo and Rosenthal 2018).

Similarly, respiration increases with increasing temperature (Dusenge et al. 2019). Here, we showed similar increases in respiration at high temperature in both species. Rashid et al. (2020) showed that rice (IR64) had limited respiratory acclimation over seven days of exposure to 40 °C, mirroring the results obtained here. Interestingly, *O. australiensis* respiration showed greater potential to acclimate to sustained high temperature at 400 ppm CO_2 than *O. sativa*. Thus, while we cannot attribute differences in assimilation to the observed differences in respiration between the species at high temperature, *O. australiensis* may provide useful germplasm to explore variation in respiratory response under heat. Finally, reduced A_n in *O. sativa* at 45 °C (both long- and short-term) comes about even though concentrations of CO_2 at the chloroplast remained stable or increase. Taken together, the results presented thus far provide evidence for the notion of biochemical limitations at the chloroplast during episodes of high temperature in *O. sativa* but not *O. australiensis*. Interestingly, the RACiR curves showed some degree of acclimation of assimilation to both long-term heat (*cf.* short-term) and elevated ambient CO_2 in *O. australiensis* though not in *O. sativa* (Fig. 6). The downregulation of photosynthetic capacity in *O. australiensis* when grown at 700 ppm CO_2 suggests that the wild species is capable of meeting sink demands even when reducing its investment in photosynthetic machinery. This is seen in its most extreme form in *O. australiensis* plants grown at 700 ppm CO_2 *and* sustained high temperature, which may indicate a metabolism that is adapted to operate most efficiently in hot environments. On the other hand, *O. sativa* appears to maintain investment in photosynthetic machinery at high CO_2, resulting in much faster A_n with CO_2 enrichment. Under a similar 700 ppm CO_2 enrichment study, *O. sativa* significantly increased A_n while the wild relative *Oryza meridionalis* had a limited response (Rahman et al. 2018). The A_n of wild relatives thus seems to be less responsive to CO_2 fertilisation relative to domestic rice. We speculate that *O. australiensis* reduced photosynthetic capacity with CO_2 enrichment because of limited sink demand, whereas *O. sativa* maintained photosynthetic capacity due to artificial selection increasing demand for assimilates needed for grain filling. Supporting this postulate, enriched CO_2 has consistently been shown to increase grain yield in commercial rice (Madan et al. 2012), demonstrating that improved crop performance by engineering faster assimilation will likely be effective only in cases where demand for assimilates (sink strength) has not already reached saturation.

The CO_2 response curves suggested Rubisco carboxylation limitations existed in *both species* at 45 °C and 400 ppm CO_2, and thus a direct effect of heat on CO_2 fixation by Rubisco is strongly implicated as the reason for differences

in photosynthetic thermotolerance between the species. Prior evidence (Scafaro et al. 2016) suggests that the properties of Rca form part of the explanation for heat tolerance in *O. australiensis*. Indeed, the photosynthetic heat tolerance of *O. sativa* can be improved with overexpression of *O. australiensis* or maize Rca (Yamori et al. 2012; Scafaro et al. 2018). The activation of Rubisco by Rca depends on the ratio of the two enzymes and the susceptibility of Rca to heat (Perdomo et al. 2017). Rca abundance can respond dynamically to heat stress in a timeframe of hours to days through transcriptional and post-transcriptional regulation (Degen et al. 2021). In relation to the ratio of Rubisco to Rca, Fukayama et al. (2012) and Fukayama et al. (2018) showed that rice overexpressing functional Rca had reduced abundance of Rubisco, while an Rca knock-out line had increased Rubisco abundance. Suganami et al. (2018) showed that as the abundance of Rubisco increases, cells disinvest in Rca, and vice versa, indicating that this feedback operates in both directions. Further, a recent report in rice demonstrated that overexpression of both Rca and Rubisco leads to improvement in assimilation, particularly under heat stress (Qu et al. 2021). Given that Rca functions in the activation of Rubisco (Portis 2003), these previous results imply that photosynthetic capacity is tightly regulated and set by the number of active Rubisco sites. As we observed changes in Rca abundance, but no significant changes in Rubisco abundance with heat, Rca was regulating photosynthetic capacity. Contrary changes in total Rca abundance between the species when grown at 45 °C complement the inherent thermostability of the Rca isoforms in each species. Specifically, the dominant Rca β isoform found in *O. australiensis* is known to have greater thermostability than its *O. sativa* ortholog (Scafaro et al. 2016). Thus, it may be the case here that at 30 °C *O. australiensis* Rca was operating below its thermal optimum. An increase in abundance of Rca presumably compensated for the loss in efficiency at this suboptimal temperature (Fig. 7). This limitation was presumably lifted at higher temperatures, allowing for increased enzyme function and a decrease in Rca abundance. Interestingly, reduced investment in Rca because of increased Rca activity may enhance plant growth over long periods of heat because after Rubisco, Rca is the second highest consumer of ATP in leaves (Li et al. 2017). Furthermore, the savings in ATP costs associated with reduced Rca may contribute to the greater respiratory acclimation potential of hot grown *O. australiensis,* as we explore above. We suggest that the opposite applies in *O. sativa*, whose Rca may have been operating outside its thermal optimum at 45 °C, likewise explaining the increased accumulation of the enzyme during exposure to heat and reduced accumulation under optimal conditions (30 °C). When grown at high temperatures, the stability of Rca may therefore explain the reduced Rca abundance, greater photosynthetic performance, and ultimately faster growth in the heat-adapted *O. australiensis*. This is supported by our earlier evidence that transgenic rice expressing the thermotolerant ortholog of Rca from *O. australiensis* had increased growth rates and seed yield at 45 °C (Scafaro et al. 2018). We further demonstrate in this study that neither photosynthetic electron transport, respiratory capacity nor photorespiratory performance could explain the susceptibility of *O. sativa* to extreme heat. Furthermore, we found no evidence for differences in Rubisco's affinity for CO_2 in the wild rice, which would have provided partial explanation for improved assimilation at high temperature. Rca as a driver of heat tolerance in wild rice therefore seems highly likely (Perdomo et al. 2017). Whether this mechanistic basis of heat tolerance applies to other important crops should be a priority of future research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11103-021-01210-3.

Acknowledgements Aspects of this research have been facilitated by access to APAF supported under the Australian Government's NCRIS. We thank Dr Mohammad Masood for maintaining the growth facilities. We extend the sincerest of thanks to Matthew Alfonzetti for helping to collect the *A:Ci*, Laisk and chlorophyll fluorescence data. We especially appreciate the thoughtful reviews of this paper that led to a greatly improved manuscript.

Author contributions ALP and BJA contributed to the study conception and design. Material preparation, data collection and analysis were performed by ALP. Additional data analyses were suggested and performed by APS. The first draft of the manuscript was written by ALP and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Not Applicable.

Data availability All data available upon request to the corresponding author.

Code availability All code and data available at: https://github.com/aaronphillips7493/Oa_heat_paper_2021.

Declarations

Conflict of interest Not Applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Ali F, Waters DLE, Ovenden B, Bundock P, Raymond CA, Rose TJ (2019) Australian rice varieties vary in grain yield response to heat stress during reproductive and grain filling stages. J Agron Crop Sci 205:179–187. https://doi.org/10.1111/jac.12312

Asseng S, Ewert F, Martre P, Rötter RP, Lobell DB, Cammarano D, Kimball BA, Ottman MJ, Wall GW, White JW, Reynolds MP, Alderman PD, Prasad PVV, Aggarwal PK, Anothai J, Basso B, Biernath C, Challinor AJ, De Sanctis G, Doltra J, Fereres E, Garcia-Vila M, Gayler S, Hoogenboom G, Hunt LA, Izaurralde RC, Jabloun M, Jones CD, Kersebaum KC, Koehler A-K, Müller C, Naresh Kumar S, Nendel C, O'Leary G, Olesen JE, Palosuo T, Priesack E, Eyshi Rezaei E, Ruane AC, Semenov MA, Shcherbak I, Stöckle C, Stratonovitch P, Streck T, Supit I, Tao F, Thorburn PJ, Waha K, Wang E, Wallach D, Wolf J, Zhao Z, Zhu Y (2015) Rising temperatures reduce global wheat production. Nat Clim Chang 5:143–147. https://doi.org/10.1038/nclimate2470

Atkin OK, Tjoelker MG (2003) Thermal acclimation and the dynamic response of plant respiration to temperature. Trends Plant Sci 8:343–351. https://doi.org/10.1016/S1360-1385(03)00136-5

Atwell BJ, Wang H, Scafaro AP (2014) Could abiotic stress tolerance in wild relatives of rice be used to improve *Oryza sativa*? Plant Sci 215–216:48–58. https://doi.org/10.1016/j.plantsci.2013.10.007

Avenson TJ, Saathoff AJ (2018) Sub-saturating multiphase flash irradiances to estimate maximum fluorescence yield. Methods Mol Biol 1770:105–120. https://doi.org/10.1007/978-1-4939-7786-4_7

Benomar L, Lamhamedi MS, Pepin S, Rainville A, Lambert M-C, Margolis HA, Bousquet J, Beaulieu J (2018) Thermal acclimation of photosynthesis and respiration of southern and northern white spruce seed sources tested along a regional climatic gradient indicates limited potential to cope with temperature warming. Ann Bot 121:443–457. https://doi.org/10.1093/aob/mcx174

Busch FA, Sage RF (2017) The sensitivity of photosynthesis to O_2 and CO_2 concentration identifies strong Rubisco control above the thermal optimum. New Phytol 213:1036–1051. https://doi.org/10.1111/nph.14258

Campbell C, Atkinson L, Zaragoza-Castells J, Lundmark M, Atkin O, Hurry V (2007) Acclimation of photosynthesis and respiration is asynchronous in response to changes in temperature regardless of plant functional group. New Phytol 176:375–389. https://doi.org/10.1111/j.1469-8137.2007.02183.x

Chi Y, Xu M, Shen R, Yang Q, Huang B, Wan S (2013) Acclimation of foliar respiration and photosynthesis in response to experimental warming in a temperate steppe in Northern China. PLoS ONE 8:e56482. https://doi.org/10.1371/journal.pone.0056482

Coursolle C, Otis Prud'homme G, Lamothe M, Isabel N (2019) Measuring rapid A-Ci curves in boreal conifers: black spruce and balsam fir. Front Plant Sci 10:1276. https://doi.org/10.3389/fpls.2019.01276

CSIRO (2018) State of the Climate. https://www.csiro.au/en/Showcase/state-of-the-climate. Accessed 3 Sept 2020

Degen GE, Orr DJ, Carmo-Silva E (2021) Heat-induced changes in the abundance of wheat Rubisco activase isoforms. New Phytol 229:1298–1311. https://doi.org/10.1111/nph.16937

Dhir B (2018) Crop productivity in changing climate. In: Lichtfouse E (ed) Sustainable agriculture reviews 27. Springer, Cham, pp 213–241

Dusenge ME, Duarte AG, Way DA (2019) Plant carbon metabolism and climate change: elevated CO_2 and temperature impacts on photosynthesis, photorespiration and respiration. New Phytol 221:32–49. https://doi.org/10.1111/nph.15283

Ebrahim MK, Zingsheim O, El-Shourbagy MN, Moore PH, Komor E (1998) Growth and sugar storage in sugarcane grown at temperatures below and above optimum. J Plant Physiol 153:593–602. https://doi.org/10.1016/S0176-1617(98)80209-5

Farquhar GD, von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. Planta 149:78–90. https://doi.org/10.1007/BF00386231

Feller U, Crafts-Brandner SJ, Salvucci ME (1998) Moderately high temperatures inhibit ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) activase-mediated activation of Rubisco. Plant Physiol 116:539–546. https://doi.org/10.1104/pp.116.2.539

Fischer T, Byerlee D, Edmeades G (2014) Crop yields and global food security: will yield increase continue to feed the world? Agricultural Research, Canberra, p 634

Fukayama H, Ueguchi C, Nishikawa K, Katoh N, Ishikawa C, Masumoto C, Hatanaka T, Misoo S (2012) Overexpression of Rubisco activase decreases the photosynthetic CO2 assimilation rate by reducing Rubisco content in rice leaves. Plant Cell Physiol 53:976–986. https://doi.org/10.1093/pcp/pcs042

Fukayama H, Mizumoto A, Ueguchi C, Katsunuma J, Morita R, Sasayama D, Hatanaka T, Azuma T (2018) Expression level of Rubisco activase negatively correlates with Rubisco content in transgenic rice. Photosynth Res 137:465–474. https://doi.org/10.1007/s11120-018-0525-9

Giuliani R, Koteyeva N, Voznesenskaya E, Evans MA, Cousins AB, Edwards GE (2013) Coordination of leaf photosynthesis, transpiration, and structural traits in rice and wild relatives (Genus *Oryza*). Plant Physiol 162:1632–1651. https://doi.org/10.1104/pp.113.217497

Gomathi R, Shiyamala S, Vasantha S, Johnson DE, Janani PK (2013) Kinetics of metabolism in sugarcane (*Saccharum officinarum* L.) under heat stress. Indian J Plant Physiol 18:41–47. https://doi.org/10.1007/s40502-013-0011-5

Harsh A, Sharma YK, Joshi U, Rampuria S, Singh G, Kumar S, Sharma R (2016) Effect of short-term heat stress on total sugars, proline and some antioxidant enzymes in moth bean (*Vigna aconitifolia*). Ann Agric Sci 61:57–64. https://doi.org/10.1016/j.aoas.2016.02.001

Hermida-Carrera C, Kapralov MV, Galmés J (2016) Rubisco catalytic properties and temperature response in crops. Plant Physiol 171:2549. https://doi.org/10.1104/pp.16.01846

Horváth I, Glatz A, Nakamoto H, Mishkind ML, Munnik T, Saidi Y, Goloubinoff P, Harwood JL, Vigh L (2012) Heat shock response in photosynthetic organisms: membrane and lipid connections. Prog Lipid Res 51:208–220. https://doi.org/10.1016/j.plipres.2012.02.002

IPCC (2019) Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. IPCC, Geneva

Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 4:1265–1272. https://doi.org/10.1074/mcp.M500061-MCP200

Jagadish SVK, Septiningsih EM, Kohli A, Thomson MJ, Ye C, Redoña E, Kumar A, Gregorio GB, Wassmann R, Ismail AM, Singh RK (2012) Genetic advances in adapting rice to a rapidly changing climate. J Agron Crop Sci 198:360–373. https://doi.org/10.1111/j.1439-037X.2012.00525.x

Jagadish SVK, Murty MVR, Quick WP (2015) Rice responses to rising temperatures—challenges, perspectives and future directions. Plant Cell Environ 38:1686–1698. https://doi.org/10.1111/pce.12430

Kaplan F, Guy CL (2004) β-amylase induction and the protective role of maltose during temperature shock. Plant Physiol 135:1674–1684. https://doi.org/10.1104/pp.104.040808

Kaplan F, Sung DY, Guy CL (2006) Roles of β-amylase and starch breakdown during temperatures stress. Physiol Plant 126:120–128. https://doi.org/10.1111/j.1399-3054.2006.00604.x

Karwa S, Bahuguna RN, Chaturvedi AK, Maurya S, Arya SS, Chinnusamy V, Pal M (2020) Phenotyping and characterization of heat stress tolerance at reproductive stage in rice (*Oryza sativa* L.). Acta Physiol Plant 42:29. https://doi.org/10.1007/s11738-020-3016-5

Laisk AK (1977) Kinetics of photosynthesis and photorespiration of C3 in plants. Nauka, Moscow

Lee JC, Timasheff SN (1981) The stabilization of proteins by sucrose. J Biol Chem 256:7193–7201

Li L, Nelson CJ, Trösch J, Castleden I, Huang S, Millar AH (2017) Protein degradation rate in *Arabidopsis thaliana* leaf growth and development. Plant Cell 29:207–228. https://doi.org/10.1105/tpc.16.00768

Loriaux SD, Avenson TJ, Welles JM, McDermitt DK, Eckles RD, Rienshe B, Genty B (2013) Closing in on maximum yield of chlorophyll fluorescence using a single multiphase flash of subsaturating intensity. Plant Cell Environ 36:1755–1770. https://doi.org/10.1111/pce.12115

Madan P, Jagadish SVK, Craufurd PQ, Fitzgerald M, Lafarge T, Wheeler TR (2012) Effect of elevated CO_2 and high temperature on seed-set and grain quality of rice. J Exp Bot 63:3843–3852. https://doi.org/10.1093/jxb/ers077

Mahlstein I, Knutti R, Solomon S, Portmann RW (2011) Early onset of significant local warming in low latitude countries. Environ Res Lett 6:034009. https://doi.org/10.1088/1748-9326/6/3/034009

Peng S, Huang J, Sheehy JE, Laza RC, Visperas RM, Zhong X, Centeno GS, Khush GS, Cassman KG (2004) Rice yields decline with higher night temperature from global warming. Proc Natl Acad Sci USA 101:9971–9975. https://doi.org/10.1073/pnas.0403720101

Perdomo JA, Conesa M, Medrano H, Ribas-Carbó M, Galmés J (2015) Effects of long-term individual and combined water and temperature stress on the growth of rice, wheat and maize: relationship with morphological and physiological acclimation. Physiol Plant 155:149–165. https://doi.org/10.1111/ppl.12303

Perdomo JA, Capó-Bauçà S, Carmo-Silva E, Galmés J (2017) Rubisco and Rubisco activase play an important role in the biochemical limitations of photosynthesis in rice, wheat, and maize under high temperature and water deficit. Front Plant Sci 8:490. https://doi.org/10.3389/fpls.2017.00490

Portis AR (2003) Rubisco activase—Rubisco's catalytic chaperone. Photosynth Res 75:11–27. https://doi.org/10.1023/A:1022458108678

Qu Y, Sakoda K, Fukayama H, Kondo E, Suzuki Y, Makino A, Terashima I, Yamori W (2021) Overexpression of both Rubisco and Rubisco activase rescues rice photosynthesis and biomass under heat stress. Plant Cell Environ 44:2308–2320. https://doi.org/10.1111/pce.14051

R Core Team (2017) R: a language and environment for statistical computing. R Foundation forStatistical Computing, Vienna

Rahman S, Duursma RA, Muktadir MA, Roberts TH, Atwell BJ (2018) Leaf canopy architecture determines light interception and carbon gain in wild and domesticated *Oryza* species. Environ Exp Bot 155:672–680. https://doi.org/10.1016/j.envexpbot.2018.08.008

Rashid FAA, Crisp PA, Zhang Y, Berkowitz O, Pogson BJ, Day DA, Masle J, Dewar RC, Whelan J, Atkin OK, Scafaro AP (2020) Molecular and physiological responses during thermal acclimation of leaf photosynthesis and respiration in rice. Plant Cell Environ 43:594–610. https://doi.org/10.1111/pce.13706

Reddy KR, Reddy VR, Hodges HF (1992) Temperature effects on early season cotton growth and development. Agron J 84:229. https://doi.org/10.2134/agronj1992.00021962008400020021x

Sage RF, Way DA, Kubien DS (2008) Rubisco, Rubisco activase, and global climate change. J Exp Bot 59:1581–1595. https://doi.org/10.1093/jxb/ern053

Salvucci ME, Crafts-Brandner SJ (2004) Relationship between the heat tolerance of photosynthesis and the thermal stability of rubisco activase in plants from contrasting thermal environments. Plant Physiol 134:1460–1470. https://doi.org/10.1104/pp.103.038323

Salvucci ME, Osteryoung KW, Crafts-Brandner SJ, Vierling E (2001) Exceptional sensitivity of Rubisco activase to thermal denaturation in vitro and in vivo. Plant Physiol 127:1053–1064. https://doi.org/10.1104/pp.010357

Scafaro AP, Haynes P, Atwell B (2010) Physiological and molecular changes in *Oryza meridionalis* Ng., a heat-tolerant species of wild rice. J Exp Bot 61:191–202. https://doi.org/10.1093/jxb/erp294

Scafaro AP, Von Caemmerer S, Evans JR, Atwell BJ (2011) Temperature response of mesophyll conductance in cultivated and wild *Oryza* species with contrasting mesophyll cell wall thickness. Plant Cell Environ 34:1999–2008. https://doi.org/10.1111/j.1365-3040.2011.02398.x

Scafaro AP, Gallé A, Van Rie J, Carmo-Silva E, Salvucci ME, Atwell BJ (2016) Heat tolerance in a wild *Oryza* species is attributed to maintenance of Rubisco activation by a thermally stable Rubisco activase ortholog. New Phytol 211:899–911. https://doi.org/10.1111/nph.13963

Scafaro AP, Atwell BJ, Muylaert S, Van RB, Ruiz GA, Van RJ, Gallé A (2018) A thermotolerant variant of Rubisco activase from a wild relative improves growth and seed yield in rice under heat stress. Front Plant Sci 9:1–11. https://doi.org/10.3389/fpls.2018.01663

Scafaro AP, Bautsoens N, den Boer B, Van Rie J, Galle A (2019) A conserved sequence from heat-adapted species improves Rubisco activase thermostability in wheat. Plant Physiol 181:43–54. https://doi.org/10.1104/pp.19.00425

Scharf KD, Berberich T, Ebersberger I, Nover L (2012) The plant heat stress transcription factor (Hsf) family: structure, function and evolution. Biochim Biophys Acta Gene Regul Mech 1819:104–119. https://doi.org/10.1016/j.bbagrm.2011.10.002

Shi P, Zhu Y, Tang L, Chen J, Sun T, Cao W, Tian Y (2016) Differential effects of temperature and duration of heat stress during anthesis and grain filling stages in rice. Environ Exp Bot 132:28–41. https://doi.org/10.1016/j.envexpbot.2016.08.006

Smertenko A, Dráber P, Viklický V, Opatrný Z (1997) Heat stress affects the organization of microtubules and cell division in *Nicotiana tabacum* cells. Plant Cell Environ 20:1534–1542. https://doi.org/10.1046/j.1365-3040.1997.d01-44.x

Stinziano JR, Morgan PB, Lynch DJ, Saathoff AJ, McDermitt DK, Hanson DT (2017) The rapid *A-Ci* response: photosynthesis in the phenomic era. Plant Cell Environ 40:1256–1262. https://doi.org/10.1111/pce.12911

Stinziano JR, McDermitt DK, Lynch DJ, Saathoff AJ, Morgan PB, Hanson DT (2019) The rapid *A/Ci* response: a guide to best practices. New Phytol 221:625–627. https://doi.org/10.1111/nph.15383

Stitt M, Zeeman SC (2012) Starch turnover: pathways, regulation and role in growth. Curr Opin Plant Biol 15:282–292. https://doi.org/10.1016/j.pbi.2012.03.016

Suganami M, Suzuki Y, Sato T, Makino A (2018) Relationship between Rubisco activase and Rubisco contents in transgenic rice plants with overproduced or decreased Rubisco content. Soil Sci Plant Nutr 64:352–359. https://doi.org/10.1080/00380768.2018.1433455

Tilman D, Balzer C, Hill J, Befort BL (2011) Global food demand and the sustainable intensification of agriculture. Proc Natl Acad Sci USA 108:20260–20264. https://doi.org/10.1073/pnas.1116437108

Tjoelker MG (2018) The role of thermal acclimation of plant respiration under climate warming: putting the brakes on a runaway

train? Plant Cell Environ 41:501–503. https://doi.org/10.1111/pce.13126

Tomeo NJ, Rosenthal DM (2018) Photorespiration differs among *Arabidopsis thaliana* ecotypes and is correlated with photosynthesis. J Exp Bot 69:5191–5204. https://doi.org/10.1093/jxb/ery274

Vincent C, Pierre MO, Stinziano JR (2020) Racing against stomatal attenuation: rapid CO_2 response curves more reliably estimate photosynthetic capacity than steady state curves in a low conductance species. bioRxiv. https://doi.org/10.1101/2020.08.28.270785

Wahid A, Gelani S, Ashraf M, Foolad MR (2007) Heat tolerance in plants: an overview. Environ Exp Bot 61:199–223. https://doi.org/10.1016/j.envexpbot.2007.05.011

Walker BJ, VanLoocke A, Bernacchi CJ, Ort DR (2016) The costs of photorespiration to food production now and in the future. Annu Rev Plant Biol 67:107–129. https://doi.org/10.1146/annurev-arplant-043015-111709

Wright IJ, Dong N, Maire V, Prentice IC, Westoby M, Díaz S, Gallagher RV, Jacobs BF, Kooyman R, Law EA, Leishman MR, Niinemets Ü, Reich PB, Sack L, Villar R, Wang H, Wilf P (2017) Global climatic drivers of leaf size. Science 357:917–921. https://doi.org/10.1126/science.aal4760

Yamori W, Masumoto C, Fukayama H, Makino A (2012) Rubisco activase is a key regulator of non-steady-state photosynthesis at any leaf temperature and to a lesser extent, of steady-state photosynthesis at high temperature. Plant J 71:871–880. https://doi.org/10.1111/j.1365-313x.2012.05041.x

Yamori W, Hikosaka K, Way DA (2014) Temperature response of photosynthesis in C3, C4, and CAM plants: temperature acclimation and temperature adaptation. Photosynth Res 119:101–117. https://doi.org/10.1007/s11120-013-9874-6

Yemm EW, Willis AJ (1954) The estimation of carbohydrates in plant extracts by anthrone. Biochem J 57:508–514. https://doi.org/10.1042/bj0570508

Yichie Y, Hasan MT, Tobias PA, Pascovici D, Goold HD, Van Sluyter SC, Roberts TH, Atwell BJ (2019) Salt-treated roots of *Oryza australiensis* seedlings are enriched with proteins involved in energetics and transport. Proteomics 19:1900175. https://doi.org/10.1002/pmic.201900175

Zhang CX, Feng BH, Chen TT, Fu WM, Li HB, Li GY, Jin QY, Tao LX, Fu GF (2018) Heat stress-reduced kernel weight in rice at anthesis is associated with impaired source-sink relationship and sugars allocation. Environ Exp Bot 155:718–733. https://doi.org/10.1016/j.envexpbot.2018.08.021

Zhu W, Smith JW, Huang C-M (2010) Mass spectrometry-based label-free quantitative proteomics. J Biomed Biotechnol 2010:840518. https://doi.org/10.1155/2010/840518

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

INTENTIONALLY LEFT BLANK

INTENTIONALLY LEFT BLANK

Chapter 4

Exploring Responses of *Oryza australiensis* Under Individual and Combined Drought and Nitrogen Deficit

Wombo Prompt:

“Rice seedlings in dry soil”

Wombo Theme

“Comic”

Statement of Authorship

Title of Paper	Exploring Responses of *Oryza australiensis* Under Individual and Combined Drought and Low Nitrogen Stress
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Aaron L. Phillips		
Contribution to the Paper	Conceived the study, performed experiments, analysed and interpreted data, wrote the manuscript.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	2nd September 2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);
ii. permission is granted for the candidate in include the publication in the thesis; and
iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Darren Plett
Contribution to the Paper	Conceived the study and assisted in data analysis

Signature		Date	7th September 2022

Name of Co-Author	Bettina Berger		
Contribution to the Paper	Conceived the study and assisted in data analysis		
Signature		Date	7th September 2022

Name of Co-Author	Brian J. Atwell		
Contribution to the Paper	Conceived the study and edited the manuscript		
Signature		Date	9th September 2022

Name of Co-Author	Rachel A. Burton		
Contribution to the Paper	Conceived the study and edited the manuscript		
Signature		Date	12th October 2022

Target Journal:
Journal of Experimental Botany

Title:
Exploring Responses of *Oryza australiensis* Under Individual and Combined Drought and Low Nitrogen Deficit

Authors:
Aaron L. Phillips[1,4], Darren Plett[1,2], Bettina Berger[1,2], Brian J. Atwell[3], and Rachel A. Burton[1,4]

Affiliations:
[1]University of Adelaide, School of Agriculture, Food, and Wine, Adelaide, South Australia, Australia

[2]Australian Plant Phenomics Facility, The Plant Accelerator, Waite Research Institute, University of Adelaide, Adelaide, Australia

[3]Macquarie University, Department of Biological Sciences, Sydney, New South Wales, Australia

[4]ARC Centre of Excellence in Plant Energy Biology, School of Agriculture, Food, and Wine, The University of Adelaide, Adelaide, South Australia, Australia

Highlight

Seedlings of *Oryza australiensis* maintain rapid growth when grown under low soil N or drought conditions. Changes in root architecture of older plants are consistent with stress avoidance strategies.

Abstract

Rain-fed and/or dry-seeded rice crops are susceptible to drought and low soil nitrogen (N). Both stresses often co-occur in rain-fed systems – a scenario likely to worsen as rainfall becomes more variable – resulting in yield reductions. The interaction between concurrent drought and low soil N has received little attention so here we investigated the above-ground responses of domestic rice (*Oryza sativa*) and a wild Australian rice relative (*O. australiensis*) to both drought and low soil N stresses using state-of-the-art non-destructive phenotyping approaches. Changes in root architecture in these species under these stresses were also characterised for potted plants. We found that *O. australiensis* seedlings (~20-25 DAP) were able to tolerate both low soil N and drought, but *O. sativa* seedlings could not. However, *O. australiensis* could not tolerate these conditions if they persisted past the seedling stage, nor could plants at any age withstand the pressures of both stresses imposed concurrently. *Oryza australiensis* root systems were far more responsive to the imposed stresses than *O. sativa*, and the observed changes are consistent with stress avoidance strategies. The wild rice investigated in this study therefore represents a valuable source of traits for individual low-N and drought tolerance at the seedling stage.

Keywords and Abbreviations

Oryza sativa; *Oryza australiensis*; drought; low soil nitrogen; plant imaging; root architecture; water use efficiency; nitrogen use efficiency

ANOVA = Analysis of Variance

DAP = Days After Planting

ESV = Early Seedling Vigour

GSR = Green Super Rice

HSI = Hue Saturation Intensity

N = Nitrogen

NUE = Nitrogen Use Efficiency

PSA = Projected Shoot Area

RAG = rapid above-ground growth

RGB = Red-Green-Blue

RGR = relative growth rates

SA:V = Surface Area:Volume

SE = Standard Error

TPA = The Plant Accelerator

WUE = Water Use Efficiency

Introduction

Water and nitrogen are two of the most important resources that limit plant growth and crop yield. Rain-fed rice crops face fluctuating water availability, and the soils of these systems often have highly variable nitrogen (N) availability (Homma *et al.*, 2003). As the frequency and duration of drought events increases and rainfall regimes change, rain-fed cropping regions are particularly vulnerable. Changes in farming management practices also mean that dry seeding is becoming a more popular option for establishing a rice crop due to reduced time and labour costs (Farooq *et al.*, 2011) rendering young plants more susceptible to environmental fluctuations (Quinones *et al.*, 2017). Further, as crop yields approach potential yield (yield under totally optimal conditions) more inputs, particularly soil nitrogen, are required to support both vegetative growth and reproductive success. Individually, both drought and N deficit constrain plant growth and have negative impacts on yield. While N application can alleviate the negative impacts of drought on rice yield, N use efficiency (NUE) is lower in rice than the other major cereals (Zhang *et al.*, 2015; Cao *et al.*, 2018). Inadequate N application and poor NUE partly explain why 'superior' rice varieties do not perform well in some farming scenarios (Zhang *et al.*, 2012; Feng *et al.*, 2018; Chen *et al.*, 2020). 'Green Super Rice' (GSR) varieties that can tolerate low-N (i.e., have higher NUE) and drought conditions are under development and will be critical for solving the issues of food security and environmental degradation associated with excessive N usage (Zhang *et al.*, 2015; Feng *et al.*, 2018). Because drought and low-N stress are likely to co-occur in low-input rain-fed cropping scenarios it is important to understand how these factors will interact and what effect this will have on rice physiology.

Drought and low-nitrogen stress act individually in rice to reduce above-ground growth, alter chlorophyll synthesis/turnover, can lead to reduced yield (Lipiec *et al.*, 2013; Xiong *et al.*, 2018; Zhang *et al.*, 2018; Saha *et al.*, 2020; Peng *et al.*, 2021) and

promote the elongation of axial roots in rice (Uga *et al.*, 2013; Xin *et al.*, 2021). While both drought and N deficit responses can manifest in some similar phenotypes in rice, these conditions are fundamentally different and lead to differences in plant morphology. For example, drought tolerant rice accessions (e.g., *aus*) are reported to have thicker roots than drought sensitive accessions (Liao *et al.*, 2022) but root thickness decreases under low-N (Deng *et al.*, 2020). How might such a tension play out when both stresses are imposed in concert? Investigation of how these abiotic stresses might interact have thus far been limited to restoring water adsorption in drought stressed rice roots with ammonium application (Cao et al. 2018a,b,c). The paucity of research into this interaction and the fact that root responses to abiotic conditions, like drought, are highly heterogenous and species-specific (Lozano *et al.*, 2020) limit the understanding of how water and N deficits will affect rice crops when experienced concurrently and therefore necessitate further investigation.

How might one design a rice plant that is tolerant to low soil N and drought and which traits would be necessary? The answers to these questions likely depend on species, soil type, and water regime, and doubtless there are more than one means to this end. As young rice plants are susceptible to environmental variation, particularly dry-seeded rice seedlings in rain-fed systems, we will focus on the seedling stage. Early seedling vigour (ESV) has received much attention as a desirable trait (or complex of traits) for rice crops because it facilitates seedling establishment, and it enables young plants to better tolerate harsh conditions early in their development (e.g., drought, salt, heat and cold; Zhang *et al.*, 2005, 2021; Zhi-Hong *et al.*, 2005; Kumar *et al.*, 2009). Rapid above-ground growth (RAG) and maintenance of green tissue are hallmarks of ESV and therefore plant health and seedling establishment. Abscisic acid produced by roots exposed to drying soils causes stomata to close, which allows seedlings to prevent water loss during episodes of drought (Pirasteh-Anosheh

et al., 2016). Reduced stomatal conductance has also been observed in species growing on low-N soils (see, Araus *et al.*, 2020), and stomata are more sensitive to drought when N is limiting (Pirasteh-Anosheh *et al.*, 2016). Thus, these are desirable above-ground traits for a hypothetical rice plant that is tolerant to low-N and drought conditions. Given that plants interact with soil N and water at the root surface, root responses to low-N and drought conditions must also be considered.

Below-ground, the accumulation of osmolytes has been shown to improve seedling tolerance to drought stress (Sharma *et al.*, 2019), where the osmolytes likely function in the maintenance of root growth allowing them to access water deeper in the soil profile (Serraj and Sinclair, 2002). Further, because increases in root biomass come with substantial energetic costs (Colombi *et al.*, 2019), we must consider that tolerance to combined drought and N deficit may be achieved by decreasing overall carbon investment in below-ground growth by changing the architecture of the root system (Lynch, 2019). For example, increasing the crown or axial root angle and axial root length while decreasing axial root thickness would allow a rice plant to access and obtain soil N and water deep in the soil profile (Lynch, 2022). In rice, the DRO1 locus simultaneously provides a mechanism for drought avoidance and N acquisition through the modification of rooting depth driven by increases in rooting angle (Uga *et al.*, 2013).

Clues to understanding how the drought × N deficit interaction plays out can be found in species tolerant to water and nitrogen deficits, individually or in combination. Perhaps the best candidates to explore mechanisms of abiotic stress tolerance in rice are the wild relatives that have adapted to numerous niches around the world (Atwell *et al.*, 2014). For example, *Oryza australiensis* is native to the northern savannahs of Australia and has proven to be a valuable resource for improving the tolerance of domestic rice (and other cereals) to stresses including heat, brown leaf hopper, and salt. This species is particularly interesting for understanding drought and low-N

tolerance in rice because: 1) *O. australiensis* has for a long time been assumed to be drought tolerant with, to the best of our knowledge, no evidence to support the claim; 2) salt tolerance is often a proxy for drought tolerance given the shared osmotic challenges and subsequent responses from tolerant plants experiencing drought; and 3) the soils in which it occurs naturally are nitrogen-poor (Spain *et al.*, 1983; ~1-6 mg N kg-1 soil, or 1-6 ppm; Smolinski *et al.*, 2015).

The present study aimed to explore both the individual and combined effects of drought and low nitrogen on the growth and physiology of domestic rice, *Oryza sativa*, and the wild rice relative *O. australiensis* (from Keep River National Park in the Northern Territory of Australia) seedlings (up to ~7 weeks old). Two cultivars of *O. sativa*, which are specialised to vastly different water management procedures, were used during these experiments: *O. sativa* cv. Calrose (the once predominant cultivar grown in Australia) was used for the individual stress experiments, while *O. sativa* cv. Nipponbare (the once predominant cultivar grown in Japan) was used for the fully factorial experiments. Australian rice is grown using ~50% of the water used by other rice farms (e.g., those in Japan) and so we were interested in the detection of cultivar-specific differences in response to low soil N and drought. The response of both cultivars to the individual stresses was essentially identical and so we focus here on the fully factorial experiments. We hypothesised that the wild rice would be more tolerant to changes in soil N and water because of its evolutionary history in the harsh conditions of northern Australia, such that *O. australiensis* would reveal traits similar to those described above for a hypothetical drought and low-N tolerant rice. State-of-the-art phenotyping infrastructure was used to measure proxies of above-ground growth, greenness, water use and root architecture. We also collected data on shoot and root biomass, leaf nitrogen levels, and chlorophyll content.

Materials and Methods

Plant Material and Growth Conditions

Individual stresses: In late February 2020, *Oryza sativa ssp*. *japonica cv*. Calrose and the wild species *O. australiensis* (sourced from Keep River, in Western Australia; 15° 58′ S, 129° 03′ E) were germinated at 34°C for two days. Germinated seed were sown (at a density of two seeds per pot) in 2.5 L pots containing a nitrogen-free soil mix (50% UC Davis sand + 35% cocopeat + 15% clay loam). Seedlings were grown for 15 days under a 30/22 °C day/night cycle in the Northwest Smart House at The Plant Accelerator (TPA; Australian Plant Phenomics Facility, University of Adelaide, Adelaide, Australia). During this time, the soil of all pots was kept moist by sitting each pot in a 20 mm x 20 mm tub. Water was supplied by filling the tub, not by watering the soil directly. After one week, seedlings were thinned to one plant per pot.

Combined stresses; These experiments were carried out in March 2022 and made use of *O. sativa ssp. japonica cv.* Nipponbare instead of *O. sativa ssp. japonica cv.* Calrose.

Treatments

Individual stresses: At 13 days after planting (DAP), control and 'drought' plants were watered once with 100 ml of a 42.85 mM ammonium nitrate (NH_4 NO_3) solution (120 mg N), while the Low N plants received 100 ml of a 10.7 mM NH_4 NO_3 solution (40 mg N). Pots were moved onto the TPA imaging system at 15 DAP. The experiment employed a split-unit design (i.e., the physical positioning of the pots), which was constructed and randomized using *dae* (Brien, 2020b), a package for the R statistical computing environment (R Core Team, 2020). Treatments were maintained and plants were imaged daily for 33 days (plants were 48 days old at the conclusion of the experiments). Water levels were monitored and adjusted daily by the Scanalyzer 3D

weighing and watering system (LemnaTec GmbH, Aachen, Germany). Pot water content was adjusted to a set target weight to ensure that moisture levels were held constant. Re-adding lost water enabled an estimate of water use during the experiments. Control and Low N plants were watered to achieve 24.3% soil moisture from 15 DAP. To induce moderate drought stress, water was initially withheld from the 'drought' cohort of plants from 15 DAP until the weight of the pot reached the target weight, and then they were watered to maintain the target soil moisture content of 12.2%.

Combined stresses: Nitrogen was applied and plants were moved onto the TPA imaging system and imaged for 32 days. Drought was induced as above, except this time water was additionally withheld from half of the low-N plants to generate a drought by nitrogen stress interaction.

*Projected Shoot Area (*PSA*) and Relative Growth Rates (*RGR*)*

Individual stresses: Red-Green-Blue (RGB)/visible spectrum cameras at TPA were used to image the plants daily for 33 days. Standard RGB images had a resolution of 8 M pixels (Berger *et al.*, 2012). From these images the Projected Shoot Area (PSA) of each plant was obtained. PSA is calculated as the sum of the areas, measured in kilopixels, from 3 camera views, comprising two side views and a view from above. Relative PSA was calculated by dividing treatment PSA values by the mean control PSA values. Due to significant species (i.e., biological) differences in PSA throughout the imaging period we found it necessary to calculate relative growth rates (RGRs) for each species/treatment group to determine the impact of treatment on growth. Changes in PSA through time were used to estimate RGRs for five-day windows (e.g., 15–20 DAP, 20–25 DAP, etc.) using the following formula:

$$RGR(day^{-1}) = \frac{\ln PSA_2 - \ln PSA_1}{t_2 - t_1}$$

Where PSA_1 is the PSA at an initial time (t_1), and PSA_2 is the PSA five days after t_1 (t_2).

Combined stresses: PSA and RGR were determined for plants grown under combined drought and low soil N stresses as above but for 32 days.

Greenness

Individual and Combined stresses: The 'greenness' of each plant, which is a proxy for chlorophyll content, was also determined from the high-throughput images described above. *RGB* values were converted to the Hue Saturation Intensity (HSI) colour management system using the formula:

$$\frac{RGB}{255} * 360$$

Where *RGB* refers to the *RGB* colour values. Dividing by 255 converts the *RGB* colour value to the 8-bit format. Multiplying by 360 converts the 8-bit format to a hue angle compliant with the HSI colour management system. Hue angle is then taken as a measure of greenness (Tackenberg, 2007).

Root and Shoot Fresh and Dry Weight

Individual stresses: After 33 days on TPA imaging system shoots were excised from the roots at soil level. Excised shoots were oven dried at 80ºC for 72 hours and weighed. Relative shoot dry weight was calculated by dividing the shoot dry weight of drought and low-N treatment plants by the mean shoot dry weight of the control groups for each species.

Combined stresses: Shoot fresh weight was determined for these plants before dry weight was determined as above but after 32 days of imaging. The roots were washed

clean of soil debris and weighed to estimate root fresh mass. Roots were oven dried as above to determine root dry weight.

Shoot dry weight (SDW) measurements were used to calibrate Projected Shoot Area (PSA) values. Correlation analysis was used to estimate a Pearson r coefficient and R2 for the relationship between SDW and PSA. Based on the results of this analysis, we show that PSA is a reliable proxy for SDW in these experiments ($p < 0.0001$; Pearson r = 0.98; R2 = 0.9683; Supplementary Figure 4.12). Thus, it was not necessary to convert PSA values into "digital biomass" to account for differences in leaf mass area and self-shading between the genotypes and/or treatments.

Leaf Nitrogen and Nitrogen Use Efficiency (NUE)

Combined stresses: Leaf nitrogen was estimated as a percent of leaf mass using a Rapid N Exceed nitrogen analyser (Elementar Analysensysteme GmbH, Langenselbold, Germany) with aspartic acid standards, as per manufacturer's instructions, by incinerating 150 mg of ground dried leaf material encapsulated in N-free paper (Thomas Scientific, Swedesboro, USA). A random subset of the six replicates were used for N analysis ($n = 3$). Leaf nitrogen levels and shoot dry weights were used to estimate NUE using the formula:

$$\frac{DW_{shoot}}{N_{shoot}}$$

Where, DW_{shoot} is the dry weight (in grams) of the above-ground material for the analysed samples, and N_{shoot} is the amount of N in the leaves (in grams) of the analysed samples.

Root Architecture

Combined stresses: Prior to drying the below-ground material, the intact washed roots were scanned with an Epson Expression 10000XL scanner (Epson, North Ryde, Australia). Intact root systems were floated in custom-built Perspex trays containing water to prevent dehydration and to allow individual roots to be detangled and imaged while minimising root overlapping. High quality images of the root systems were obtained and supplied to the program *RootGraph* (Cai *et al.*, 2015) for the estimation of axial and lateral root surface area, length, volume, and number, and the average axial root thickness.

Water Use and Water Use Efficiency (WUE)

Combined stresses: Since the Smart House at TPA automatically waters pots to a pre-defined weight (corresponding to a desired soil moisture level) daily and records the volume of water required by each pot. Thus, the volume of water used per unit leaf area, and the water use efficiency (*WUE*) can be estimated. Here, we estimated daily water use per unit shoot area using the following equation:

$$\frac{V_{water}}{PSA}$$

Where, V_{water} is the volume (ml) of water added to a pot each day, and PSA is the projected shoot area of the pot (expressed in kilopixels). *WUE* was estimated using the following equation:

$$\frac{PSA}{V_{water}}$$

Statistical Analyses

All data collected for the present study were analysed using either one-way, two-way or three-way ANOVAs. Duncan's Multiple Range post-hoc tests were performed to detect significant differences between the experimental groups of this study. Unless

otherwise specified, all statistics were performed in GraphPad Prism (v9.0.0) and GenStat (v19.1.0.21390). ANOVA results are summarised in Table 4.1.

Results

O. australiensis Seedlings Maintain Robust Growth Under Low N and Drought

Under control conditions, *O. australiensis* plants had smaller projected shoot area (PSA) compared to *O. sativa* throughout the imaging period (Fig. 4.1, Fig. 4.2a, Fig. 4.3a, Supplementary Figure 4.6, and Supplementary Figure 4.7a). Control plants of both species had similar relative growth rates (RGRs) throughout the imaging period (Fig. 4.2e and Supplementary Figure 4.7d).

Low-N soils significantly reduced the growth of well-watered *O. sativa* and *O. australiensis* plants (Fig. 4.1, Fig. 4.2b, Fig. 4.3a, Supplementary Figure 4.6, and Supplementary Figure 4.7b). After 34 days of growing on low-N soil (i.e., at 47 DAP; low-N fertiliser applied at 13 DAP), the PSA of *O. sativa* and *O. australiensis* plants was reduced by ~58 and ~74%, respectively (Fig. 4.2b, Fig. 4.3a; see also, Supplementary Figure 4.6, and Supplementary Figure 4.7a & b). While growth of *O. australiensis* was more strongly inhibited by low soil N later in the vegetative growth phase, the RGR of *O. australiensis* plants was significantly higher than *O. sativa* in the early vegetative growth phase (15-20 DAP; $p < 0.05$; Fig. 4.2f and Supplementary Figure 4.7e). However, the steep decline of RGR of *O. australiensis* between 20-30 DAP and the relatively small decline in *O. sativa* RGR over the same period resulted in *O. sativa* having a higher RGR ($p < 0.05$; Fig. 4.2f and Supplementary Figure 4.7e).

Drought conditions similarly reduced the growth of both *O. sativa* and *O. australiensis* (Fig. 4.1, Fig. 4.2c, Fig. 4.3a & c, and Supplementary Figure 4.7c and 8a & b). The domestic and wild rice species grew at similar rates, in absolute terms, under drought conditions and only diverged significantly at 47 DAP (Fig. 4.2c). After 33 days of drought (47 DAP), the PSA of *O. sativa* was reduced by 48% ($p < 0.05$) and *O. australiensis* by 53% ($p < 0.05$; Fig. 4.3a). The final difference (~23%) in PSA between droughted *O. sativa* and *O. australiensis* plants was significant ($p < 0.05$; Fig. 4.3a), though *O. australiensis* maintained growth up to 29 DAP while *O. sativa* suffered growth penalties as early as 20 DAP (Supplementary Figure 4.10a & b). The RGR of *O. australiensis* was significantly larger than *O. sativa* under drought conditions early in the vegetative growth phase but this dissipated after 20 DAP (Fig. 4.2g). The RGR of the droughted *O. sativa* group was significantly smaller than its control RGR from 25-35 (DAP Supplementary Figure 4.10c).

The individual effects of low-N and drought on above-ground phenotypes were additive when both stresses were imposed concurrently (Fig. 4.1, Fig. 4.2d, Fig. 4.3, and Supplementary Figure 4.10). At 47 DAP the PSA of *O. sativa* was reduced by ~77% relative to the controls under drought/low-N conditions, while the PSA of *O. australiensis* was reduced by ~85% relative to controls (Fig. 4.3a, and Supplementary Figure 4.10a & b). The PSA of droughted/low-N *O. sativa* plants first diverged from the control plants at 19 DAP (6 days after low-N fertiliser application, and 4 days after water was withheld; $p < 0.05$; Supplementary Figure 4.10a), whereas the *O. australiensis* group diverged from control at 25 DAP ($p < 0.05$; Supplementary Figure 4.10b). The RGR of *O. sativa* droughted/low-N plants was significantly lower than control plants between 15-40 DAP (Supplementary Figure 4.10c), while the RGR of *O. australiensis* was only significantly less than its control from 20-45 DAP (Supplementary Figure

4.10d). Both species stopped growing by 30 DAP (Supplementary Figure 4.10c & d), but growth resumed after 30 DAP in both species.

While *O. australiensis* exhibited signs of early vigour, this did not translate to growth benefits at 47 DAP. A two-way interaction between water status and soil N content indicated that the effect of drought on PSA at 47 DAP changed depending on the amount of N in the soil ($p < 0.01$; Fig. 4.3a). Here, while both drought and low N reduced the PSA of both species at 47 DAP ($p < 0.01$; $p < 0.01$, respectively), the greatest reduction in PSA was observed when both drought and low N co-occurred.

Interestingly, a two-way interaction between species and soil N content indicated that the root:shoot of *O. australiensis* plants was responsive to low N soils ($p < 0.05$; Fig. 4.3b-d). Root dry mass increased relative to shoot dry mass in *O. australiensis* when it was grown on low N soils, while *O. sativa* did not respond to any imposed stress (Fig. 4.3b). Changes in *O. australiensis* root:shoot was mostly driven by relatively large reductions in shoot dry weight on low N soils (Fig. 4.3c), though root dry weight was also responsive to the imposed stresses (Fig. 4.3d).

O. australiensis Plants Resist Changes in Greenness Under Low N and Drought

Control *O. australiensis* plants were significantly greener than *O. sativa* plants at 15 and 47 DAP (Fig. 4.2i, and Fig. 4.3e). Greenness tended to decline through the imaging period. The greenness of low-N grown *O. sativa* leaves first became significantly lower than that of control leaves at 21 DAP (i.e., 8 days after low-N fertiliser application), while low-N grown *O. australiensis* leaves were only consistently less green than control leaves at 36 DAP (i.e., 23 days after low-N fertiliser application; Supplementary Figure 4.10e & f). The greenness of droughted *O. australiensis* leaves never changed

relative to control plants throughout the imaging period (Supplementary Figure 4.10f), while the leaves of droughted *O. sativa* plants were less green than control leaves at 25 DAP (Supplementary Figure 4.10e). However, an increase in the greenness of droughted *O. sativa* leaves from 34 DAP onwards and the steady decrease in control leaf greenness meant that droughted leaves were significantly greener than control leaves from 41-47 DAP (Supplementary Figure 4.10e). Similar observations were made for the droughted/low-N *O. sativa* group, though this increase resulted in no difference in leaf greenness at 47 DAP (Supplementary Figure 4.10e). Changes in leaf greenness at 47 DAP largely follow changes in leaf N content (Fig. 4.3f). Interestingly, the nitrogen-use efficiency (NUE) of *O. australiensis* increased when soil N was low (Fig. 4.3g).

O. australiensis Root Systems Were Highly Responsive to Low N and Drought

Main effects of species, water status, and soil N content indicated that *O. sativa* had more axial roots than *O. australiensis* ($p < 0.01$; Fig. 4.4a), well-watered plants had more axial roots than droughted plants ($p < 0.01$), and plants grown in low-N soils had fewer axial roots than plants grown with adequate soil N ($p < 0.01$). There were significant interactions between species and water status ($p < 0.01$), species and soil N content ($p < 0.01$), and water status and soil N content ($p < 0.01$) for lateral root number (Fig. 4.4b). Low soil N and drought reduced the number of lateral roots in both *O. sativa* and *O. australiensis* root systems. The effects of low soil N and drought were additive for *O. sativa*, though not for *O. australiensis*, which had the same number of lateral roots under both low N and drought/low-N conditions. As a result, the total length of axial roots and lateral roots in the root systems of *O. sativa* and *O. australiensis* follow similar trends (Fig. 4.4c & d).

A significant interaction between species and water status ($p < 0.01$) for axial root SA:V indicated species specific responses to drought (Fig. 4.4e). The SA:V of *O. australiensis* axial roots increased significantly in response to drought ($p = 0.02$), while no change in axial root SA:V was detected for *O. sativa* under drought. The axial roots of *O. australiensis* had a significantly lower SA:V compared to *O. sativa* when both species were well-watered ($p < 0.01$), though this difference was lost when drought was imposed. There was a main effect of soil N content on axial root SA:V ($p < 0.01$), in which SA:V increased in plants grown on low-N soils.

There was a significant interaction between species and water status for lateral root SA:V ($p < 0.01$; Fig. 4.4f). The interaction comes about because under well-watered conditions the SA:V of *O. sativa* and *O. australiensis* was significantly different ($p < 0.05$) but this difference was lost under drought due to a significant decrease in *O. sativa* SA:V. There was also a main effect of soil N content on lateral root SA:V ($p < 0.01$), in which the SA:V of lateral roots grown on low-N soils was significantly larger than plants grown on control N soils.

A significant interaction between species and water status suggested that the thickness of axial roots of *O. sativa* and *O. australiensis* changed differently in response to drought ($p < 0.01$; Fig. 4.4g). The axial roots of *O. australiensis* became significantly thinner under drought conditions ($p < 0.05$), while the thickness of *O. sativa* axial roots did not change in response to drought. Under well-watered conditions, the axial roots of *O. australiensis* were significantly thicker than those of *O. sativa*, though this difference was lost under drought. There was a significant main effect of soil N content on axial root thickness ($p < 0.01$), in which the axial roots of plants grown on low-N soils were significantly thinner than those grown on control N soils.

There was a significant three-way interaction between species, water status, and soil N content for average axial root length ($p < 0.05$). Thus, the interaction between species and water status ($p < 0.01$) changed depending on the amount of N in the soil. Here, low-N soils reduced the length of individual axial roots in the root system of well-watered *O. australiensis* plants but not *O. sativa*, and a non-significant increase in individual axial root length occurred when *O. australiensis* plants were grown on low-N soils and were exposed to drought. Overall, the species by water status interaction showed that the length of the average axial root increased under drought for *O. australiensis,* but no change was observed for *O. sativa*.

Droughted O. australiensis Plants Use Less Water Than O. sativa Plants, But Uses More Water Under Low N Conditions

Low soil N and drought appear to affect how the two species use water differently throughout the imaging period (Fig. 4.5a-d). Water use did not appear to be different between *O. sativa* and *O. australiensis* under control conditions (Fig. 4.5a). However, when soil N was low under both well-watered and drought conditions, *O. australiensis* water consumption appeared to increase, while no change was observed for *O. sativa* (Fig. 4.5b & d). Under drought conditions, water use was lower for both species relative to control, but *O. australiensis* water use was lower than *O. sativa* (Fig. 4.5c).

A three-way ANOVA was used to examine the interaction and main effects of species, water status, and soil N content on the amount of water used by plants at 47 DAP (Fig. 4.5e). There was a significant species by soil N content interaction ($p < 0.01$), which indicated that *O. sativa* and *O. australiensis* used water differently when soil N was low. When soil N was low, *O. australiensis* used more water than control-N plants ($p <$

0.01), while *O. sativa* did not change water use under low-N soils. Thus, *O. australiensis* consumed more water per kpixel than *O. sativa* under low-N conditions ($p < 0.01$).

A three-way ANOVA was used to examine the interaction and main effects of species, water status, and soil N content on water use efficiency (WUE; Fig. 4.5f). No significant interactions were detected, however there were main effects of water ($p < 0.01$) and nitrogen ($p < 0.01$). WUE was higher under drought conditions but lower under low soil N conditions.

Discussion

We set out to identify some optimal traits for the selective breeding of a rice genotype that is tolerant to low soil N and drought. More stringent agronomic constraints in rice cultivation will be driven by fertiliser costs and erratic climate regimes, amplifying the need for more stress-tolerant cultivars. Literature searches identified several above- and below-ground traits that are candidates for this goal, informing the detailed phenotypic analysis reported here using The Plant Accelerator. These earlier reports include early seedling vigour (ESV; as measured by rapid above-ground growth (RAG)), maintenance of green tissue, accumulation of osmolytes, reduction in root system size, but an increased proportion of long axial roots with high SA:V, and increased crown root angle. In the present study, we report on the identification of some of these traits in *O. australiensis*, a wild rice relative from the monsoonal northern savannah of Australia, with the reasoning that some of these traits have been naturally selected in wild relatives but not *O. sativa*.

Oryza australiensis had robust growth and maintained leaf greenness early in its development under individual stresses, as well as when low-N and drought were combined. Notwithstanding it being an unselected wild relative, under control and individual low-N and drought conditions *O. australiensis* had higher RGR than *O. sativa*. Rapid early growth, often termed as 'early vigour', is an important trait for crop species, especially in dry-seeded cereals. It is characterised by leaf area expansion, allowing a seedling to become photosynthetically competent and competitive with weeds (Mahender *et al.*, 2015). Once leaf area has been established plants are also less susceptible to climatic variation because they can access soil water and nutrients (Reynolds *et al.*, 2007; Mahender *et al.*, 2015). Given the high variability and small volume of rainfall that Kununurra receives in the months when *O. australiensis* seed is germinating and seedlings are growing (see Literature Review section: "*Oryza australiensis* is a Wild Rice Relative with Multi-Stress Tolerance"), rapid seedling growth would allow it to develop deep roots and extract sparse soil water and limited soil N. We therefore speculate that the tolerance of the seedlings to low soil N and drought (as determined by PSA expressed relative to control plants, and RGR) come about because of superior shoot and root development (i.e., early vigour) when the plants were young. For instance, *O. australiensis* axial roots were longer than *O. sativa* axial roots under low-N and drought, and their thickness was much more responsive to both conditions than *O. sativa*. This could be characterised as a stress-avoidance strategy allowing *O. australiensis* to access and, importantly, capture water and N deep in the soil profile.

However, while fast-growing 'deep root' phenotypes are beneficial in a natural soil profile, the advantage this trait confers might be less obvious in the restricted soil volume of a pot experiment (Passioura, 2006; Poorter *et al.*, 2016). In the experiments

reported here, *O. australiensis* appeared to be vulnerable to low-N and drought conditions if they persist into the late stages of root development (beyond ~25 DAP), while *O. sativa* demonstrated some ability to 'recover' after soil water status equilibrated to 12% SWC. Further, we report what appears to be sensitivity to low-N in 47-day old *O. australiensis* seedlings – that is, while both low-N and drought substantially affected the physiology of *O. australiensis*, the impact of low-N soils was often far greater than water deficits (Fig. 4.3 & 4). It is possible that the 'susceptibility' of *O. australiensis* shoot growth to low-N was at least in part because of increased investment below-ground (i.e., changing root:shoot; Fig. 4.3b). Furthermore, the restrictive soil volumes in the 2 L pots used in this study would also explain why the effects of low-N and drought were more severe for *O. australiensis* late in the last stages of the experiment. The long axial roots of *O. australiensis* have likely evolved to access resources deep in their native uncultivated soil profiles, a phenomenon that would have been neutralised in the pots of soil.

Root systems are metabolically costly to maintain because they impose a significant respiratory burden (Colombi *et al.*, 2019). During times of stress, particularly water stress, reducing total root system size but altering other features of the root system to optimise resource acquisition is essential (Fry *et al.*, 2018; Lynch, 2022; Liao *et al.*, 2022). We found that *O. australiensis* had a smaller root system (dry mass basis) than *O. sativa* under control conditions, with fewer roots and shorter *total* length of axial and lateral roots. This contrast in root dry mass and axial root number and length disappeared under the droughted conditions demonstrating that drought had a greater impact upon roots (and shoots) of *O. sativa* than *O. australiensis*. Previously, we found that the size of rice root systems is responsive to growth temperature, and that these species have different temperature optima for root growth (Phillips *et al.*, 2022). Thus,

it may be that in the present study the growth temperature (30°C) was not optimal for *O. australiensis* root growth. Interestingly, *O. sativa* root systems included more lateral roots than *O. australiensis* in every treatment group, suggesting it could extract nutrients and water intensively from a smaller soil volume. While both *O. sativa* and *O. australiensis* root systems reduced in size, some qualitative changes in root architecture in *O. australiensis* were consistent with tolerance to low-N and drought. For example, the axial roots of *O. australiensis* became thinner under low-N and drought conditions, with a higher SA:V, and drought caused the already long axial roots to grow even longer. In contrast, the thickness, SA:V, and length of *O. sativa* axial roots, which were relatively thin and short in control pots, did not change in response to drought, though the axial root SA:V of *O. sativa* grown on low-N soils increased in part because the roots became slightly thinner. Further, the number of lateral roots relative to the number of axial roots appeared to increase in *O. australiensis* under drought but not in *O. sativa* (Supplementary Figure 4.11). Responsiveness of root traits to drought has been associated with tolerant rice genotypes, while susceptible rice varieties had static responses to drought (Henry *et al.*, 2012). Long, deep-drilling roots under control conditions were associated with high and stable yield for drought tolerant *aus* rice genotypes (Liao *et al.*, 2022). However, the roots of these plants grew thicker than 'control roots' and so grain yield under drought was also positively associated with thick roots, not thinner, because of their branching potential (Liao *et al.*, 2022). The discrepancy between the response of root thickness in the 'drought tolerant' *aus* genotypes with that of *O. australiensis* may come down to differences in age/development. The plants in the present study were sampled at 47 DAP, while those of Liao et al. (2022) were sampled at 70 DAP. The significance of deep rooting lies in access to resources deep in the soil profile, while the significance of a high SA:V in deep roots (i.e., relatively thin roots) may lie in the uptake of those resources (Lynch,

2022). With a larger SA:V under low-N and drought, *O. australiensis* axial roots could increase the uptake of N (as well as other resources) from the soil. Interestingly, the number of lateral roots increased relative to the number of axial roots in *O. australiensis* but not *O. sativa* under drought. With higher lateral:axial, *O. australiensis* roots may have been able to absorb more water. This is in accordance with the observation of Liao et al. (2022), who showed that root branching was associated with yield stability under drought. However, the present study included pot experiments, which come with significant limitations (Passioura, 2006), and water was automatically applied from the top of the soil to achieve a target pot weight. So, it may be that droughted *O. australiensis* plants increased the number of lateral roots at the soil surface to capitalise on water when it became available. This trait may have been naturally selected in *O. australiensis* to allow it to rapidly absorb water in the soil surface following monsoonal rain events.

We also found that, while NUE increased in *O. australiensis* plants grown under low-N conditions, WUE declined because of increased water use per unit leaf area. During individual stress events, plants can adjust the consumption of specific resources to survive. When exposed to drought, plants typically act to conserve limited water by reducing stomatal conductance and therefore transpiration (Praba *et al.*, 2009; Xiong *et al.*, 2012; Ouyang *et al.*, 2017). Similarly, rice plants grown in low-N soils tend to reduce water consumption by reducing stomatal conductance (Xiong *et al.*, 2015). Water uptake by wheat is also reduced when grown on low-N soils (Barraclough *et al.*, 1989). In this study, while water use per unit leaf area changed over the course of the measurement period (measured from 30 to 47 DAP, when daily watering was relatively consistent for all groups), there were strong species differences and treatment effects (Fig. 4.5). Interestingly, *O. sativa* did not adjust its water

consumption in response to individual low-N and drought stress, though water use appears to increase slightly when both stresses are imposed concurrently. *Oryza australiensis* was particularly responsive to low-N soils where, overall, water use per unit leaf area increased substantially under both well-watered and drought conditions. Conversely, the amount of water consumed per unit leaf area was typically lower across time for *O. australiensis* plants grown under drought (up to ~30% lower at 32 DAP) when they had access to sufficient soil N. While these reductions in water use under drought, particularly early on (~30-42 DAP), are small, they demonstrate an adaptive response of *O. australiensis* to water limiting conditions. In the absence of gas exchange data, and thus with no information on the response of stomatal conductance (nor rates of transpiration) to the imposed stresses, it is difficult to understand the observed water use of *O. australiensis* in response to low soil N. It is generally accepted that there is a positive relationship between leaf nitrogen content and stomatal conductance in rice (Ishihara *et al.*, 1979). However, if we accept water use as a proxy for stomatal conductance to water, the leaf N (Fig. 4.3f) and water use (Fig. 4.5e) data from 47 DAP suggest that *O. australiensis* breaks from this traditional view of rice response to low soil N (Ishihara *et al.*, 1979). Given the intimate link between water movement and N uptake and distribution (Plett *et al.*, 2020), it may be that *O. australiensis* increases water movement through its body to facilitate N translocation. This might explain why *O. australiensis* tended to have higher NUE but lower WUE when grown on soils with low-N content (Fig. 4.3e and Fig. 4.5f), though this hypothesis needs further investigation.

Even though *O. australiensis* suffered growth penalties under individual and combined low-N and drought conditions, the results presented here indicate that *O. australiensis* will be a useful resource for traits associated with seedling level tolerance

to low-N and drought conditions. The reported changes in root architecture require further investigation, particularly under field conditions or in deeper pots, though the preliminary data presented here suggest that *O. australiensis* may also possess tolerance traits later in its development.

Acknowledgements

We thank the hardworking staff at The Plant Accelerator for their support throughout these experiments. In particular, we thank Lidia Michis for the help she gave during the root scanning part of the project. We thank Vanessa Milano for valuable discussions about plant nitrogen use.

Author contributions

AP, BA, RB, DP, and BB conceived the study. AP established the plants used in this study. AP and DP harvested root material. AP took all measurements (with the help of LM for root architecture – see acknowledgements) analysed all data and prepared all figures. AP wrote the initial manuscript. BA and RB edited the manuscript.

Conflict of interest

The authors declare no conflicts of interest.

Funding

This work was supported by an Australian Plant Phenomics Facility PhD Internship. AP acknowledges support from an Australian Research Training Program Scholarship and the FJ Sandoz Scholarship from the University of Adelaide.

Data availability

All data available upon request.

References

Araus V, Swift J, Alvarez JM, Henry A, Coruzzi GM. 2020. A balancing act: how plants integrate nitrogen and water signals. Journal of Experimental Botany **71**, 4442–4451.

Atwell BJ, Wang H, Scafaro AP. 2014. Could abiotic stress tolerance in wild relatives of rice be used to improve Oryza sativa? Plant Science **215–216**, 48–58.

Barraclough PB, Kuhlmann H, Weir AH. 1989. The Effects of Prolonged Drought and Nitrogen Fertilizer on Root and Shoot Growth and Water Uptake by Winter Wheat. Journal of Agronomy and Crop Science **163**, 352–360.

Berger B, de Regt B, Tester M. 2012. High-Throughput Phenotyping of Plant Shoots BT - High-Throughput Phenotyping in Plants: Methods and Protocols. In: Normanly J, ed. Totowa, NJ: Humana Press, 9–20.

Cai J, Zeng Z, Connor JN, Huang CY, Melino V, Kumar P, Miklavcic SJ. 2015. RootGraph: a graphic optimization tool for automated image analysis of plant roots. Journal of Experimental Botany **66**, 6551–6562.

Cao X, Zhu C, Zhong C, Hussain S, Zhu L, Wu L, Jin Q. 2018. Mixed-nitrogen nutrition-mediated enhancement of drought tolerance of rice seedlings associated with photosynthesis, hormone balance and carbohydrate partitioning. Plant Growth Regulation **84**, 451–465.

Chen Z, Wang Q, Ma J, Zou P, Jiang L. 2020. Impact of controlled-release urea on rice yield, nitrogen use efficiency and soil fertility in a single rice cropping system. Scientific Reports **10**, 10432.

Colombi T, Herrmann AM, Vallenback P, Keller T. 2019. Cortical Cell Diameter Is Key To Energy Costs of Root Growth in Wheat. Plant Physiology **180**, 2049–2060.

Deng J, Feng X, Wang D, Lu J, Chong H, Shang C, Liu K, Huang L, Tian X, Zhang Y. 2020. Root morphological traits and distribution in direct-seeded rice under dense planting with reduced nitrogen. PLoS ONE **15**, e0238362.

Farooq M, Siddique KHM, Rehman H, Aziz T, Lee DJ, Wahid A. 2011. Rice direct seeding: Experiences, challenges and opportunities. Soil and Tillage Research **111**, 87–98.

Feng B, Chen K, Cui Y, *et al.* 2018. Genetic Dissection and Simultaneous Improvement of Drought and Low Nitrogen Tolerances by Designed QTL Pyramiding in Rice . Frontiers in Plant Science **9**.

Fry EL, Evans AL, Sturrock CJ, Bullock JM, Bardgett RD. 2018. Root architecture governs plasticity in response to drought. Plant and Soil **433**, 189–200.

Henry A, Cal AJ, Batoto TC, Torres RO, Serraj R. 2012. Root attributes affecting water uptake of rice (Oryza sativa) under drought. Journal of Experimental Botany **63**, 4751–4763.

Homma K, Horie T, Shiraiwa T, Supapoj N, Matsumoto N, Kabaki N. 2003. Toposequential Variation in Soil Fertility and Rice Productivity of Rainfed Lowland Paddy Fields in Mini-Watershed *(Nong)* in Northeast Thailand. Plant Production Science **6**, 147–153.

Ishihara K, Iida O, Hirasawa T, Ogura T. 1979. Relationship between Nitrogen Content in Leaf Blades and Photosynthetic Rate of Rice Plants with Reference ot Stomatal Aperture and Conductance. Japanese journal of crop science **48**, 543–550.

Kumar A, Verulkar S, Dixit S, Chauhan B, Bernier J, Venuprasad R, Zhao D, Shrivastava MN. 2009. Yield and yield-attributing traits of rice (Oryza sativa L.) under lowland drought and suitability of early vigor as a selection criterion. Field Crops Research **114**, 99–107.

Liao Q, Chebotarov D, Islam MS, *et al.* 2022. Aus rice root architecture variation contributing to grain yield under drought suggests a key role of nodal root diameter class. Plant, Cell & Environment **45**, 854–870.

Lipiec J, Doussan C, Nosalewicz A, Kondracka K. 2013. Effect of drought and heat stresses on plant growth and yield: A review. International Agrophysics **27**, 463–477.

Lozano YM, Aguilar-Trigueros CA, Flaig IC, Rillig MC. 2020. Root trait responses to drought are more heterogeneous than leaf trait responses. Functional Ecology **34**, 2224–2235.

Lynch JP. 2019. Root phenotypes for improved nutrient capture: an underexploited opportunity for global agriculture. New Phytologist **223**, 548–564.

Lynch JP. 2022. Harnessing root architecture to address global challenges. The Plant Journal **109**, 415–431.

Mahender A, Anandan A, Pradhan SK. 2015. Early seedling vigour, an imperative trait for direct-seeded rice: an overview on physio-morphological parameters and molecular markers. Planta **241**, 1027–1050.

Ouyang W, Struik PC, Yin X, Yang J. 2017. Stomatal conductance, mesophyll conductance, and transpiration efficiency in relation to leaf anatomy in rice and wheat genotypes under drought. Journal of Experimental Botany **68**, 5191–5205.

Passioura JB. 2006. The perils of pot experiments. Functional Plant Biology **33**, 1075–1079.

Peng J, Feng Y, Wang X, Li J, Xu G, Phonenasay S, Luo Q, Han Z, Lu W. 2021. Effects of nitrogen application rate on the photosynthetic pigment, leaf fluorescence

characteristics, and yield of indica hybrid rice and their interrelations. Scientific Reports **11**, 7485.

Phillips AL, Scafaro AP, Atwell BJ. 2022. Photosynthetic traits of Australian wild rice (Oryza australiensis) confer tolerance to extreme daytime temperatures. Plant Molecular Biology.

Pirasteh-Anosheh H, Saed-Moucheshi A, Pakniyat H, Pessarakli M. 2016. Stomatal responses to drought stress. Water Stress and Crop Plants. John Wiley & Sons, Ltd, 24–40.

Plett DC, Holtham LR, Okamoto M, Garnett TP. 2018. Nitrate uptake and its regulation in relation to improving nitrogen use efficiency in cereals. Seminars in Cell & Developmental Biology **74**, 97–104.

Plett DC, Ranathunge K, Melino VJ, Kuya N, Uga Y, Kronzucker HJ. 2020. The intersection of nitrogen nutrition and water use in plants: new paths toward improved crop productivity. Journal of Experimental Botany **71**, 4452–4468.

Poorter H, Fiorani F, Pieruschka R, Wojciechowski T, van der Putten WH, Kleyer M, Schurr U, Postma J. 2016. Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. New Phytologist **212**, 838–855.

Praba ML, Cairns JE, Babu RC, Lafitte HR. 2009. Identification of Physiological Traits Underlying Cultivar Differences in Drought Tolerance in Rice and Wheat. Journal of Agronomy and Crop Science **195**, 30–46.

Quinones C, Mattes N, Faronilo J, Sudhir-Yadav, Jagadish KSV. 2017. Drought Stress Reduces Grain Yield by Altering Floral Meristem Development and Sink Size under Dry-Seeded Rice Cultivation. Crop Science **57**, 2098–2108.

Reynolds MP, Pierre CS, Saad ASI, Vargas M, Condon AG. 2007. Evaluating Potential Genetic Gains in Wheat Associated with Stress-Adaptive Trait Expression in Elite Genetic Resources under Drought and Heat Stress. Crop Science **47**, S-172-S-189.

Saha S, Begum HH, Nasrin S, Samad R. 2020. Effects of drought stress on pigment and protein contents and antioxidant enzyme activities in five varieties of rice (Oryza sativa L.). Bangladesh Journal of Botany **49**, 997–1002.

Serraj R, Sinclair TR. 2002. Osmolyte accumulation: can it really help increase crop yield under drought conditions? Plant, Cell & Environment **25**, 333–341.

Sharma A, Shahzad B, Kumar V, Kohli SK, Sidhu GPS, Bali AS, Handa N, Kapoor D, Bhardwaj R, Zheng B. 2019. Phytohormones Regulate Accumulation of Osmolytes Under Abiotic Stress. Biomolecules **9**, 285.

Smolinski H, Pathan S, Galloway P, Kuswardiyanto K, Laycock J. 2015. *'Cockatoo Sands in the Victoria Highway and Carlton Hill areas, East Kimberley: land capability assessment for developing irrigated agriculture', Resource management technical report 391*. Perth.

Spain A, Isbell RF, Probert ME. 1983. Organic matter contents of Australian soils. Soils: An Australian Viewpoint. Melbourne: CSIRO, 551–563.

Tackenberg O. 2007. A New Method for Non-destructive Measurement of Biomass, Growth Rates, Vertical Biomass Distribution and Dry Matter Content Based on Digital Image Analysis. Annals of Botany **99**, 777–783.

Uga Y, Sugimoto K, Ogawa S, *et al.* 2013. Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. Nature Genetics **45**, 1097–1102.

Xin W, Zhang L, Gao J, Zhang W, Yi J, Zhen X, Bi C, He D, Liu S, Zhao X. 2021. Adaptation Mechanism of Roots to Low and High Nitrogen Revealed by Proteomic Analysis. Rice **14**, 5.

Xiong Q, Tang G, Zhong L, He H, Chen X. 2018. Response to Nitrogen Deficiency and Compensation on Physiological Characteristics, Yield Formation, and Nitrogen Utilization of Rice. Frontiers in Plant Science **9**.

Xiong D, Yu T, Ling X, Fahad S, Peng S, Li Y, Huang J. 2015. Sufficient leaf transpiration and nonstructural carbohydrates are beneficial for high-temperature tolerance in three rice (Oryza sativa) cultivars and two nitrogen treatments. Functional Plant Biology **42**, 347.

Xiong J, Zhang L, Fu G, Yang Y, Zhu C, Tao L. 2012. Drought-induced proline accumulation is uninvolved with increased nitric oxide, which alleviates drought stress by decreasing transpiration in rice. Journal of Plant Research **125**, 155–164.

Zhang F, Cui Z, Chen X, *et al.* 2012. Chapter one - Integrated Nutrient Management for Food Security and Environmental Quality in China. In: Sparks DLBT-A in A, ed. Academic Press, 1–40.

Zhang X, Davidson EA, Mauzerall DL, Searchinger TD, Dumas P, Shen Y. 2015. Managing nitrogen for sustainable development. Nature **528**, 51–59.

Zhang R, Hussain S, Wang Y, Liu Y, Li Q, Chen Y, Wei H, Gao P, Dai Q. 2021. Comprehensive Evaluation of Salt Tolerance in Rice (Oryza sativa L.) Germplasm at the Germination Stage. Agronomy **11**, 1569.

Zhang Z-H, Qu Z-S, Wan S, Chen L-H, Zhu Y-G. 2005. Comparison of QTL Controlling Seedling Vigour under Different Temperature Conditions Using Recombinant Inbred Lines in Rice (Oryza sativa). Annals of Botany **95**, 423–429.

Zhang J, Zhang S, Cheng M, Jiang H, Zhang X, Peng C, Lu X, Zhang M, Jin J. 2018. Effect of Drought on Agronomic Traits of Rice and Wheat: A Meta-Analysis. International Journal of Environmental Research and Public Health **15**, 839.

Zhi-Hong Z, Li S, Wei L, Wei C, Ying-Guo Z. 2005. A major QTL conferring cold tolerance at the early seedling stage using recombinant inbred lines of rice (Oryza sativa L.). Plant Science **168**, 527–534.

Tables and Figures

Table 4.1. ANOVA summary table for the interactions and main effects of species (A), water status (B), and soil N content (C) on the above and below-ground traits measured at 47 DAP.

	Species (A)	Water (B)	Nitrogen (C)	A × B	A × C	B × C	A × B × C
PSA	****	****	****	ns	ns	****	ns
R:S Dry Weight	**	ns	****	ns	*	ns	ns
Shoot Dry Weight (SR)	****	****	****	*	*	****	ns
Root Dry Weight (SR)	****	****	****	*	ns	***	ns
Hue Angle	**	*	****	***	ns	***	ns
Leaf N	**	**	****	**	ns	*	ns
Nitrogen Use Efficiency (LN)	**	ns	****	**	*	*	ns
No. Axial Roots (SR)	****	****	***	ns	ns	ns	ns
No. Lateral Roots	****	****	****	****	**	***	ns
No. Lateral:No. Axial (LN)	ns	ns	ns	*	ns	ns	ns
Total Axial Root Length	****	****	****	**	ns	***	ns
Total Lateral Root Length	****	****	****	****	****	****	*
Axial Root SA:V	**	ns	****	**	ns	ns	ns
Lateral Root SA:V	ns	ns	****	**	ns	ns	ns
Axial Root Thickness	****	*	****	*	ns	ns	ns
Av. Axial Root Length (LN)	***	ns	ns	**	ns	ns	*
Water Use (OR)	**	ns	****	ns	***	ns	ns
Water Use Efficiency	ns	****	****	ns	ns	ns	ns

*p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001. (SR) = square-root transformed; (OR) = outliers removed by Grubbs test; (LN) = natural logarithm transformed.

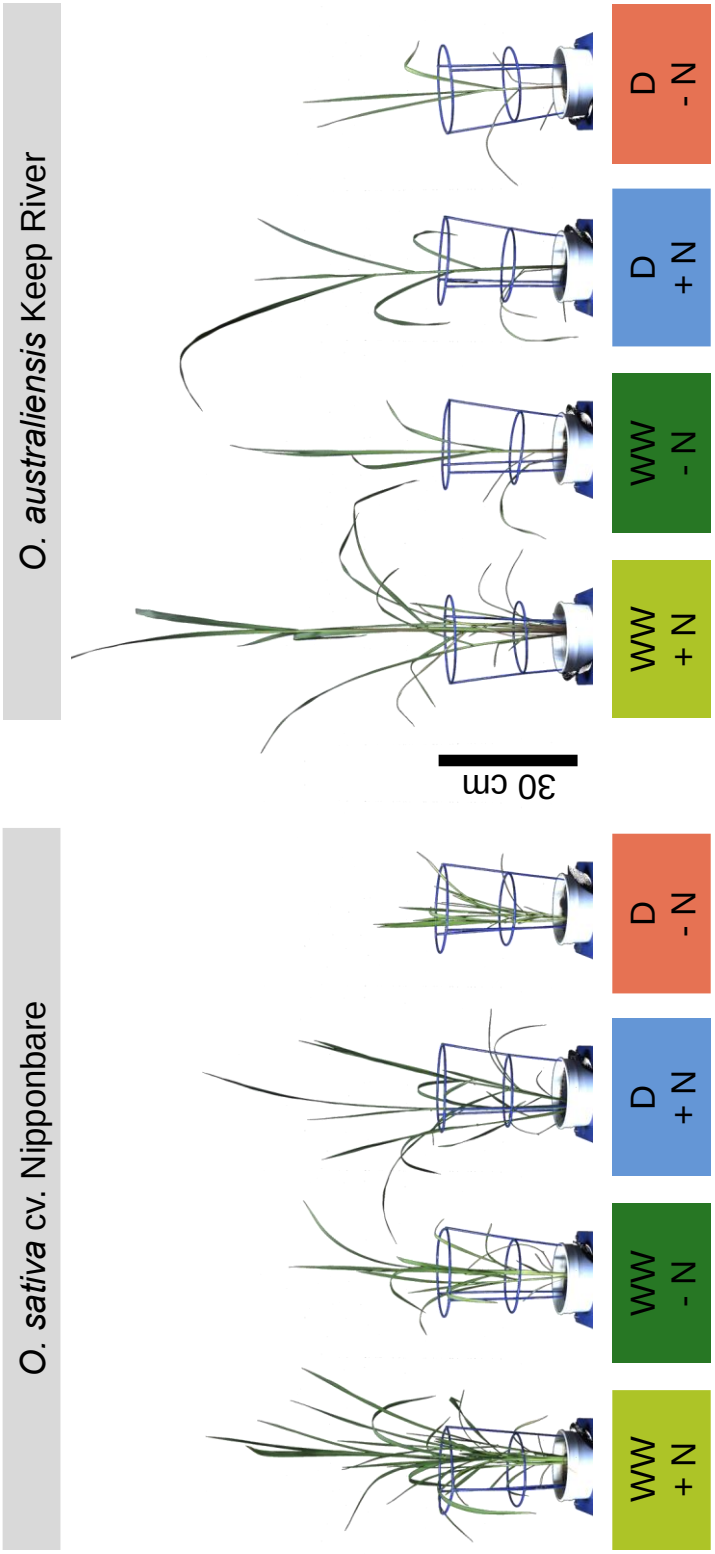


Figure 4.1. Representative images of *O. sativa* and *O. australiensis* plants grown under different combinations of water and nitrogen availability at 47 DAP (32 days after imaging started). Images were obtained from The Plant Accelerator.

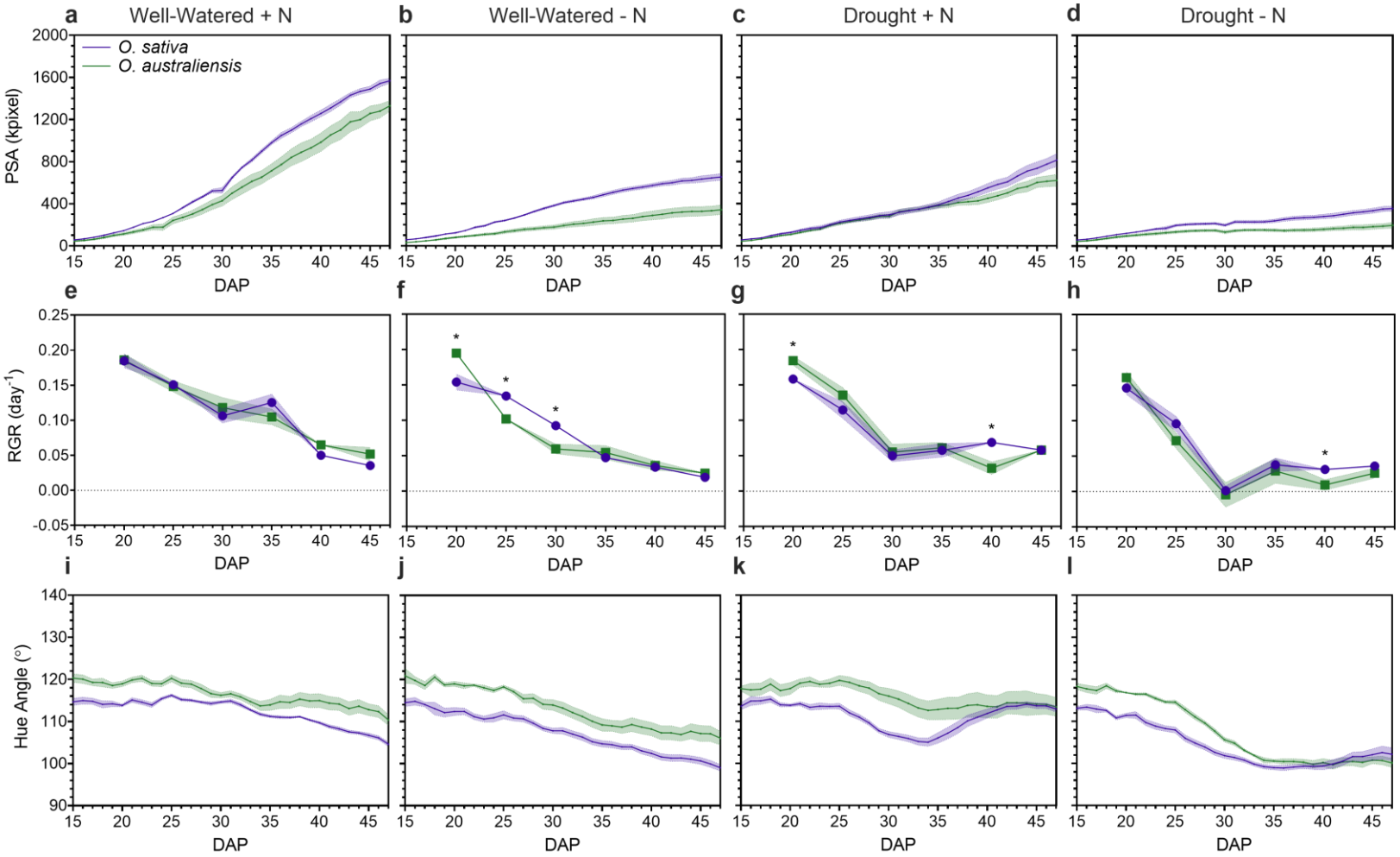


Figure 4.2. Mean (SE; n = 6) projected shoot area (PSA; a-d), relative growth rate (RGR; e-h), and hue angle (greenness; i-l) of *O. sativa* and *O. australiensis* plants grown under different combination of water and nitrogen availability for 33 days (15-47 DAP).

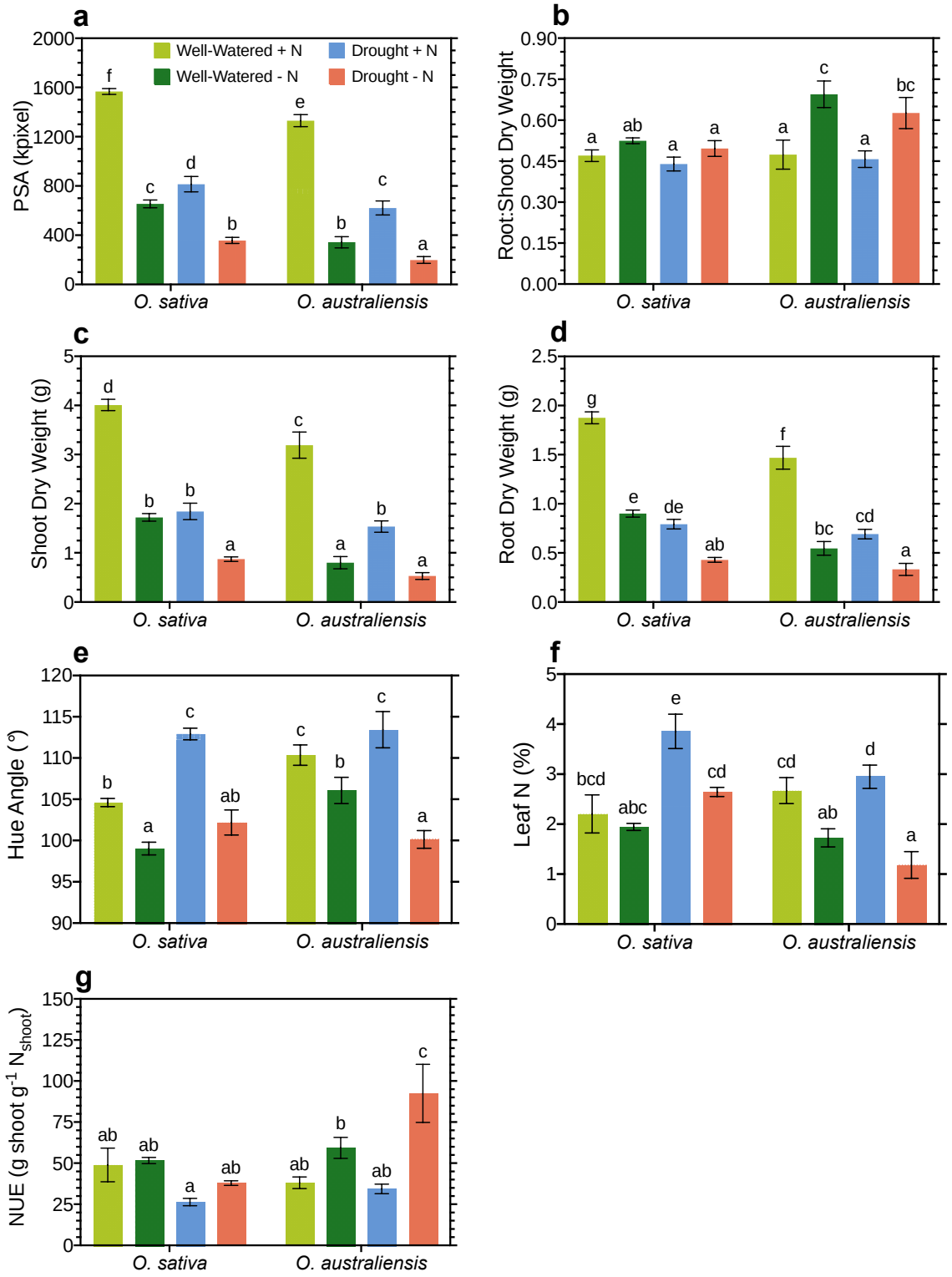


Figure 4.3. Responses of *O. sativa* and *O. australiensis* plants to different combinations of water and nitrogen availability 47 DAP. (a) PSA, (b) root to shoot dry weight ratio, (c) shoot dry weight, (d) root dry weight, (e) hue angle (greenness), (d) leaf N content, and (g) nitrogen use efficiency. Values are means SE (n = 6). Letters denote means that differ significantly, as determined by Duncan's Multiple Range Test.

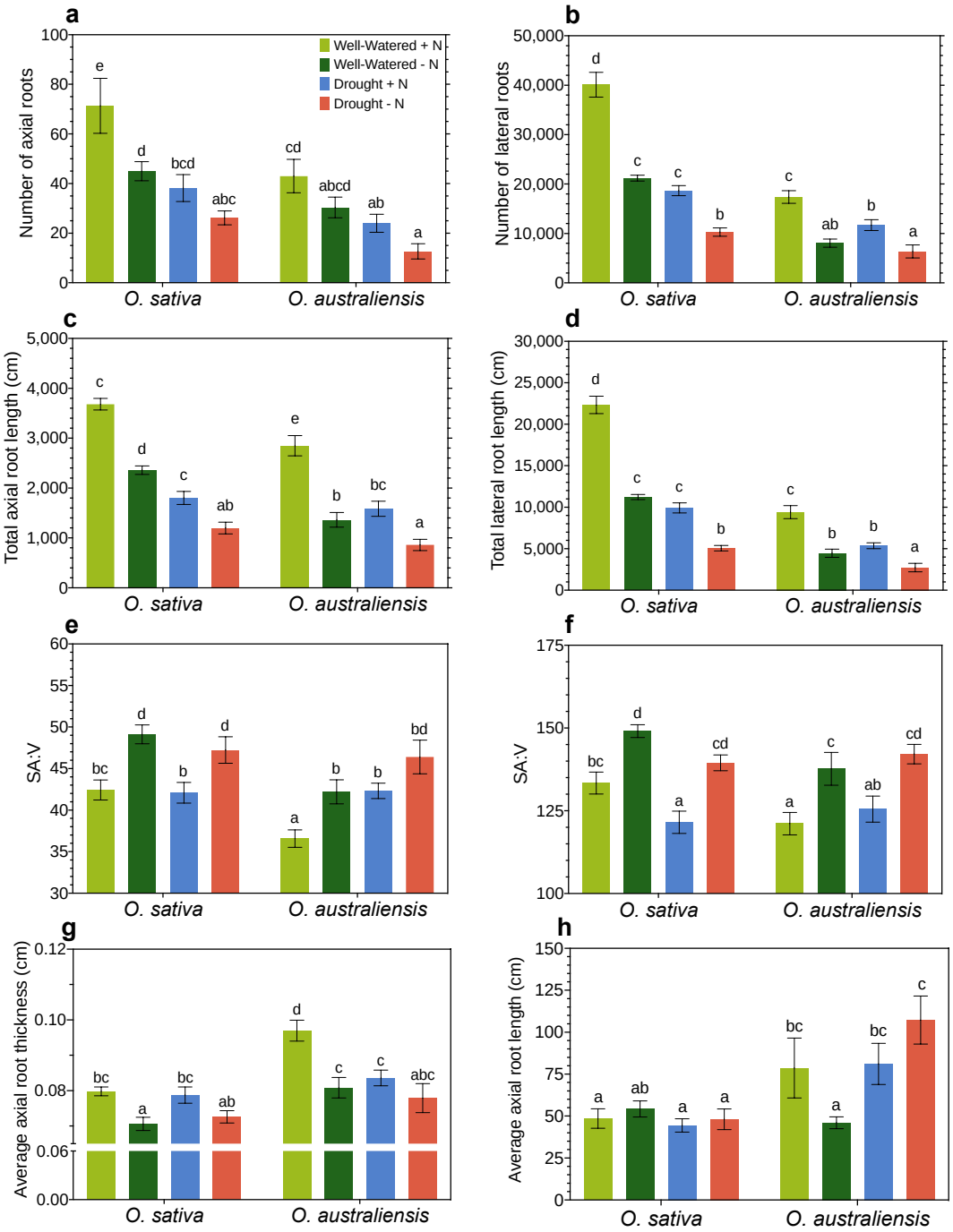


Figure 4.4. Root architecture responses of *O. sativa* and *O. australiensis* plants grown under different combinations of water and soil N availabilities for 33 days. Values are means SE (n = 4-6) for (a-b) number of axial and lateral roots, (c-d) total axial and lateral root length, (e-f) axial and lateral root surface area to volume ratio, (g) the average thickness of individual axial roots, and (h) the average length of individual axial roots. Outliers were identified and removed using Grubb's test. Letters denote means that differ significantly, as determined by Duncan's Multiple Range Test.

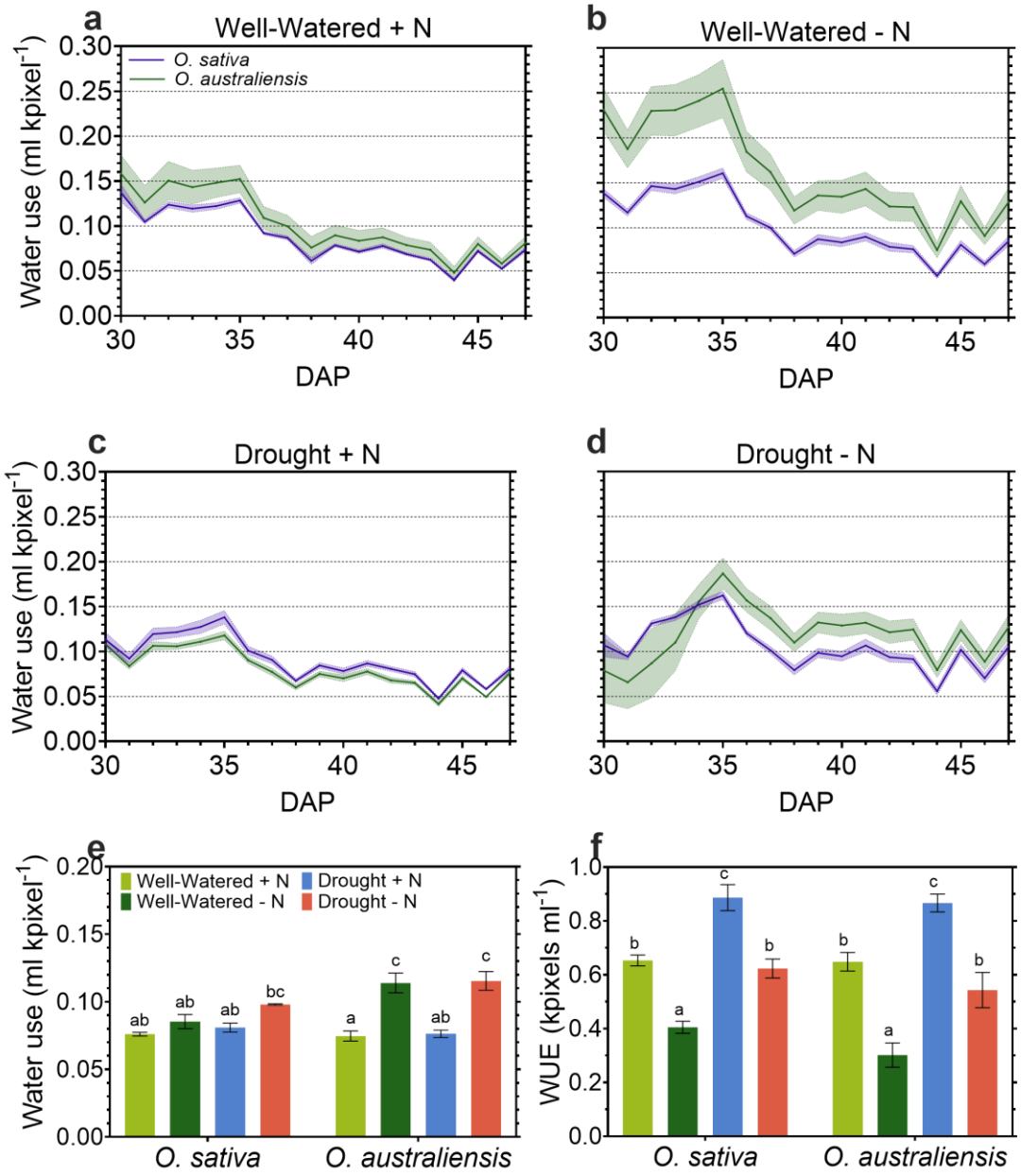
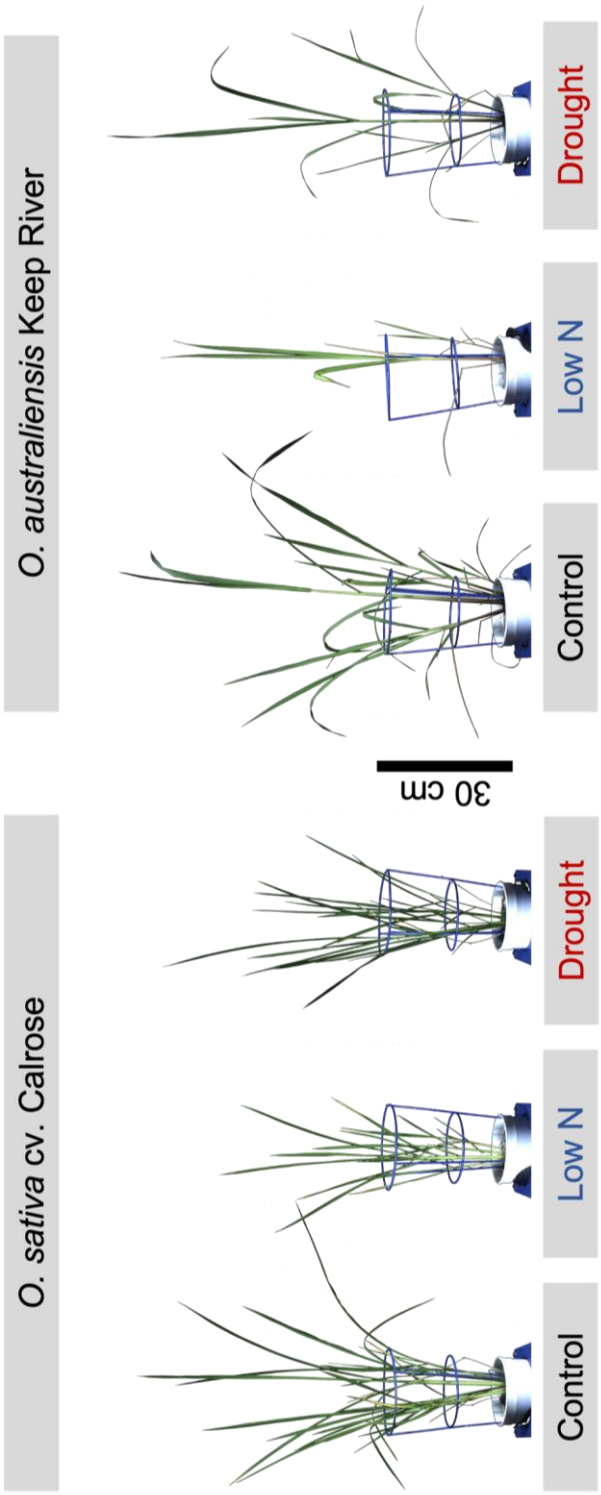
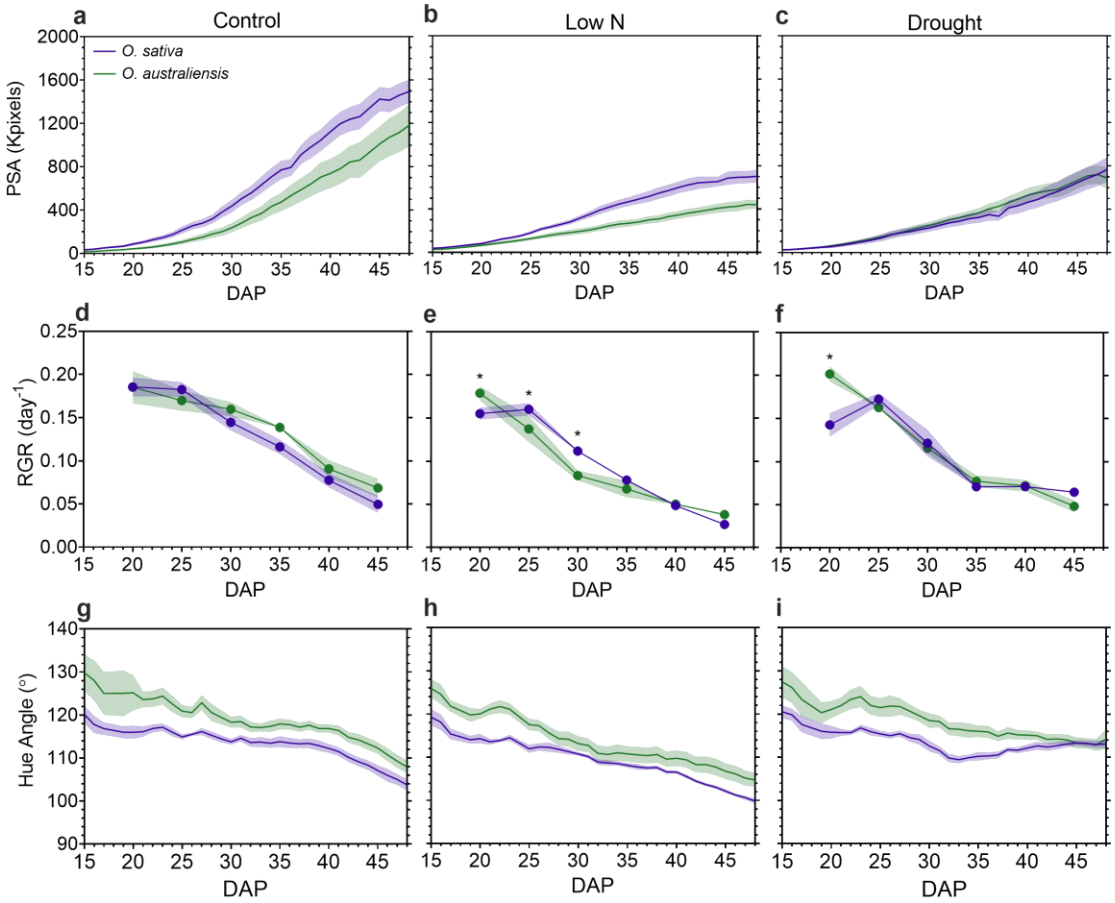


Figure 4.5. Mean (SE; n = 6) water use per unit shoot area (kpixel; a-d) of *O. sativa* and *O. australiensis* plants grown under different combinations of water and nitrogen availability for 33 days (15-47 DAP), (e) water use at 47 DAP (n = 4-6), and (f) water use efficiency at 47 DAP (n =6). Note that a-d display data from 30-47 DAP only. This is because droughted plants did not start receiving water until 27 DAP (for *O. sativa*) and 28 DAP for (*O. australiensis*). Note the different scales of a-d and e. Outliers were identified and removed using the Grubb's test. Letters denote means that differ significantly, as determined by Duncan's Multiple Range Test. The horizontal dashed lines were added to aid in comparisons between time course panels.

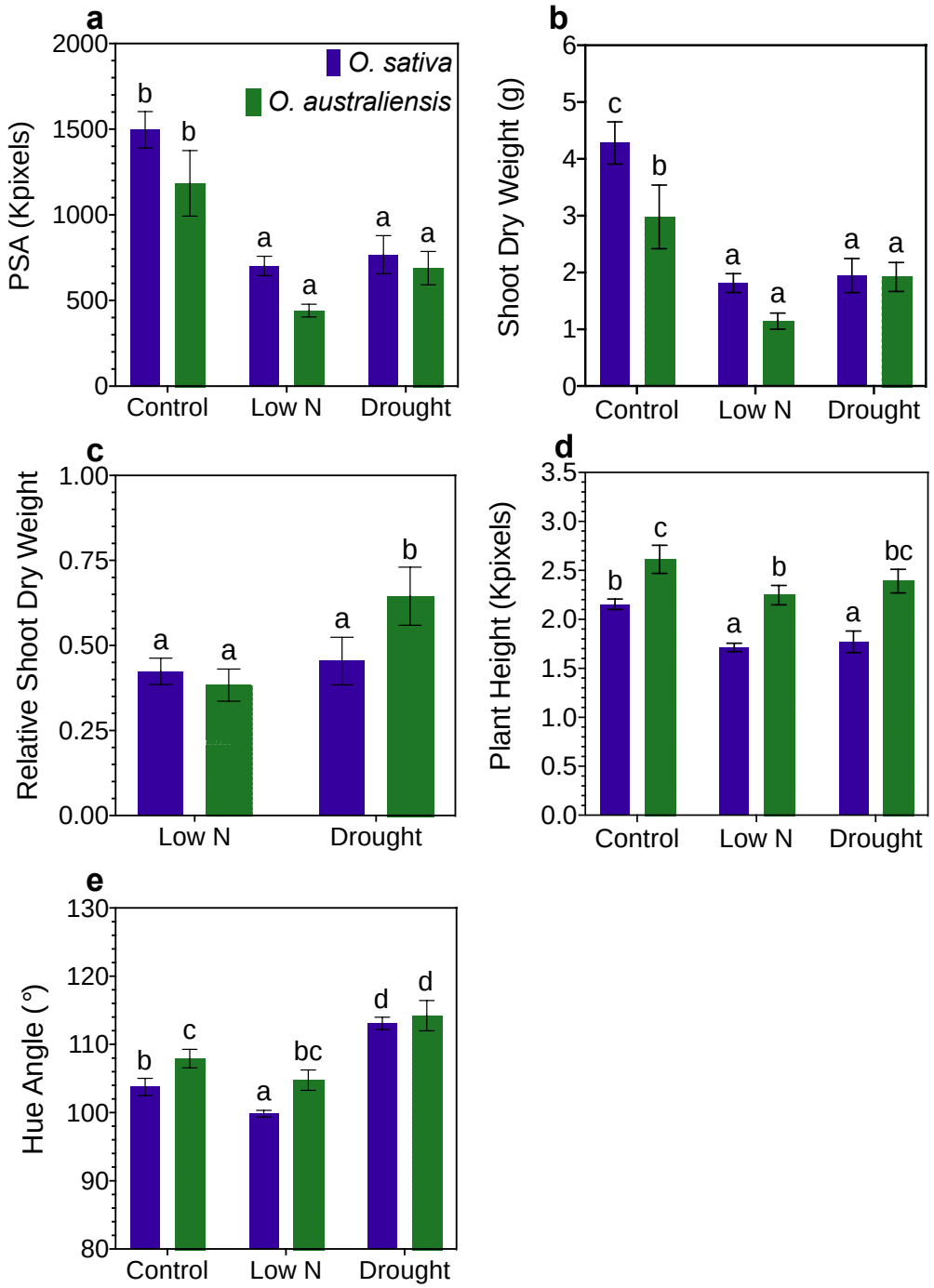
Supplementary Data



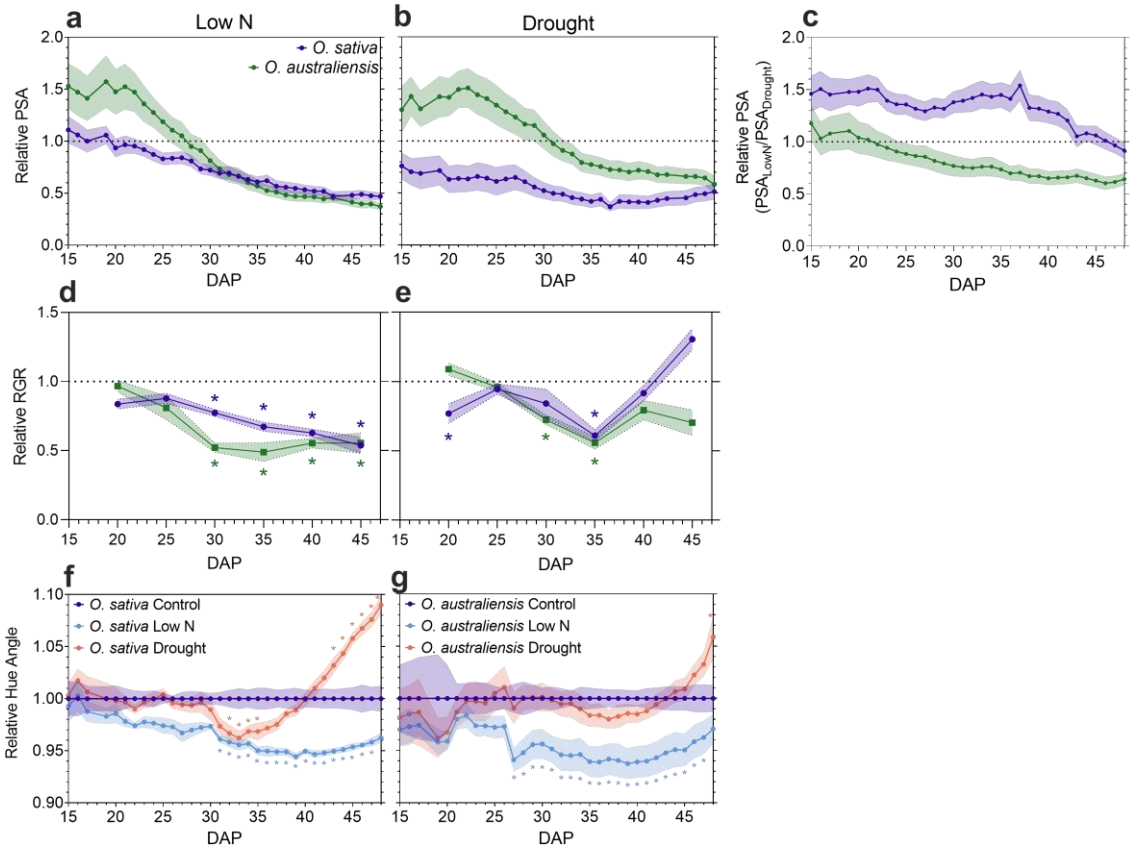
Supplementary Figure 4.6. Representative images of *O. sativa* and *O. australiensis* plants grown under individual low-N and drought stress at 48 DAP (34 days after imaging started) from the pilot phenotyping study. Images were obtained from The Plant Accelerator. Note that a different cultivar of *O. sativa* (Calrose, instead of Nipponbare) was used in the pilot study.



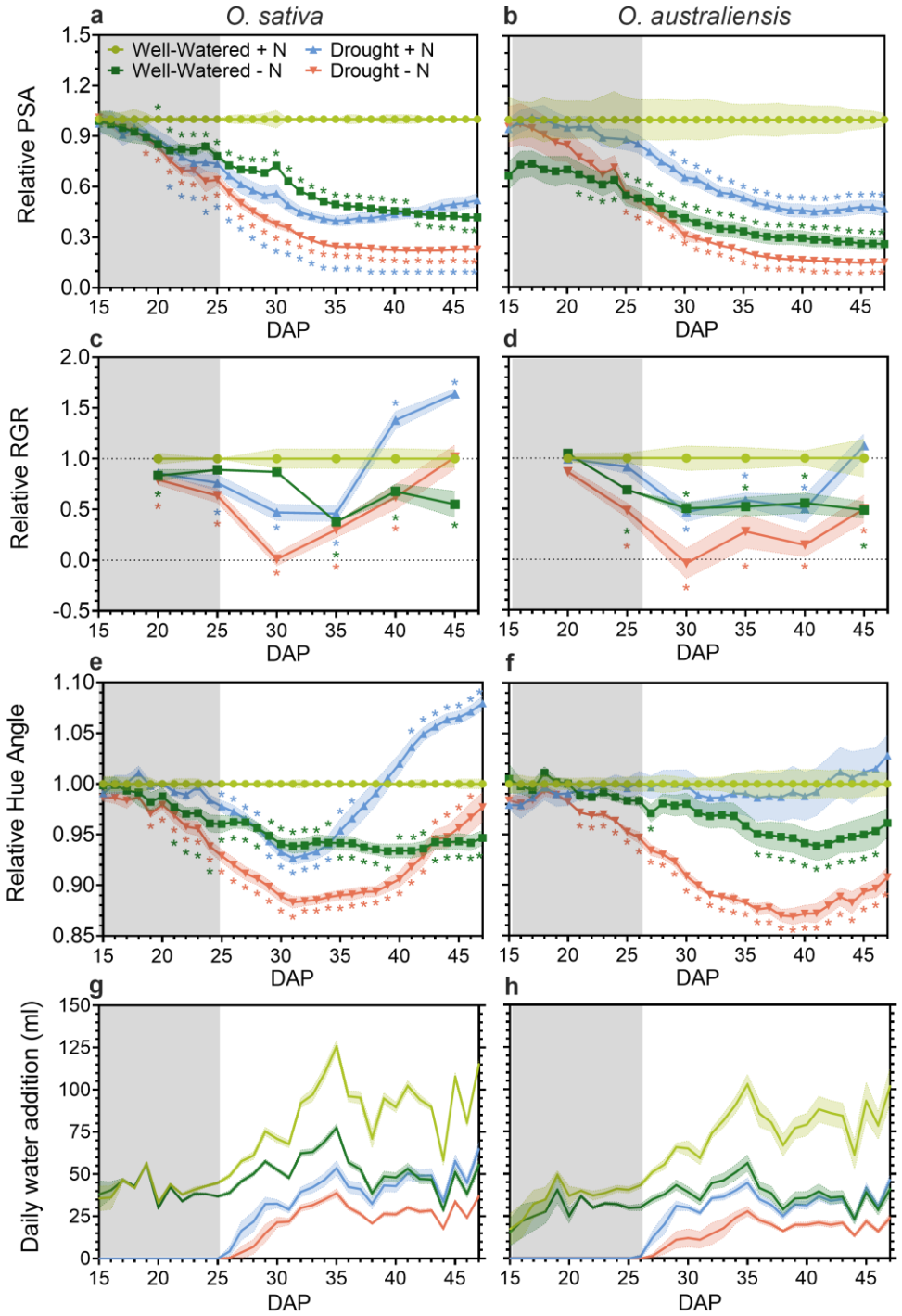
Supplementary Figure 4.7. Mean (SE; n = 6) projected shoot area (PSA; a-c), relative growth rate (RGR; d-f), and hue angle (greenness; g-i) of *O. sativa* and *O. australiensis* plants grown under individual low-N and drought stresses for 34 days (15-48 DAP) for the phenotyping pilot study.



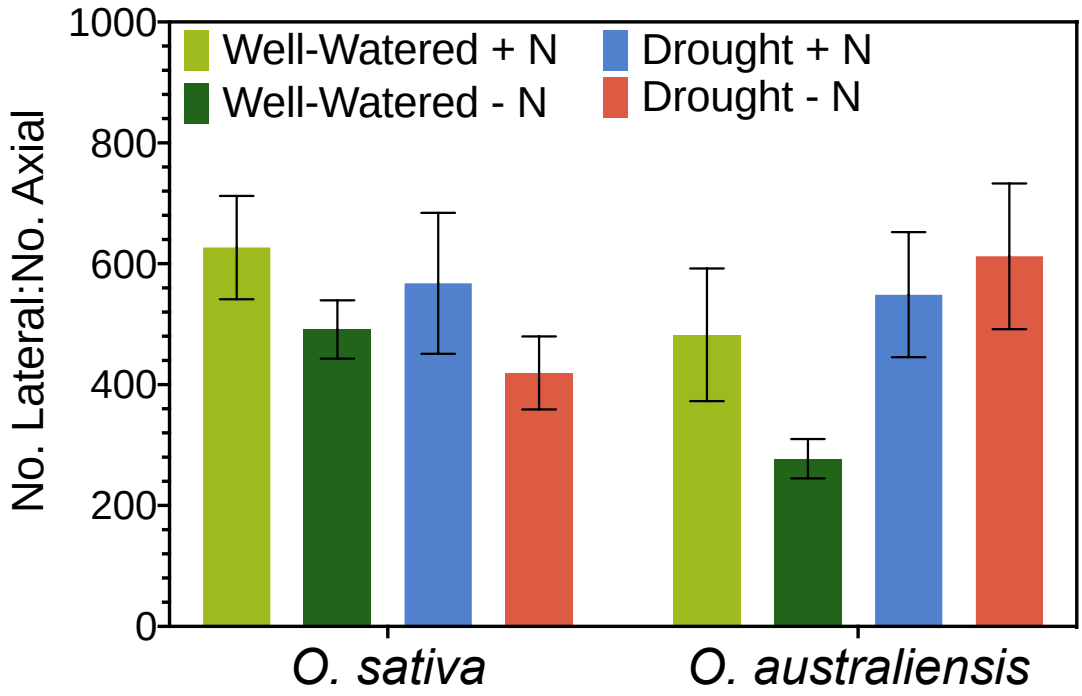
Supplementary Figure 4.8. Responses of *O. sativa* and *O. australiensis* plants to low-N or drought stress 48 DAP. (a) PSA, (b) shoot dry weight, (c) relative shoot dry weight, (d) plant height, (e) hue angle (greenness). Values are means SE (n = 6). Letters denote means that differ significantly, as determined by Duncan's Multiple Range Test.



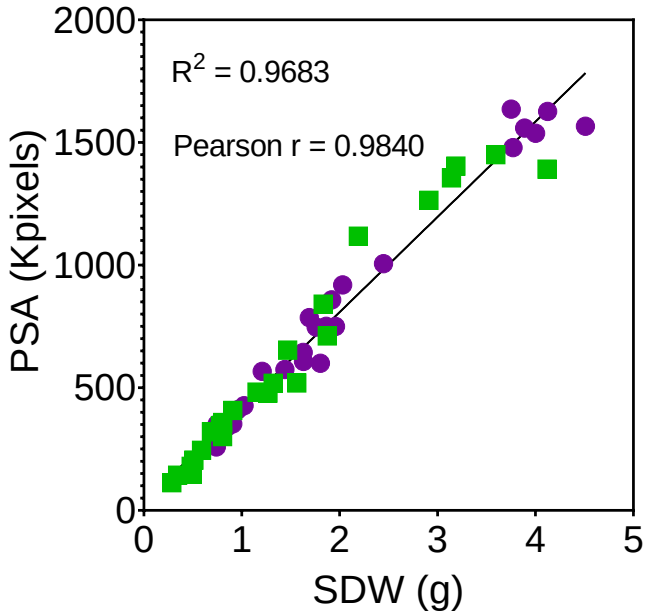
Supplementary Figure 4.9. Above-ground growth for *O. sativa* and *O. australiensis* plants grown on low-N or droughted soils expressed relative to the control group for each species. (a) and (b) show relative PSA of low and droughted plants, respectively. (c) shows the PSA of low-N grown plants relative to the PSA of droughted plants. (d) and (e) show the relative growth rates of low-N and droughted plants. (f) and (g) show the relative hue angle (greenness) of low-N and droughted plants. Values are means (SE; n = 6). In (a)-(e) the dotted horizontal line denotes the control line to which the treatment groups were compared. In (f) and (g) the variation around the control lines (purple) are shown. Asterix's denote means that are significantly different to the control.



Supplementary Figure 4.10. Above-ground growth for *O. sativa* and *O. australiensis* plants grown different combinations of soil N and water availability expressed relative to the control group for each species (a-f), and daily water addition over the course of the imaging period (g-h). (a) and (b) show relative PSA, (c) and (d) show relative RGR, and (g) and (h) show relative hue angle (greenness). Values are means ± SE (n = 6). The control groups are the horizontal lines at y = 1, and variation around the control is also shown. Asterix's denote means that are significantly different to the control. The grey shaded boxes indicate the drying down period for the droughted plants. Note that water was first administered to *O. sativa* by the automatic watering system at ~25 DAP, while *O. australiensis* received water from ~26 DAP.



Supplementary Figure 4.11. Mean (±SE; n = 6) ratio of lateral root number to axial root number in *O. sativa* and *O. australiensis* plants grown under individual and combined low soil N and drought conditions.



Supplementary Figure 4.12. Relationship between shoot dry weight (SDW) and projected shoot area (PSA) for *O. australiensis* (green) and *O. sativa* (purple) plants grown in this experiment. The results of the correlation analysis indicate that PSA is a good proxy for SDW.

Chapter 5

The first long-read nuclear genome assembly of *Oryza australiensis*, a wild rice from northern Australia

Wombo Prompt

“Grass nuclear genome assembly”

Wombo Theme

“Comic”

Statement of Authorship

Title of Paper	The first long-read nuclear genome assembly of *Oryza australiensis*, a wild rice from northern Australia
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Phillips, A.L., Ferguson, S., Watson-Haigh, N.S., Jones, A.W., Borevitz, J.O., Burton, R.A., and Atwell, B.J. The first long-read nuclear genome assembly of *Oryza australiensis*, a wild rice from northern Australia. Sci Rep 12, 10823 (2022). https://doi.org/10.1038/s41598-022-14893-5

Principal Author

Name of Principal Author (Candidate)	Aaron L. Phillips		
Contribution to the Paper	Conceived the study, carried out bench work, ran analyses, wrote the manuscript		
Overall percentage (%)	60%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	7th September 2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Scott Ferguson		
Contribution to the Paper	Assisted with data analysis and edited the manuscript		
Signature		Date	16th September 2022

Name of Co-Author	Nathan S. Watson-Haigh		
Contribution to the Paper	Conceived the study, assisted with data analysis, edited the manuscript		
Signature		Date	7th September 2022

Name of Co-Author	Ashley W. Jones		
Contribution to the Paper	Conceived the study, assisted with laboratory work, edited the manuscript		
Signature		Date	16th September 2022

Name of Co-Author	Justin O. Borevitz		
Contribution to the Paper	Conceived the study		
Signature		Date	16th September 2022

Name of Co-Author	Rachel A. Burton		
Contribution to the Paper	Conceived the study, edited the manuscript		
Signature		Date	12th October 2022

Name of Co-Author	Brian J. Atwell		
Contribution to the Paper	Conceived the study, funded the study, edited the manuscript		
Signature		Date	9th September 2022

Title:

The first long-read nuclear genome assembly of *Oryza australiensis*, a wild rice from northern Australia

Author List:

Aaron L. Phillips[1,2], Scott Ferguson[3,4], Nathan S. Watson-Haigh[5,6], Ashley W. Jones[3,4], Justin O. Borovitz[3,4], Rachel A. Burton[1,2], and Brian J. Atwell[7*]

***Corresponding Author:**

Professor Brian J. Atwell

Institutional Addresses:

1. Department of Food Science, University of Adelaide, Adelaide, South Australia, Australia
2. ARC Centre of Excellence in Plant Energy Biology, Adelaide, South Australia
3. Research School of Biology, Australian National University, Canberra, Australian Capital Territory, Australia
4. ARC Centre of Excellence in Plant Energy Biology, Canberra, Australian Capital Territory
5. South Australian Genomics Centre, University of Adelaide, Adelaide, South Australia, Australia
6. Australian Genome Research Facility, Victorian Comprehensive Cancer Centre, Melbourne, Victoria, Australia
7. School of Natural Sciences, Macquarie University, Sydney, New South Wales, Australia

Author Emails:

aaron.phillips@adelaide.edu.au

scott.ferguson@anu.edu.au

nathan.watson-haigh@sahmri.com

ashley.jones@anu.edu.au

justin.borevitz@anu.edu.au

rachel.burton@adelaide.edu.au

brian.atwell@mq.edu.au

Abstract:

Oryza australiensis is a wild rice native to monsoonal northern Australia. The International *Oryza* Map Alignment Project emphasises its significance as the sole representative of the EE genome clade. Assembly of the *O. australiensis* genome has previously been challenging due to its high Long Terminal Repeat (LTR) retrotransposon (RT) content. Oxford Nanopore long reads were combined with Illumina short reads to generate a high-quality ~858 Mbp genome assembly within 850 contigs with 46x long read coverage. Reference-guided scaffolding increased genome contiguity, placing 88.2% of contigs into 12 pseudomolecules. After alignment to the *Oryza sativa* cv. Nipponbare genome, we observed several structural variations. PacBio Iso-Seq data were generated for five distinct tissues to improve the functional annotation of 34,587 protein-coding genes and 42,329 transcripts. We also report SNV numbers for three additional *O. australiensis* genotypes based on Illumina re-sequencing. Although genetic similarity reflected geographical separation, the density of SNVs also correlated with our previous report on variations in salinity tolerance. This genome re-confirms the genetic remoteness of the *O. australiensis* lineage within the *O. officinalis* genome complex. Assembly of a high-quality genome for *O. australiensis* provides an important resource for the discovery of critical genes involved in development and stress tolerance.

Keywords:

Rice, *Oryza*, Genome, Nanopore, PacBio, Long Terminal Repeat, Stress Tolerance

Background:

The genus *Oryza* is made up of ~27 species that have been classified into 11 genome groups[1,2]. Two of these species (*O. sativa* and *Oryza glaberrima*), both of

which belong to the AA genome clade, were independently domesticated in Asia and Africa, respectively. Together, these two domestic rice species serve as a staple food source for about 60% of the global human population. The remaining ~25 species occur in wild populations and have become specialised to occupy a diverse set of environments around the world[3].

To date, the species and cultivars that make up the AA genome clade have provided most of the genetic variation and resources that have led to the success of domesticated rice[4], while other genome clades have contributed relatively little to domestication. Encouragingly, rich diversity is apparent, even within landraces of just one *O. sativa* subspecies. For example, in a GWAS study for 14 agronomic traits (including tolerance to drought and degree of seed shattering) in 517 *indica* sub-species landraces, Huang et al.[5] revealed ~3.6 million SNPs. Currently nine of the 11 *Oryza* genome assemblies available on EnsemblPlants (accessed 26/2/2021)[6] belong to the AA genome clade, one belongs to the BB genome (*Oryza punctata*), and one to the FF genome (*Oryza brachyantha*). Therefore, wild species of rice occupy a number of genomic clades that are not yet represented by an assembly and undoubtedly include traits that would be desirable in domestic cultivars. The wild species therefore represent vast reservoirs of untapped genetic variation that could be harnessed for domestic rice improvement[2–4,7]. Recently, Hiromi et al.[8] provided an update to the OryzaGenome database which includes short-read sequencing data for several non-AA genome *Oryza* species. The availability of *Oryza* '-omics' resources is increasing, yet the number of whole-genome assemblies for non-AA genome wild species remains low. Whole-genome assemblies of wild *Oryza* species will provide new sources of genes and gene variants/haplotypes, enhancing the ability of breeding programs to introduce desirable traits into domestic cultivars. However care is necessary to navigate issues such as linkage drag and suppressed recombination

frequencies that may be associated with use of wild germplasm. Additionally, such projects could lead to the domestication of wild rice populations (e.g., through the targeted modification of domestication-associated traits, such as seed shattering)[9]. Overall, harnessing the genetics of wild *Oryza* species has wide application for cereal crop improvement and specifically for crop resilience. In fact, the *Oryza* Map Alignment Project (OMAP) and the International OMAP (IOMAP) recognised the value that the genome assemblies of wild *Oryza* species would provide to research in abiotic and biotic stress tolerance, and domestication, as long ago as 2003[10–12]. These projects recommended sequencing the genomes of several wild rice species but progress in this respect has been slow.

Recently, several non-AA genetic resources have become available, including genomes within the *O. officinalis* complex[13] and the genome of *Oryza granulata*[14]. *O. granulata,* an upland wild rice species, has a demonstrated ability to tolerate drought stress and bacterial blight, providing a valuable new source of *Oryza* genetic variation. However, owing to its large genome size, high repeat content, and being a short-read assembly, the *O. granulata* genome assembly is highly fragmented, reducing its utility. These same difficulties (i.e., genome size and repeat content) have been encountered during assembly of the EE genome, represented solely by *O. australiensis,* which is reported to be the largest diploid *Oryza* genome.

Oryza australiensis is a perennial wild relative of rice endemic to the tropical regions of northern Australia. Due to the extremely hot and sporadically dry environment in which *O. australiensis* exists, this species has been described as an extremophile[3]. *Oryza australiensis* is the sole member of the EE genome clade and its genome has been estimated to be 965 Mbp in size, more than double that of *O. sativa* cv. *Japonica*[15,16]. The size of the *O. australiensis* genome is the result of the expansion of long terminal repeat (LTR) retrotransposon (RT) families (e.g., *Gypsy*

and *Copia*), which comprise an estimated 65% of the genome. In comparison, the genome of *O*. *sativa* is composed of approximately 10% LTR-RT[15]. It is uncertain what led to the rapid genome-wide expansion of these LTR-RT families. However, it has been noted that the expansion of LTR-RTs facilitated rapid adaptive genome evolution in the *Oryza* genus, leading to speciation events[17]. These same processes may also play a role in the evolution of biotic and abiotic stress tolerances. Heat tolerance has been well-characterised in *O. australiensis*[18], and transformation of domestic rice with a key gene (Rubisco activase) involved in heat tolerance fortifies grain yield during episodes of heat[19]. Further, Yichie et al.[20] identified population-specific tolerance to salinity stress in *O. australiensis* accessions, and Hamzelou et al.[21] report on drought tolerance in this species. The unique anatomy and photosynthetic efficiency of *O. australiensis* leaves also make it an important species for understanding how rice photosynthesis can be augmented[22]. In terms of food quality, the grain of *O. australiensis* has unique pigmentation and starch composition, distinct from Asian species[23,24], that could prove to be a profitable food item in and of itself.

A highly-quality reference genome for *O. australiensis* could reveal novel loci for abiotic stress tolerance and help shed light on the genome evolution of the *Oryza* genus. However, since the conception of OMAP in 2003[11], and despite international sequencing efforts since[12], no research group has successfully assembled a high-quality, chromosome-level genome for *O. australiensis.* A previous attempt to assemble the *O. australiensis* genome with short reads resulted in a highly fragmented assembly due to the genome size and high repeat content[25]. Efforts such as those by Joly-Lopez et al.[25] highlight the difficulties involved in assembling highly contiguous plant genomes with significant repeat content[26].

In this study, we report on the assembly of a reference-quality *O. australiensis* nuclear genome from Keep River (KR; Northern Territory, Australia) into 1,956 contigs using Oxford Nanopore Technology (ONT) long-reads. The contigs were subsequently scaffolded into 12 pseudomolecules representing 12 chromosomes using the *O. sativa* Nipponbare genome assembly as a guide.

Methods:

Growth Conditions for Plant Material used for Genome Sequencing

Seed of four *O. australiensis* populations was collected from the monsoonal savannah region of northern Australia (Fig. 5.1), compliant with the relevant national and international guidelines and legislation for wild species collections. All seed was obtained from the Australian Tropical Crops and Forages Collection (AusTRCF) of the Australian Grains Genebank (previously the Australian Plant Genetic Resource Information System) and as part of the Australian Tropical Crops Genetic Resource Centre Collection (ATCGRC). The Keep River accession used for the reference genome assembly bears the accession number AusTRCF 318143 and was collected by Dr I. Cowie at Darwin Herbarium, Northern Territory, Australia (Voucher ID# 9583) on 2nd May 2002 from a drying creek bed in the Keep River National Park after the passing of the summer monsoon. The other three accessions used for re-sequencing (CH, D and VR) are held in the same collection with accession numbers of AusTRCF 300134, 300137 and 300131. These lines are vouchered in the Jeff Corfield Collection in Townsville (Queensland, Australia) under the accession numbers JC 2317, JC 2336 and JC 2312, respectively. CH, D and VR were collected from the sites denoted in Fig. 5.1 between 1995 and 1997 by Mr I. Watson. Seed was stored at 4 °C until used. Seeds were imbibed in water for one hour and then surface sterilised in a 50% sodium hypochlorite bleach solution for 30 minutes followed by

rinsing. Sterile seeds were sown onto absorbent cotton-lined petri dishes and incubated at 34 ^{o}C for two days to trigger germination. Germinated seeds were sown into a fine textured krasnozem soil (sourced from Robertson, NSW, Australia) and covered in a thin layer of vermiculite. Plants were grown in glasshouses at Macquarie University (Sydney, Australia) under a 30 ^{o}C/22 ^{o}C day/night cycle in April-June 2019.

DNA Extraction

Roughly 1 g of healthy, fully expanded leaf material was collected from ~2-month-old *O. australiensis* plants (one plant per genotype) and ground in a mortar and pestle under liquid nitrogen until a fine powder was achieved. High molecular-weight DNA was extracted at the Australian National University (ANU; Canberra, Australia) following the method of Jones et al.[27].

Nanopore Library Preparation, Sequencing, and Base Calling

The sequencing library was prepared as per the manufacturer's (Oxford Nanopore Technologies, Oxford, United Kingdom) instructions for genomic DNA by ligation method (SQK-LSK109) with small modifications. The manufacturer recommends using 0.2 pmol of DNA, however this is optimised for DNA fragments 8 kbp in length. As we were working with substantially longer (~20 kbp) DNA fragments we loaded 3 µg (0.23 pmol) of DNA.

The library was run on FLO-MIN106 (R9.4.1 revD) flowcells as per the manufacturer's instructions. To obtain a good level of coverage of the *O. australiensis KR* genome, the same library preparation for this sample was loaded onto three separate flow cells (i.e., the genomic DNA from one plant was sequenced

on three flowcells). The raw signal intensities obtained from the flow cells were used by *guppy* (version 3.0.3) from ONT to perform base calling.

Illumina Sequencing

Aliquots (containing 2.24 ng DNA) of the same DNA samples from the four *O. australiensis* genotypes described above were prepared for Illumina short read sequencing. Short-read libraries were created using the Illumina Nextera tagment DNA enzyme (TDE1) according to Jones *et al.*[28]. Libraries were size selected for 350-600 bp fragments using a PippinHT (Sage Science). Sequencing was performed on an Illumina NovaSeq 6000 S4 flow cell 300 cycles (150 bp paired end), being multiplexed with other projects. Sequencing was performed at the Biomolecular Resource Facility, ANU.

PacBio Iso-Seq Library Preparation and Sequencing

Young grains were harvested during the milky period of development from mature *O. australiensis KR* plants from the same collection as the individual used for ONT and Illumina sequencing (i.e., different plants from the same population; germinated as above and grown in a glasshouse at 30 ^{o}C with ambient light at The Plant Accelerator, University of Adelaide, Adelaide, South Australia from February to June 2020). Some developing grain still possessed their pedicel/rachilla and pollen-containing anthers and these were included in the samples to increase the diversity of transcripts in the RNA library.

The growth zones of ten 115-day-old *O. australiensis KR* plants were excised and used for RNA extraction. The growth zone is defined here as ~10 mm from the base

of the youngest leaves emanating from the central cylinder of the plants. It was necessary to remove surrounding leaf sheathes to access the central cylinder.

A separate cohort of *O. australiensis KR* plants were germinated as above, except germination was achieved with an incubation temperature of 30 ^{o}C. Germinated seeds were transferred to a growth rack partially submerged in a ½ Hoaglands solution (pH 5.01) spiked with 25 mg/L ferrous sulphate. After seven days of growth, root tips (5 mm from the end of the root), older seminal roots with young lateral roots, and healthy fully expanded leaves were harvested, snap frozen in liquid nitrogen and stored at -80 ^{o}C. The growth solution was changed once during this time.

An additional cohort of *O. australiensis KR* seed was germinated at 30 ^{o}C and submerged in water in the dark to promote coleoptile growth. After six days under these conditions, etiolated coleoptiles were harvested as above.

Total RNA was extracted from the six tissues (7-day-old leaves, 7-day-old root tips, 7-day-old mature root sections, 6-day-old hypoxic coleoptiles, reproductive tissue, and 118-day-old growth zones) with the Sigma Spectrum Plant Total RNA Kit according to the manufacturer's instructions, with minor modifications for the developing grain sample as per Betts *et al*.[29]. RNA samples were dried in RNA Stable tubes as per the manufacturer's instructions and shipped to DNA Link (Seoul, South Korea). RNA quality and concentration of the samples were checked with a Nanodrop and with a Bioanalyzer 2100 instrument (Agilent, Santa Clara, USA). Only low-quality RNA (RIN < 7) could be extracted from the older seminal roots so this sample was omitted from further analyses.

RNA samples were barcoded upon conversion to cDNA libraries (to facilitate multiplexing) following the manufacturer's instructions of the SMARTer PCR cDNA Synthesis kit. cDNA libraries were amplified by PCR and then an equal concentration of cDNA from each was pooled into a single library. The library was prepared for SMRTbell sequencing with the Express Template Prep 2.0 kit. Iso-Seq sequencing was carried out on a Sequel II PacBio instrument using a single SMRT cell.

Iso-Seq reads were processed through the IsoSeq v3 pipeline to generate high quality polished isoforms. Briefly, raw data is converted from multiple subreads per ZMW into a single consensus CCS read (minimum read quality of 0.9), demultiplexed and clustered (minimum predicted accuracy of 0.99). The aim of the clustering step is to collapse the CCS reads derived from different isoforms of the same gene or from the same transcript isoform but which differ slightly in their 5' and 3' ends, typically as a result of RNA degradation.

Nuclear Genome Assembly and Assessment

A summary of all reads obtained during this study can be found in Table 5.1. *Jellyfish* (RRID:SCR_005491)[30] and *genomescope2* (RRID:SCR_017014)[31,32] were used to estimate the size, unique sequence content, and heterozygosity of the *O. australiensis* KR genome using 17-mers in the KR Illumina short read libraries. Raw MinION reads for *O. australiensis KR* were corrected and then *de novo* assembled using *Canu* v1.9 (RRID:SCR_015880)[33] as follows: canu -p oryza -d canu genomeSize=965m -nanopore-raw all.fastq.gz -gridOptions="--time=72:00:00" -obtovlThreads=24 -batMemory=9. The MinION long reads and the Illumina short reads were mapped to the resulting contigs using *minimap2* v2.17 (RRID:SCR_018550)[34]. Alignment files were used to polish the genome assembly

with *HyPo* v1.0.3[35]. *JVarkit* was used to visualise long-read coverage of the contigs[36]. Contigs were checked for microbial contamination by querying BLAST-generated taxon ID numbers with the lineage command of *taxonkit* v0.6.0[37]. Contigs found to contain non-eukaryotic sequences were interrogated using the online genome browser IGV-Web (RRID:SCR_011793) by mapping long reads back to the assembly and checking the coverage and overlap of long reads with non-suspect sequences[38]. *Purge Haplotigs* (v1.1.2) was used to determine whether the wild rice genome assembly contained any haplotigs or junk sequences[39].

The quality of the genome assembly was assessed by analysing the quality of the LTR-RT elements of the *O. australiensis KR* genome. LTR-RT elements are difficult to assemble. As such, LTR-RT assemblies provide a proxy for the quality of the rest of the genome. LTR regions were identified with *LTR_Finder_Parallel* v1.1 (RRID:SCR_018969)[40] and *LTR_Harvest*[41]. The outputs were used as input for LTR_Retriever v2.8.7 (RRID:SCR_017623) – to generate LTR Assembly Index (LAI) scores[42]. The quality of the assembly was also assessed with *BUSCO* v4 (RRID:SCR_015008)[43] using the *Poales* database (poales_odb10.2019-11-20; contains 4,896 sequences) as the lineage. Contigs were mapped to the *O. sativa* Nipponbare reference genome (GCA_001433935.1) with *minimap2* and visualised with *MashMap* v2.0[44].

Pseudomolecules for the assembly were generated using *RagTag* v1.0.0[45] using release 7 of the *O. sativa cv.* Nipponbare genome[46] as a guide for the ordering and orientation of contigs in to pseudomolecules (Supplementary File S1a). Contigs that could not be assigned to a pseudomolecule were concatenated (with 100 Ns at contig boundaries) into a ‘chromosome’ called ChrUn, though we used the

unassigned contigs themselves for most of the downstream analyses. To understand why the unassigned contigs could not be placed into a pseudomolecule, the identity of a random subset of the unassigned contigs was determined by BLAST searches and the average LTR content of ChrUn was determined by *LTR_Retriever*. Contigs flagged as repeats, haplotigs or junk by *Purge Haplotigs* were kept in the assembly for completeness but were not included in scaffolding (see above). *JVarkit's* *WGSplotter* was used to visualise long-read genome coverage[36]. BLAST was used to determine the identity of a random subset of the unassigned contigs. Subsequent annotation of the unassigned contigs (see below) also helped in determining the identity of the sequences that could not be assembled into pseudomolecules. The quality of the scaffolded assembly was assessed again with *BUSCO* and LAI scores, as above. *LTR_Retriever* was also used to produce estimates for LTR-RT insertion times. The quality of the assembly was further checked by mapping ONT long reads to the pseudomolecules to obtain a read mapping rate and to estimate read coverage across the assembly.

Comparison of O. australiensis *KR Assembly to the* O. sativa *Nipponbare Assembly*

The *EDTA*-masked (see below) pseudomolecules were aligned to the Nipponbare genome using *minimap2* with the 'asm20' setting. The mapping file was used as input for *dotPlotly*[47]to generate a dot plot of the whole-genome alignment with the following settings: "-m 2000 -q 500000 -k 12 -l -p 12". The resulting plot was used to interrogate suspected structural variations (SVs) in the wild rice genome. We used *O. sativa* vs *O. punctata* Chr2 (CM002489.2; a BB genome rice), *O. sativa* vs *O. australiensis* KR Chr2, and *O. australiensis* KR vs *O. punctata* Chr2 alignments to inspect a region on Chr2 in these species which has been reported as an inversion in

non-AA rice genome[1]. Alignments were made as described above, except "-m 5000" was used for *dotPlotly* visualisation. To check whether reference bias in the *O. australiensis* scaffolds was affecting the Chr2 SVs, we scaffolded the KEEP contigs using the *O. punctata* reference genome (AVCL0200000). We aligned the *O. punctata*-guided *O. australiensis* Chr2 to *O. punctata* Chr2, and then to *O. sativa* Nipponbare Chr2 as above.

To check whether the SVs were real, contig coordinates within the scaffolds were extracted from the *RagTag* AGP file and these were used to build a 'cytoband' file for *karyoploteR*[48]. The coordinates for the SVs (± 10 kbp) identified in the *O. australiensis KR* and *O. sativa* genome alignment were added to the 'cytoband' file, and these were used to visualise the boundaries between contigs and SV events using *karyoploteR*. Further, long reads were mapped to the *O. australiensis* KR scaffolds to check the level of coverage across SV junctions. SV junctions with low coverage were coloured red to identify them as suspicious.

To identify single nucleotide variations (SNVs) in the wild rice genome, *O. australiensis KR* Illumina short reads were mapped to the Nipponbare genome with *minimap2*. The resulting BAM file was used as input for the *SAMtools* (RRID:SCR_002105)[49] sub-command *mpileup,* and *BCFtools* (RRID:SCR_005227) called SNV variants. *VCFtools* (RRID:SCR_001235)[50] filtered low quality SNVs with low-read support (minQ = 30; minDP = 3). High quality SNVs were visualised using *karyoploteR* in RStudio with a window size of 70 kbp.

Genome Repeat Annotation

LTR_Retriever (see above) was used to provide an estimate of LTR-RT elements in the *O. australiensis* KR contigs and scaffolds. Repetitive DNA elements were annotated using the Extensive *de novo* TE Annotator (*EDTA;* Supplementary File S2b)[51]. A manually curated rice repeat-element library provided by the developers of *EDTA* was used to identify repeats in the *O. australiensis KR* genome. A library containing *O. sativa* Japonica transcript sequences including untranslated regions (UTRs) was downloaded from the IOMAP data repository[52] and used by *EDTA* to mask endogenous genes from its annotation pipelines. *EDTA* also identifies repeat sequences in the input genome that were not part of the repeat library used for interrogation and then constructs a repeat library specific to the input assembly. Total repeat content was also estimated and annotated using the *RepeatMasker*-based *Repeat Masking* tool included in OmicsBox version 2.0.10[53,54]. The custom *O. australiensis KR*-specific repeat library generated by *EDTA* was used as input for the RMBlast search engine, while other settings were left as default.

Protein-coding Gene Annotation

Evidence-based gene-finding was carried out with the OmicsBox Gene Finding tool with *O. sativa* set as the most closely related species[53,54]. A protein evidence file containing protein sequences from 12 species/accessions of rice downloaded from the IOMAP data repository[52] and the Iso-Seq isoforms were used as evidence for gene finding. Genes were searched on both strands of the pseudomolecules and the unassigned contigs and alternative splice variants were allowed. The resulting gene coordinates were used as input for the built-in *Functional Gene Annotation* workflow. This workflow uses RepeatMasker to mask repeat elements in the genome assembly before carrying out gene prediction with Augustus. Then, both Blast2GO and InterProScan pipelines were applied to search for gene and protein matches for the

identified genes. Annotations were further supplemented by searching orthologues of the identified genes/proteins using EggNOG-mapper in OmicsBox. Annotations for the contigs and scaffolds were exported as GFF3 files (Supplementary Files S3c and S4d). To check whether the GFF3 files contained TE gene annotations, we used a list of known plant TE families from the Atlas of Plant Transposable Elements[55].

O. australiensis *Genotype Diversity and Genetic Similarity*

To identify SNVs between four genotypes of *O. australiensis* with demonstrated differences in salt tolerance[20], the *O. australiensis* CH, D, and VR Illumina short reads (Table 5.1) were mapped to the scaffolded *O. australiensis KR* genome assembly using *minimap2* and visualised as above. Keep River short reads were also mapped back against the Keep River genome to provide a baseline of SNVs. *Mpileup* was used to call variants between the *O. australiensis* KR genome assembly and the mapped reads. VCF files were filtered using bcftools such that only reads with a mapping quality >30, and only SNVs with a depth of 5× were considered. The patterns of SNV density described here were visually inspected.

Genetic similarity between the *O. australiensis* genotypes was determined using the Illumina short reads for each genotype as input for *kWIP*[56]. *kWIP* uses a Weighted Inner Product of different k-mer hash values derived from *khmer* (RRID:SCR_001156)[57] to reduce the effect that erroneous read data has on estimates of genetic similarity. The estimate of genetic similarity was also extended to the *Oryza* genus using publicly available Illumina short read data derived from genomic DNA samples (see Supplementary File S5e for the list of accessions).

Results:

Genome Assembly, Making Pseudomolecules and Genome Quality

To generate long-read sequence data for assembly, DNA from a single *O. australiensis* Keep River (KR) plant was extracted, purified, and sequenced on three ONT MinIOn flow cells (KR1, KR2, and KR3). Additionally, the long-read DNA extract was used to generate Illumina NovaSeq short read data for *O. australiensis* KR. DNA extraction and Illumina short-read sequencing was also performed on *O. australiensis* genotypes CH, D, and VR (Fig. 5.1; Table 5.1). We generated 49 Gbp (2.01 million reads) of ONT MinION long read data and 19 Gb (138.1 million reads) of Illumina short read data for the *O. australiensis* Keep River accession. Further, we generated 3.14 Gbp (1.25 million reads) of PacBio Iso-Seq data from five distinct tissue types (Table 5.1; Supplementary Figure 5.4). Finally, we generated 20 Gbp, 12 Gbp, and 19 Gbp of Illumina short read data for three additional *O. australiensis* genotypes (CH, D, VR, respectively; Table 5.1).

K-mer analysis of *O. australiensis* KR short reads by *jellyfish* and *genomescope2* estimated the genome size to be ~822 Mbp. About 41.9% of the k-mers were unique, which suggested that the genome contained 58.1% repeat content. Heterozygosity was estimated to be 0.001%. Based on these results, we did not expect heterozygosity to have greatly affected contig assembly.

Canu assembled a genome containing 980 Mbp across 1956 contigs (Table 5.2). This represents 101.5% of the expected *O. australiensis* genome size. The contig N50 was 1.9 Mbp, the largest contig was 24.9 Mbp in length, while the smallest contig was 1324 bp in length. After polishing with *HyPo* the size of the assembly increased to 996 Mbp (103.2% of the expected genome size). We achieved an average coverage of 38× and the contigs aligned well with the *O. sativa* Nipponbare

reference genome (Supplementary Figure 5.5 & 5.6). *BUSCO* and LAI scores of 91.9% (complete) and 15.2, respectively, indicate that we have assembled a highly complete contigs. We identified 93 haplotigs, 124 junk contigs, 889 repeat contigs, and 850 primary contigs using *Purge Haplotigs* (Table 5.2). Thus, after purging the assembly of all but the primary contigs, the final assembly size was 859 Mbp (*cv*. 965 Mbp estimated by flow cytometry[15], and 822 Mbp estimated by k-mer counting – see above).

Table 5.1. Summary of reads used for each step of genome assembly, polishing, variant detection and annotation. See Figure 1 for details on the accessions. KR 1, KR 2, and KR 3 are the reads obtained from a single genomic DNA preparation sequenced on three different flowcells. These reads were derived from the same *O. australiensis KR* plant and were used for the assembly of the reference *O. australiensis KR* genome. The same DNA preparation was used for the KR Illumina NovaSeq library preparation. CH, D, and VR (accession numbers appear in parentheses) refer to different accessions of *O. australiensis* that have been shown to vary in their tolerance to salt[20]. Reads from these accessions were used to estimate genetic similarity between the genotypes. Multiple *O. australiensis* KR plants were used for RNA extraction for Iso-Seq analysis.

Purpose	Sequencing Platform	Sample	No. Reads (Millions)	Mean Read Length (kbp)	Read N50 (kbp)	No. Bases (Gbp)	Coverage (x)
Genome Assembly and Polishing	MinION FLO-MIN106 R9.4.1 revD	*O. australiensis* - KR 1	0.45	26.24	40.31	12	12
		O. australiensis - KR 2	0.94	23.22	35.14	22	23
		O. australiensis - KR 3	0.62	24.95	37.38	15	16
Genome Polishing	Illumina NovaSeq	*O. australiensis* - KR	138.81	NA	NA	19	21
Genetic similarity	Illumina NovaSeq	*O. australiensis* - CH (300134)	143.27	NA	NA	20	22
		O. australiensis - D (300137)	87.27	NA	NA	12	13
		O. australiensis - VR (300131)	139.67	NA	NA	19	21
Genome Annotation	PacBio Sequel II	*O. australiensis* - KR Leaf	0.24	2.69	NA	0.65	NA
		O. australiensis - KR Coleoptile	0.18	2.29	NA	0.42	NA
		O. australiensis - KR Root Tip	0.3	2.16	NA	0.64	NA
		O. australiensis - KR Growth Zone	0.3	2.75	NA	0.84	NA
		O. australiensis - KR Reproductive Tissue	0.24	2.46	NA	0.59	NA

Table 5.2. Genome assembly statistics for *O. australiensis* Keep River.

Stage	Feature (unit)	Value
Long Reads	Post-QC Bases (Gbp)	49.2
	Estimated Coverage (x)	51

	Expected Genome Size (Mbp)	965
Genome Assembly (contigs)	Assembled Genome Size (Mbp)	996
	Long Read Coverage (x)	38
	No. Contigs	1,956
	Contig N50 (Mbp)	1.9
	Contig L50	114
	Contig N90 (kbp)	186.7
	Contig L90	799
	BUSCO Score (100% = 4896)	91.9
	Whole-Genome LAI	15.2
Genome Scaffolds	Scaffolded Genome Size (Mbp)	860.9
	Long Read Coverage (x)	46
	Placed Contigs	693
	Unplaced Contigs	157
	Total Length of Placed Contigs (including Ns; Mbp)	812
	Total Length of Unplaced Contigs (Mbp)	46.9
	Gaps (Mbp)	2.1
	BUSCO Score (100% = 4896)	97.5
	Whole-Genome LAI Score	17.6
LTR_Retriever (scaffolds)	LTR-RT Content (Mbp)	518.2
	LTR-RT Content (%)	60.2

Scaffolding resulted in the placement of 693 of the 850 primary contigs into one of 12 pseudomolecules and increased assembly size to ~861 Mbp. Pseudomolecules contained 81.5% of all assembled contigs (~812 Mbp of 996 Mbp) and long read coverage increased to 46x (Table 5.2; Supplementary Figure 5.7; Supplementary File S1a). The remaining 157 contigs (46.9 Mbp) could not be placed into a pseudomolecule. Analysis of the unassigned contigs by *LTR_Retriever* showed that these sequences were largely made up of repeat sequences (~80% LTR-RT content, versus ~60% in the pseudomolecules), explaining why they could not be placed in the scaffolds. After scaffolding with KEEP contigs only, the *BUSCO* completeness score of the assembly was 97.5% and the LAI was 17.6. To further assess the quality of our scaffolded genome, we mapped ONT long-reads to the pseudomolecules, where 99.87 % of ONT long-reads aligned to our genome with high quality (average MAPQ = 47).

After aligning our assembly long-reads (i.e., KR1, KR2, and KR3) to our assembled genome we observed several positions in the *O. australiensis* KR assembly that have low long-read coverage (<20x, *cv*. 46x scaffold coverage; Supplementary Figure 5.7). These regions predominantly occur on chromosome 4 and coincide with regions that are composed of smaller contigs (Supplementary Figure 5.8). Additionally, chromosome 4 was constructed from a larger number of contigs than all other pseudomolecules (76 contigs compared to an average of 56 ± 9 contigs per pseudomolecule; Supplementary Figure 5.8).

Structural Variations (SVs)

We aligned the *O. australiensis* KR assembly to the *O. sativa* Nipponbare assembly and observed a high level of synteny between the two genomes (Figure 2). We also observed 378 instances of non-syntenic alignment (e.g., sequence that occurs on *O. sativa* chromosome 1 aligning to *O. australiensis* KR pseudomolecule 3). Thus, the genome–genome alignment indicated the presence of SVs. We investigated 21 SVs further (8 translocations, 7 inverted translocations, 5 inversions, and 1 duplication). The largest SV detected was an inversion that occurred on Chr11 and was 2.95 Mbp long within the *O. sativa* Nipponbare genome and 7.2 Mbp long in the *O. australiensis* KR genome assembly. We show that some of the SVs are unlikely to be caused by scaffolding errors (Supplementary Figure 5.8). Seven of the 21 reported SVs cross contig–contig boundaries. Further, by mapping the *O. australiensis* KR long reads to the pseudomolecules and observing continuous mapping of long reads at a depth of 31 - 75x coverage across SV junctions, we show that the investigated SVs are likely not artefacts of errors in the assembly of the contigs. All reported SVs require further investigation before they can be confirmed.

Recently, Stein et al.[1] showed an inversion on chromosome two of non-AA rice genomes relative to AA genome rice. This inversion was not detected in the assembly presented here (Fig. 5.2). We verified the inversion reported by Stein et al.[1] and the lack of inversion in the *O. australiensis* KR assembly by aligning *O. sativa* Nipponbare Chr2 with *O. punctata* Chr2 (Supplementary Figure 5.9A) and *O. australiensis* KR Chr2 (Supplementary Figure 5.9B), respectively. When we aligned *O. australiensis* KR Chr2 to *O. punctata* Chr2 we saw an inversion similar to the *O. sativa* Nipponbare vs *O. punctata* Chr2 alignment (Supplementary Figure 5.9C). To check the validity of this finding, we mapped *O. australiensis* KR long reads back to the assembly and observed 38 - 43x coverage at the alleged SV junction. Further, the coordinates of the alleged SV (Chr2:74,761,274-74,941,712) in the *O. australiensis* KR assembly, derived from the alignment of *O. australiensis* KR Chr2 to *O. punctata* Chr2 (Supplementary Figure 5.9C), sit well within the boundaries of a contig (tig00003125 occupies Chr2: 69,779,629-76,584,758). After aligning *O. punctata*-guided *O. australiensis* Chr2 to *O. punctata* Chr2 we again observed an inversion at these coordinates, but no inversion when *O. punctata*-guided *O. australiensis* Chr2 and *O. sativa* Nipponbare Chr2 (data not shown).

Transcripts and Genome Annotation

Using a single PacBio Sequel II 8M SMRT Cell we were able to generate 3.3 million polymerase reads. Once these data were collapsed to Circular Consensus Sequencing (CCS) reads and demultiplexed, we had 183k - 305k CCS reads per tissue. Once the demultiplexed CCS reads were clustered we had 21k - 35k polished high quality (HQ) isoforms per tissue (Table 5.3). *BUSCO* analysis indicated that the Iso-Seq transcripts from the five selected tissues accounted for ~40% of the 4,896

sequences in the *Poales* database. Only the HQ transcripts were used for annotation (low quality transcripts were discarded). A number of transcript isoforms were detected (see Supplementary Figure 5.4B-D for examples of transcript isoforms for genes involved in stress response in rice).

Table 5.3. Summary statistics of the Pac Bio Iso-Seq data.

Tissue	# CCS Reads	Mean Length (kbp)	# Polished Isoforms	
			High Quality	Low Quality
Leaf	242,841	2.69	25,805	6
Coleoptile	182,785	2.29	21,013	2
Root Tip	297,806	2.16	33,619	3
Growth Zone	304,934	2.75	35,048	12
Developing Grain	239,213	2.46	23,176	6

Following the production of pseudomolecules via reference-guided scaffolding, *LTR_Retriever* identified 608.7 Mbp (61.6%) LTR-RT content in the assembly (assigned and unassigned contigs inclusive). *EDTA* identified 730.1 Mbp of repeat sequences in the pseudomolecules (Table 5.4; Supplementary File S2b), including 53.07% LTR-RT sequences. A peak in insertion events for both *Gypsy* and *Copia* LTR-RT elements was observed at ~0.5 million years ago (MYA; Supplementary Figure 5.10). *Gypsy* repeat elements are more abundant than *Copia* elements in the *O. australiensis KR* genome. A third ‘unknown’ LTR-RT superfamily shows a slow increase in copy number over the last four million years (Supplementary Figure 5.10).

Table 5.4. Repeat elements in the *O. australiensis* KR scaffolds[51].

Class	Superfamily	Count	Masked (Mbp)	Masked (%)
LTR-RT	*Copia*	168,433	97.5	11.4
	Gypsy	299,854	319.2	37.2
	unknown	63,847	39.1	4.6
Non-LTR-RT	---	365,794	167.2	19.5
TOTAL	---	897,928	623	72.7

Using the HQ Iso-Seq transcripts and protein sequences from the *Oryza* genus as evidence*, OmicsBox* identified 52,090 genes and 60,325 transcripts in the assembly (Supplementary Files S3c and S4d). *OmicsBox* functionally annotated 35,620 genes and 43,362 transcripts. To check whether our annotation included TE genes, we searched the gene transcripts for known plant TE families[55]. Of our annotated genes 1033 sequences (829 *Gypsy* LTR-RTs, 81 *Copia* LTR-RTs, 4 unknown LINEs, 105 CACTA TIRs, 4 MuDR TIRs, 8 Pong TIRs, and 2 unknown non-LTR-RT repeat elements) were found to be TE genes and removed from the *O. australiensis* gene annotation. After accounting for the presence of TE genes we were left with 51,057 predicted genes, 34,587 functionally annotated genes, and 42,329 functionally annotated transcripts.

Genotypic Diversity and Genetic Similarity across four **O. australiensis** ***accessions and within the genus*** **Oryza**

After mapping the *O. australiensis* KR short reads to the *O. australiensis* KR assembly, we detected ~2 million SNVs. In comparison, we detected ~2.9 million, ~5 million, and ~6.4 million SNVs when mapping the CH, D, and VR accession short reads to the *O. australiensis* KR assembly. In the samples we analysed, genome wide SNV number did not increase with increasing geographic distance (Fig. 5.1). On the other hand, Figure 3 shows that the CH accession is most closely related to KR, followed by VR, and then D – that is, genetic distance (as determined by *kWIP*) increases with increasing geographic distance (Fig. 5.1 and Fig. 5.3).

Genetic distance between 15 species/accessions (representing the AA, BB, CC, BBCC, CCDD, EE, FF, GG, HHJJ, and HHKK genomes) within the *Oryza* genus was estimated with kWIP. A full list of the short read accessions used can be found in

Supplementary File S5e. The genetic distances estimated by kWIP allowed for the reconstruction of the *Oryza* phylogeny except for two 'misplaced' samples (10. *O. ridleyi*, and 13. *O. australiensis*; Fig. 5.3). We were able to re-construct the canonical *O. sativa* (AA) and the *O. officinalis* (BB, CC, CCDD, EE) genome complexes, and the remaining *Oryza* genomes (FF, GG, HHJJ, HHKK) form a third complex. The *O. australiensis* lineage diverges from the rest of the *O. officinalis* complex members (Fig. 5.3).

Mapping the *O. australiensis* KR short reads to the *O. sativa* Nipponbare reference genome revealed ~2.9 million SNVs between *O. australiensis KR* and *O. sativa* Nipponbare (Supplementary Figure 5.11). There are regions on all *O. sativa* chromosomes that, from visual inspection, have low SNV density. Also, the density of SNVs appears to increase at both terminal ends of the *O. sativa* chromosomes, though this may be due to uneven short read mapping density across the *O. sativa* Nipponbare genome (Supplementary Figure 5.12).

Discussion:

In this study we report on the first long-read assembly of an *O. australiensis* nuclear genome. Our results indicate that the assembled genome is of high quality and completeness. The assembled pseudomolecules (excluding ChrUn) are, on average, 2.2 times larger than the *O. sativa* Nipponbare reference chromosomes, which is in line with previous observations that *O. australiensis* chromosomes are larger than other *Oryza* species when seen during mitosis[58]. Our scaffolded *O. australiensis* genome assembly (858 Mbp) is smaller than estimated by flow cytometry (965 Mbp[16,58]), however our k-mer based estimation of genome size using *O. australiensis* short reads predicted a genome of 822 Mbp. The discrepancy in size between the

genome reported here and the estimated genome size for *O. australiensis*[15,16] may come about due to irresolvable LTR-RT elements.

After aligning the ONT long reads used to assemble the *O. australiensis* KR contigs to the reference-guided pseudomolecules, we observed unequal coverage distribution on some regions of the pseudomolecules (Supplementary Figure 5.7). Most often, coverage was halved in these regions (for example, see Chr4). Having purged the assembly of haplotigs it is unlikely that low coverage results from the presence of homologous maternal and paternal sequences being present in the assembly. Instead, it appears that the low coverage came about due to the relatively high number of small contigs that make up some parts of the pseudomolecules. For example, Chr4, which has the lowest median long read coverage, is composed of the most contigs compared to all other pseudomolecules (apart from ChrUn; Supplementary Figure 5.8). The fragmented nature of the contigs that have been used to build these regions of Chr4, and the other pseudomolecules, is characteristic of repetitive genome regions that are typically difficult to assemble. To overcome the limitations imposed by highly repetitive genome regions and, thus, to resolve the sequences therein, it is recommended that future studies employ long reads sufficient for a higher genome coverage.

We found that ~74% of the assembly is repeat sequence, of which ~55.5 - 61.6% was identified as LTR-RTs. These estimates are in line with previous estimates (~65%) of *O. australiensis* genome repeat content[15]. Stein et al.[1] report a maximum total repeat content of 49.6% for *O. punctata* (BB genome), and a minimum of 27.2% for *O. meridionalis* (AA genome), thus *O. australiensis* has the highest total repeat content among the few *Oryza* species sequenced to date. We show that there was a

relatively recent (~0.5 MYA) increase in LTR-RT (*Gypsy* and *Copia*) composition of the *O. australiensis KR* genome, in line with previous observations of *O. australiensis* Bacterial Artificial Chromosome (BAC) library sequences[15]. Similar trends of expansion for these LTR-RT types have been documented for both *O. sativa* indica and *O. sativa* japonica varieties[1]. However, along with a recent expansion of these LTR-RTs, the *O. sativa* genomes have undergone subsequent substantial gene loss, which has played a role in determining the genome size[59]. We also report on the expansion of an unknown LTR-RT family that appears to be in progress in modern-day *O. australiensis* plants, similar to the two *O. sativa* subspecies mentioned above and other *O. australiensis* accessions[15]. Piegu et al.[15] showed that the LTR-RT element called '*Kangourou'*, a low copy number LTR-RT relative to *Gypsy* or *Copia*, had a peak burst time of ~2.3 MYA. The magnitude of the copy number burst in the *O. australiensis* genome of *Kangourou* in previous studies using BAC libraries[15] matches that of the 'unknown' LTR-RT element we report here, however our results suggest a much more recent (~0.2 MYA) but slower increase in copy number. Other species of wild rice appear to have more ancient LTR-RT family (*Gypsy*, *Copia*, and unknown) expansions. For example, the *Gypsy*, *Copia*, and unknown LTR-RT families of *O. brachyantha* peaked in abundance around 1.8, 2, and 2.5 MYA, respectively[1]. Shenton et al.[13] show a burst in the *Mutator* MuDR2 family composition of the *O. officinalis* genome complex. Further, *O. meridionalis* seems to have more dated (~1 - 1.5 MYA) expansions and then contractions of the same LTR-RTs relative to *O. australiensis* KR[1]. However, given the differences in the assembly methods and qualities of the various *Oryza* genomes, it is difficult to make direct comparisons of LTR-RT family ages and so these speculations should be viewed with caution.

After accounting for TE genes in the annotation file, we predicted 51,057 genes and functionally annotated 34,587 of those genes. This is in line with previously reported *Oryza* genome annotations. For example, Stein et al.[1] report a maximum of 38,550 annotated loci for O*. sativa* ssp. *japonica*, and a minimum of 24,208 for *O. brachyantha*. In comparison, Release 7 of the *O. sativa* Nipponbare genome contains 39,045 non-TE annotated genes and 49,066 non-TE gene models[46]. Thus, the number of genes/transcripts annotated in this version of the *O. australiensis KR* genome assembly is comparable to that of the *O. sativa* Nipponbare version 7 genome. This is consistent with the hypothesis that the size of the *O. australiensis* genome is the result of the expansion of LTR-RT elements, rather than gene duplication events[15]. However, similar numbers of coding genes invite speculation about a functional role for expansion of LTR-RT elements. For example, it is conceivable that the especially adverse environments where *O. australiensis* evolved in northern Australia have selected for expansion of particular gene families involved in tolerance to stress.

Comparing the *O. australiensis* KR genome assembly to the *O. sativa* Nipponbare assembly, we identified a number of large structural variations. Recently, Stein et al.[1] showed an inversion on chromosome two of non-AA rice genomes relative to AA genome rice. The same inversion appears in *Leersia perrieri* and *Brachypodium distachyon* and so it may be a conserved feature in grass genomes, except in rice with the AA genome[1]. This inversion was not detected by reference-guided construction of the *O. australiensis* assembly (Fig. 5.2), suggesting a need for deeper coverage of this region in future studies. However, we can report 36 - 45x coverage at the putative inversion junction in the *O. australiensis* KR assembly, and we show that the putative SV sits well within the boundaries of a single contig. Hence, we

suggest that this region of the *O. australiensis* KR assembly is not the result of misassembly. The possibility remains that it is an artefact of reference-biased scaffolding, however we found that the inversion is missing in the *O. australiensis* assembly even after using a BB rice genome (*O. punctata*) to order and orient the contigs (data not shown), so this is unlikely. The SV reported by Stein et al.[1] in non-AA rice genomes may truly be absent from the *O. australiensis* KR genome. One hypothesis for its absence is hybridisation between the co-occurring AA-genome species (*O. meridionalis* and *O. rufipogon*) and *O. australiensis*, thus eliminating the Chr2 SV in the *O. australiensis* KR genome. However, given the difficulty in producing viable progeny from artificial crosses between the AA genome species and *O. australiensis*[60], we find this to be an improbable explanation for the observed results. Thus, in the absence of further evidence, all structural variations should be viewed with caution.

Given that the 21 large SVs we observed do not appear to be the result of either scaffolding or assembly artefacts, if these SVs prove to be real, they highlight considerable differences in the structure of the *O. australiensis KR* assembly and *O. sativa*. These SVs, and the substantial repeat content of the wild rice genome, would offer insight into why crossing the AA genome *O. sativa* and the EE genome *O. australiensis* is difficult and why successful crosses have abnormal chromosomal arrangements and the resulting progeny are sterile[61,62]. To validate the putative SVs, future scaffolding attempts could make use of other *Oryza* species for reference-guided scaffolding (e.g., the *O. officinalis* assembly may be a better candidate for scaffolding than the *O. sativa* Nipponbare assembly as *O. australiensis* is a member of the *O. officinalis* complex). However, this would still present biases in contig ordering and orientation. Therefore, we recommend that future scaffolding attempts either make use of pre-existing *O. australiensis* BAC libraries and FingerPrinted

Contigs (FPC) to order and orient the contigs[1,13] or use Hi-C technology to generate new data for *O. australiensis*. Given the still experimental nature of Hi-C and difficulties in obtaining high coverage of linked reads these efforts may require significant investment.

Along with structural variations in the *O. australiensis KR* genome relative to *O. sativa* Nipponbare, we also uncovered many SNVs (Supplementary Figure 5.11). Interestingly, there were more SNVs between some *O. australiensis* genotypes and *O. australiensis* KR than there were between *O. sativa* and *O. australiensis* KR based on short read mapping. The density of SNVs increases at the terminal regions of the *O. sativa* chromosomes. Conversely, there is a region on each chromosome where the density of SNVs is low. These regions may represent the centromeric repeat elements of the chromosome (Supplementary Figure 5.11). While centromeric regions have a common functional role in organising cell division, their sequences can be diverse between species. However, in the case of rice and other cereals, they are highly conserved[63,64]. For example, Gao et al.[65] showed that *O. australiensis* and *O. sativa* share similar centromeric retrotransposon repeat (CRR) sequences even though other rice species such as *O. brachyantha* have unique CRR profiles. This helps explain why sequence similarity between *O. australiensis KR* and *O. sativa* at centromeric regions is higher than at other regions of the genome.

Interestingly, when comparing four genotypes of *O. australiensis* that had been documented for their contrasting salt tolerance[20], we found that SNV density was not well correlated with the geographical distance between each of the populations (Fig. 5.1). For example, VR had at least 1.5 million more SNVs than the other two genotypes (when compared with KR), yet it was collected from a site far closer to

Keep River than the D genotype. We speculate that SNV densities may be more related to the diversity of alleles required for adaptation to stresses than physical separation between populations. For example 'relative salt tolerance'[20] (Fig. 5.1) appeared to correlate better with the abundance of SNV abundance than simple distance between populations.

Contrary to the SNV data, the genetic distances estimated by kWIP *do* reflect the geographic distance between the *O. australiensis* genotypes (Fig. 5.3). That is, populations of *O. australiensis* genotypes that occur closer to each other in space are more genetically similar than genotypes that occur further apart. This is expected as *O. australiensis* is a mostly inbreeding species (~3% outcrossing), so gene flow and introgressions between genotypes should be low. Thus, we assume that these populations of *O. australiensis* are reproductively isolated and, as such, their genomes are evolving independently. If this is true, we expect to see local adaptation in these genotypes[66], a hypothesis which supports the findings of Yichie et al.[20].

The genetic-distance tree generated by *kWIP* follows previously published *Oryza* phylogenies[13]. Importantly, the *O. australiensis* lineage diverges from the rest of the *O. officinalis* complex members (the B, C, and D genomes) suggesting that we can expect to see lineage specific adaptation in the *O. australiensis* genome. Given that this species occupies a diverse yet adverse suite of environmental niches, the genomic variation that underpins the divergence of the *O. australiensis* lineage may be associated with tolerance to multiple abiotic and biotic stresses. The functionally annotated genome that we present for *O. australiensis* KR here is therefore a resource that can be used to explore this genomic variation to uncover genes involved in stress tolerance.

Conclusions:

Here we present the first reference-quality genome of *O. australiensis KR* – a wild species of rice native to the northern regions of Australia. While *O. australiensis* co-occurs with other *Oryza* species, it is genetically and phenotypically unique among the *Oryza* genus and is the sole member of the EE genome clade[67]. It demonstrates unrivalled tolerance to extreme environmental conditions, including heat, drought and soil salinity[3,18,21,68]. The leaf anatomy and associated photosynthetic efficiency of *O. australiensis* are also unique traits[22] and its grain has distinct pigmentation and starch composition[23,24]. Such traits make it an important genetic resource in the *Oryza* genus, as recognised formally by the OMAP and IOMAP[11,12]. However, no group has successfully assembled the nuclear genome for this species until now. It is expected that this genomic resource will enable phenomics in wild cereal relatives and lead to the selection of useful traits for the improvement of domestic rice cultivars, and in the breeding and domestication of *O. australiensis* itself. For example, given that we now know the positions of genes in the *O. australiensis* KR assembly, one can use the assembly alongside short reads from diverse *O. australiensis* accessions to identify SNVs in genes associated with tolerance in accessions from adverse soils. Similarly, with further Iso-Seq data one could identify novel transcript isoforms across different *O. australiensis* accessions under salt stress. Furthermore, given the role that gene regulation plays in response to stress, the *O. australiensis* KR assembly will uncover promoters of genes of interest. This work could extend to the use of long read data (either PacBio or ONT reads) to investigate the response of the epigenome to abiotic stresses. Finally, an expanded collection of *O. australiensis* accessions with contrasting phenotypes under extreme conditions will facilitate the identification of markers for stress tolerance via strategic inclusion of tolerant accessions to make mapping populations.

Abbreviations:

AGP: A Golden Path; ANU: Australian National University ; BAM: Binary Alignment Map; BCF: Binary Variant Call Format; BLAST: Basic Local Alignment Search Tool; bp: Base Pairs; BUSCO: Benchmarking Universal Single-Copy Orthologs; Chr: Chromosome; ChrUn: Chromosome Unassigned; DVC(R): Deputy Vice Chancellor (Research); EDTA: Ethylenediaminetetraacetic acid; EDTA: Extensive de novo TE Annotator; GFF3: General Feature Format 3; GO: Gene Ontology; HPC: High Performance Computer; HQ: High Quality; IGV: Integrative Genomics Viewer; IOMAP: International *Oryza* Map Alignment Project; Iso-Seq: Isoform Sequencing; Kbp: Kilo Base Pairs; LAI: LTR Assembly Index; LQ: Low Quality; LTR: Long Terminal Repeat; RT: Retrotransposon; Mbp: Mega Base Pairs; MYA: Million Years Ago; NCBI: National Center for Biotechnology Information; NSW: New South Wales; OMAP: *Oryza* Map Alignment Project; ONT: Oxford Nanopore Technologies; PAF: Pairwise mApping Format; PEG: Polyethylene glycol; PIF: P Instability Factor; PVP: Polyvinylpyrrolidone; RIN: RNA Integrity Number; RMBlast: RepeatMasker compatible NCBI BLAST suite; SAM: Sequence Alignment Map; SNV: Single Nucleotide Variation; SV: Structural Variation; TIR: Terminal Inverted Repeat ; VCF: Variant Call Format.

Additional Information

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Data Availability

This Whole Genome project (i.e., the *Canu*-generated purged contigs reported here) has been deposited at DDBJ/ENA/GenBank under the accession JAIFGZ000000000. The version described in this paper is version JAIFGZ010000000. All long- and short-read data associated with the reported genome assembly have been deposited at the Sequence Read Archive and GenBank in NCBI under BioProject PRJNA743927. All Supplementary Figures and Supplementary Files (the AGP file required to construct the pseudomolecules from the *Canu* contigs, the annotation for protein-coding genes in the contigs, the annotation for repeat elements in the scaffolds, and the annotation for protein-coding genes in the scaffolds (all GFF3 files)) can be accessed via figshare:

https://doi.org/10.6084/m9.figshare.c.5875592.v2

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded by the Macquarie University DVC(R) Fund with contributions from the Department of Biological Sciences. This project was supported by the ARC Centre of Excellence in Plant Energy Biology (CE140100008). AP acknowledges support from an Australian Research Training Program Scholarship and the FJ Sandoz Scholarship from the University of Adelaide.

Authors' contributions

AP, SF, NWH, AJ, JB, RB, and BA contributed to technical planning after BA initiated the project with a successful application for a Macquarie University DVC(R)

Discretionary grant. JB and AJ provided technical expertise for execution of the sequencing. AP and AJ performed all DNA and RNA preparations and performed all ONT MinION sequencing steps. AJ performed Illumina sequencing steps, gave initial bioinformatics support, and supplied FASTQ files. NWH performed Iso-Seq analysis and provided bioinformatics support. AP and SF performed all data analysis, genome assembly, and subsequent analyses. NWH and RB provided access to the computing resources required. NWH, RB, and BA provided guidance throughout. BA grew the plants for sequencing prior to extractions in Canberra. AP created all figures and wrote the initial draft of the manuscript. All authors contributed to the final manuscript.

Acknowledgements

We would like to acknowledge Ming-Dao Chia for providing initial bioinformatics support. We thank Scott Ferguson, Lina Herliana, Chelsea Matthews, and Farhad Masoomi-Aladizgeh for the discussions about bioinformatics analyses. We thank Dr Sally Norton from the Australian Grains Genebank in Horsham Victoria, Australia for providing the seed a decade ago that led to the present study. We thank Dr. James Cowley (Founder of 'The Figure Legend') for his advice on designing figures. We would like to deeply thank the reviewers, whose thoughtful comments helped substantially improve the manuscript.

References:

1. Stein, J. C. *et al.* Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. *Nature Genetics* **50**, 285–296 (2018).

2. Brar, D. S. & Khush, G. S. Wild relatives of rice: a valuable genetic resource for genomics and breeding research. in *The Wild Oryza Genomes* (eds. Mondal, T. K. & Henry, R. J.) 1–25 (Springer International Publishing, 2018). doi:10.1007/978-3-319-71997-9_1.
3. Atwell, B. J., Wang, H. & Scafaro, A. P. Could abiotic stress tolerance in wild relatives of rice be used to improve Oryza sativa? *Plant Science* **215–216**, 48–58 (2014).
4. Bisht, D. S., Solanke, A. U. & Mondal, T. K. Informatics of wild relatives of rice. in *The Wild Oryza Genomes* (eds. Mondal, T. K. & Henry, R. J.) 27–40 (Springer International Publishing, 2018). doi:10.1007/978-3-319-71997-9_2.
5. Huang, X. *et al.* Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics* **42**, 961–967 (2010).
6. Bolser, D., Staines, D. M., Pritchard, E. & Kersey, P. Ensembl Plants: integrating tools for visualizing, mining, and analyzing plant genomics data. *Methods in molecular biology (Clifton, N.J.)* **1374**, 115–140 (2016).
7. Henry, R. J. *et al.* Australian Oryza: utility and conservation. *Rice* **3**, 235–241 (2010).
8. Hiromi, K.-K. *et al.* OryzaGenome2.1: database of diverse genotypes in wild Oryza species. *Rice* **14**, (2021).
9. Yu, H. *et al.* A route to de novo domestication of wild allotetraploid rice. *Cell* **184**, 1156-1170.e14 (2021).
10. Song, S., Tian, D., Zhang, Z., Hu, S. & Yu, J. Rice genomics: over the past two decades and into the future. *Genomics, Proteomics & Bioinformatics* **16**, 397–404 (2018).

11. Wing, R. A. *et al.* The Oryza map alignment project: the golden path to unlocking the genetic potential of wild rice species. *Plant Molecular Biology* **59**, 53–62 (2005).
12. Jacquemin, J., Bhatia, D., Singh, K. & Wing, R. A. The International Oryza Map Alignment Project: development of a genus-wide comparative genomics platform to help solve the 9 billion-people question. *Current Opinion in Plant Biology* **16**, 147–156 (2013).
13. Shenton, M. *et al.* Evolution and Diversity of the Wild Rice Oryza officinalis Complex, across Continents, Genome Types, and Ploidy Levels. *Genome Biology and Evolution* **12**, 413–428 (2020).
14. Shi, C. *et al.* The draft genome sequence of an upland wild rice species, Oryza granulata. *Scientific Data* **7**, 131 (2020).
15. Piegu, B. *et al.* Doubling genome size without polyploidization: Dynamics of retrotransposition-driven genomic expansions in Oryza australiensis, a wild relative of rice. *Genome Research* **16**, 1262–1269 (2006).
16. Ammiraju, J. S. S. *et al.* The Oryza bacterial artificial chromosome library resource: construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genome types of the genus Oryza. *Genome research* **16**, 140–147 (2006).
17. Zhang, Q. J. & Gao, L. Z. Rapid and recent evolution of LTR retrotransposons drives rice genome evolution during the speciation of AA-genome Oryza Species. *G3: Genes, Genomes, Genetics* **7**, 1875–1885 (2017).
18. Phillips, A. L., Scafaro, A. P. & Atwell, B. J. Photosynthetic traits of Australian wild rice (Oryza australiensis) confer tolerance to extreme daytime temperatures. *Plant Molecular Biology* (2022) doi:10.1007/s11103-021-01210-3.

19. Scafaro, A. P. *et al.* A thermotolerant variant of Rubisco activase from a wild relative improves growth and seed yield in rice under heat stress. *Frontiers in Plant Science* **9**, 1–11 (2018).
20. Yichie, Y., Brien, C., Berger, B., Roberts, T. H. & Atwell, B. J. Salinity tolerance in Australian wild Oryza species varies widely and matches that observed in O. sativa. *Rice* **11**, 1–14 (2018).
21. Hamzelou, S. *et al.* Wild and Cultivated Species of Rice Have Distinctive Proteomic Responses to Drought. *International Journal of Molecular Sciences* vol. 21 Preprint at https://doi.org/10.3390/ijms21175980 (2020).
22. Mathan, J., Singh, A., Jathar, V. & Ranjan, A. High photosynthesis rate in two wild rice species is driven by leaf anatomy mediating high Rubisco activity and electron transport rate. *Journal of Experimental Botany* (2021) doi:10.1093/jxb/erab313.
23. Tikapunya, T., Fox, G., Furtado, A. & Henry, R. Grain physical characteristic of the Australian wild rices. *Plant Genetic Resources* **15**, 409–420 (2016).
24. Tikapunya, T. *et al.* Molecular structures and properties of starches of Australian wild rice. *Carbohydrate polymers* **172**, 213–222 (2017).
25. Joly-Lopez, Z. *et al.* An inferred fitness consequence map of the rice genome. *Nature Plants* **6**, 119–130 (2020).
26. Michael, T. P. & VanBuren, R. Building near-complete plant genomes. *Current Opinion in Plant Biology* **54**, 26–33 (2020).
27. Jones, A. *et al.* High-molecular weight DNA extraction, clean-up and size selection for long-read sequencing. *PloS One* **16**, (2021).
28. Jones, A., Borevitz, J., Warthmann, N. & Jones, A. Cost-conscious generation of multiplexed short-read DNA libraries for whole genome sequencing. *protocols.io* 1–13 (2018) doi:10.17504/protocols.io.unbevan.

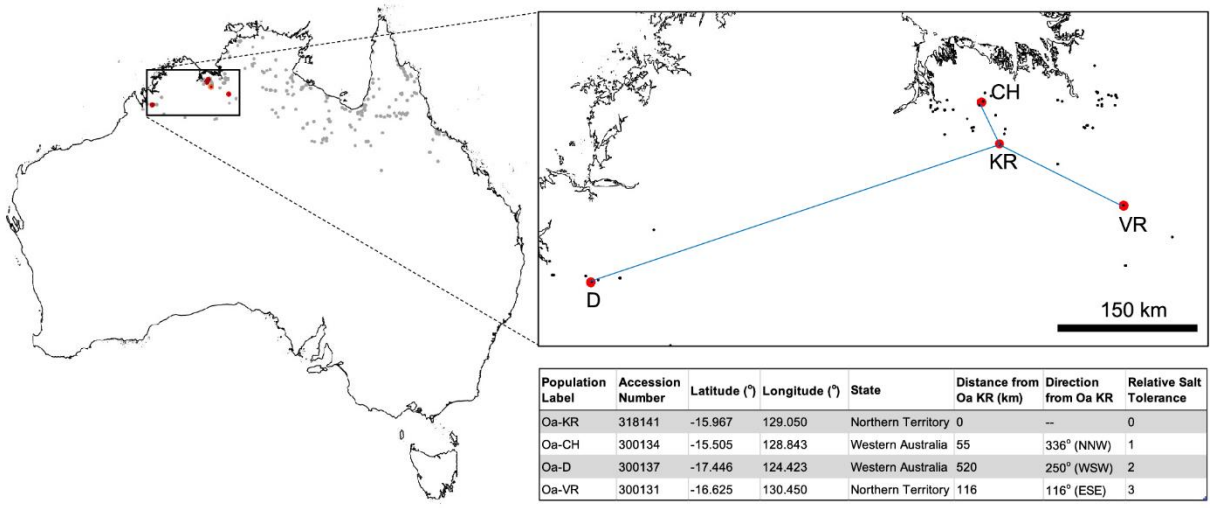
29. Betts, N. S. *et al.* Isolation of tissues and preservation of RNA from intact, germinated barley grain. *The Plant Journal* **91**, 754–765 (2017).
30. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764–770 (2011).
31. Vurture, G. W. *et al.* GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* **33**, 2202–2204 (2017).
32. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nature Communications* **11**, 1432 (2020).
33. Koren, S. *et al.* Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome research* **27**, 722–736 (2017).
34. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
35. Kundu, R., Casey, J. & Sung, W.-K. Hypo: super fast & accurate polisher for long read genome assemblies. *bioRxiv* 2019.12.19.882506 (2019) doi:10.1101/2019.12.19.882506.
36. Lindenbaum, P. JVarkit: java-based utilities for bioinformatics. Preprint at https://doi.org/10.6084/m9.figshare.1425030.v1 (2015).
37. Shen, W. & Xiong, J. TaxonKit: a cross-platform and efficient NCBI taxonomy toolkit. *bioRxiv* 513523 (2019) doi:10.1101/513523.
38. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* **14**, 178–192 (2012).
39. Roach, M. J., Schmidt, S. A. & Borneman, A. R. Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics* **19**, 460 (2018).

40. Ou, S. & Jiang, N. LTR_FINDER_parallel: parallelization of LTR_FINDER enabling rapid identification of long terminal repeat retrotransposons. *Mobile DNA* **10**, 48 (2019).
41. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* **9**, 18 (2008).
42. Ou, S., Chen, J. & Jiang, N. Assessing genome assembly quality using the LTR Assembly Index (LAI). *Nucleic Acids Research* **46**, e126–e126 (2018).
43. Seppey, M., Manni, M. & Zdobnov, E. M. BUSCO: Assessing Genome Assembly and Annotation Completeness. in *Gene Prediction* (ed. Kollmar, M.) 227–245 (Springer New York, 2019). doi:10.1007/978-1-4939-9173-0_14.
44. Jain, C., Koren, S., Dilthey, A., Phillippy, A. M. & Aluru, S. A fast adaptive algorithm for computing whole-genome homology maps. *Bioinformatics* **34**, i748–i756 (2018).
45. Alonge, M. *et al.* RaGOO: fast and accurate reference-guided scaffolding of draft genomes. *Genome Biology* **20**, 224 (2019).
46. Kawahara, Y. *et al.* Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. *Rice* **6**, 4 (2013).
47. Poorten, T. dotPlotly. Preprint at (2018).
48. Gel, B. & Serra, E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **33**, 3088–3090 (2017).
49. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
50. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

51. Ou, S. *et al.* Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biology* **20**, 275 (2019).
52. Wing, R. International Oryza Map Alignment Project. Preprint at https://doi.org/10.7946/P2FC9Z (2017).
53. Götz, S. *et al.* High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research* **36**, 3420–3435 (2008).
54. BioBam. OmicsBox – Bioinformatics Made Easy. Preprint at (2019).
55. Pedro, D. L. F. *et al.* An Atlas of Plant Transposable Elements [version 1; peer review: 2 approved]. *F1000Research* **10**, (2021).
56. Murray, K. D., Webers, C., Ong, C. S., Borevitz, J. & Warthmann, N. kWIP: The k-mer weighted inner product, a de novo estimator of genetic similarity. *PLOS Computational Biology* **13**, e1005727 (2017).
57. Crusoe, M. R. *et al.* The khmer software package: enabling efficient nucleotide sequence analysis [version 1; peer review: 2 approved, 1 approved with reservations]. *F1000Research* **4**, (2015).
58. Uozu, S. *et al.* Repetitive sequences: cause for variation in genome size and chromosome morphology in the genus Oryza. *Plant Molecular Biology* **35**, 791–799 (1997).
59. Vitte, C., Panaud, O. & Quesneville, H. LTR retrotransposons in rice (Oryza sativa, L.): recent burst amplifications followed by rapid DNA loss. *BMC genomics* **8**, 218 (2007).
60. Multani, D. S. *et al.* Development of monosomic alien addition lines and introgression of genes from Oryza australiensis Domin. to cultivated rice O. sativa L. *Theoretical and Applied Genetics* **88**, 102–109 (1994).

61. Abbasi, F. M., Ahmad, H., Sajid, M., Inamullah & Brar, D. S. Detecting Nature of Chromosome Pairing in A and E Genomes of Oryza. *Cytologia* **74**, 223–227 (2009).
62. Abbasi, F. M. *et al.* Assessment of genomic relationship between Oryza sativa and Oryza australinesis. *African Journal of Biotechnology* **9**, 1312–1316 (2010).
63. Aragón-Alcaide, L., Miller, T., Schwarzacher, T., Reader, S. & Moore, G. A cereal centromeric sequence. *Chromosoma* **105**, 261–268 (1996).
64. Jiang, J. *et al.* A conserved repetitive DNA element located in the centromeres of cereal chromosomes. *Proceedings of the National Academy of Sciences* **93**, 14210 LP – 14213 (1996).
65. Gao, D. *et al.* A lineage-specific centromere retrotransposon in Oryza brachyantha. *The Plant Journal* **60**, 820–831 (2009).
66. Lascoux, M., Glémin, S. & Savolainen, O. Local Adaptation in Plants. *eLS* 1–7 Preprint at https://doi.org/10.1002/9780470015902.a0025270 (2016).
67. Henry, R. J. Oryza australiensis Domin. in *The Wild Oryza Genomes* (eds. Mondal, T. K. & Henry, R. J.) 61–66 (Springer International Publishing, 2018). doi:10.1007/978-3-319-71997-9.
68. Yichie, Y. *et al.* Salt-treated roots of Oryza australiensis seedlings are enriched with proteins involved in energetics and transport. *Proteomics* **19**, 1900175 (2019).

Figures & Figure Legends:



Population Label	Accession Number	Latitude (°)	Longitude (°)	State	Distance from Oa KR (km)	Direction from Oa KR	Relative Salt Tolerance
Oa-KR	318141	-15.967	129.050	Northern Territory	0	--	0
Oa-CH	300134	-15.505	128.843	Western Australia	55	336° (NNW)	1
Oa-D	300137	-17.446	124.423	Western Australia	520	250° (WSW)	2
Oa-VR	300131	-16.625	130.450	Northern Territory	116	116° (ESE)	3

Figure 5.1. Locations of *O. australiensis* seed collection sites for this study. List of wild rice accessions (obtained from the Australian Grain Genebank) used in this study reported by Yichie *et al*.[19] to demonstrate varying degrees of salt tolerance. Dots on the map show the occurrence of *O. australiensis* (retrieved from the Atlas of Living Australia). In-set table: accessions are ordered by salt tolerance, with Oa-KR being the most sensitive to salt, and Oa-VR being most tolerant[19].

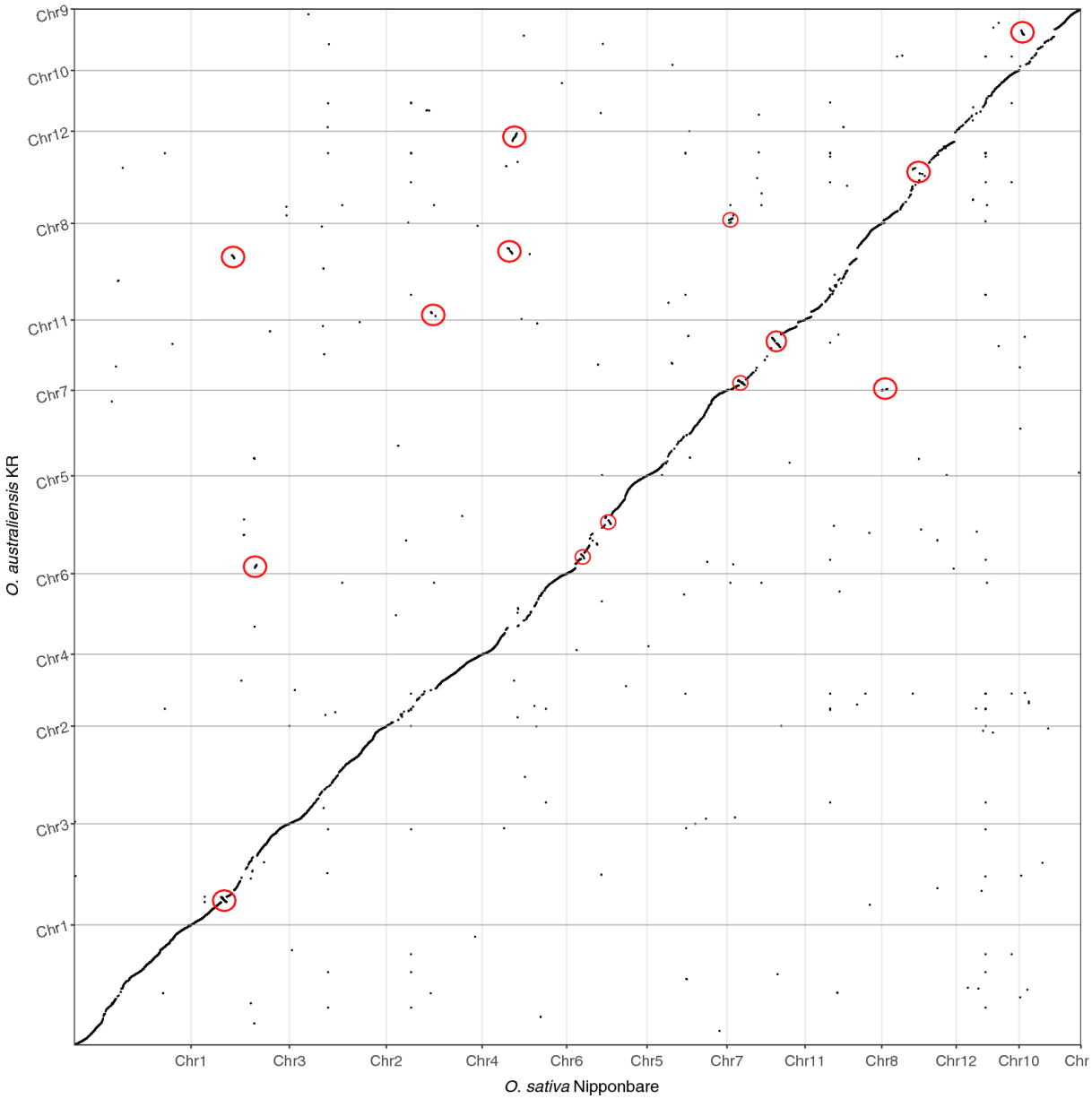


Figure 5.2. Alignment of scaffolded *O. australiensis* KR contigs (named Chr1-12) to the *O. sativa* Nipponbare reference genome. The wild rice genome was aligned to the domestic rice genome using *minimap2* and visualised using *dotplotly*. The 21 SVs that were investigated further are circled in red (some circles contain multiple SVs). Note: chromosomes do not appear in numerical order – *dotPlotly* orders the target sequence (*O. sativa*, here) by chromosome size; ChrUn is not included as it did not contain any large alignments.

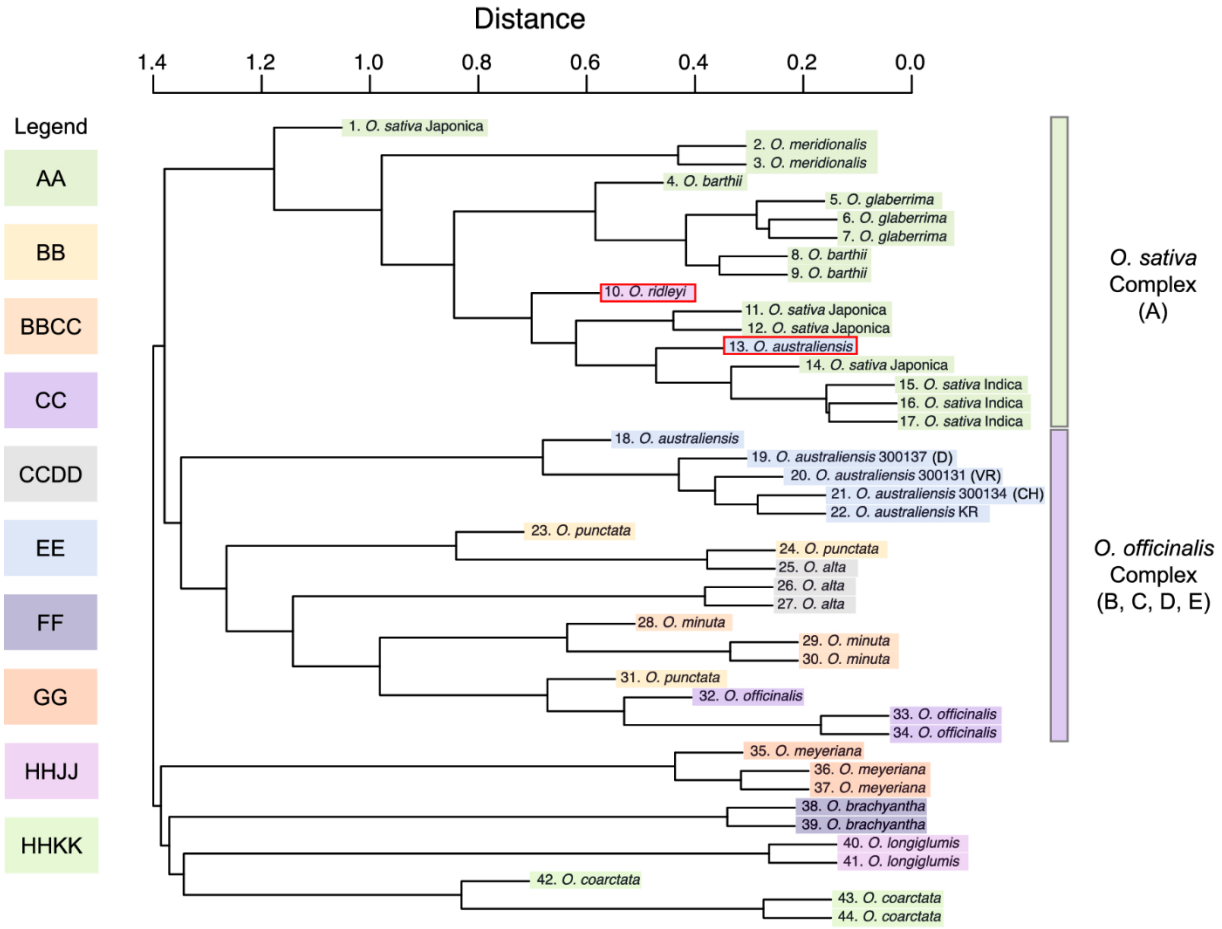
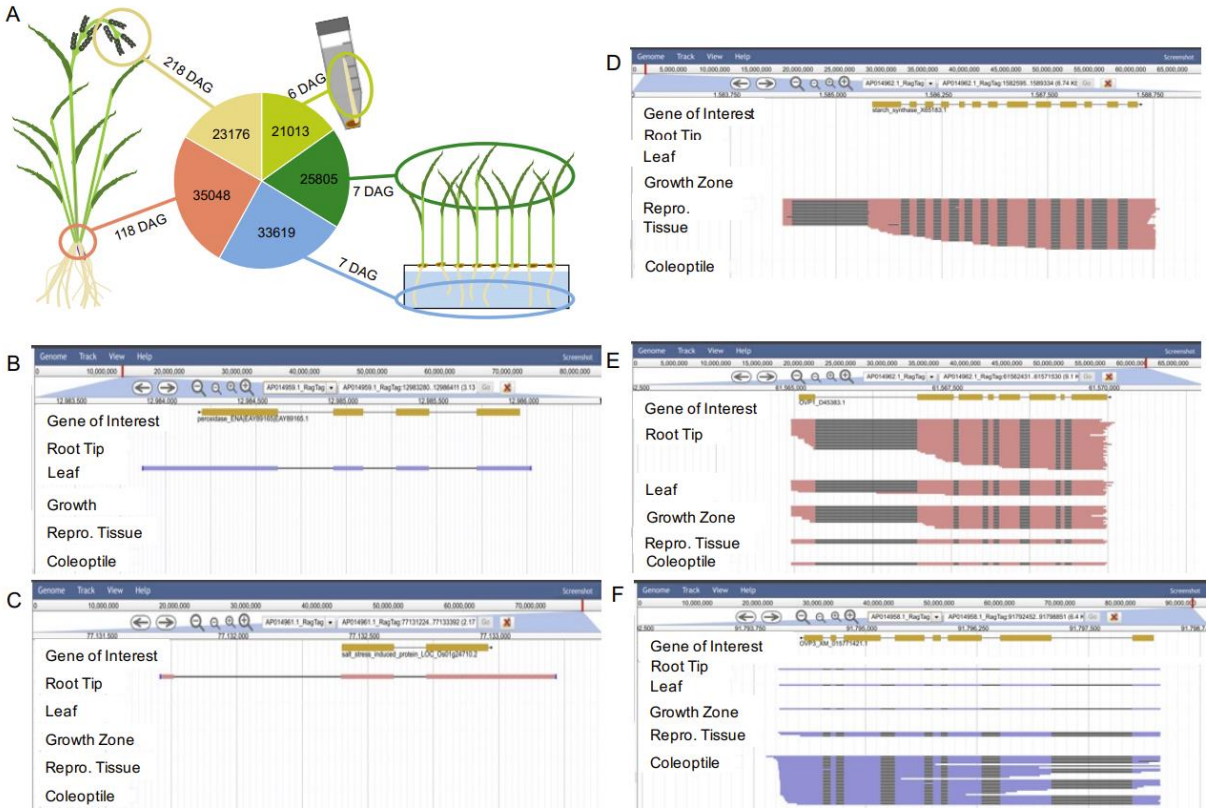


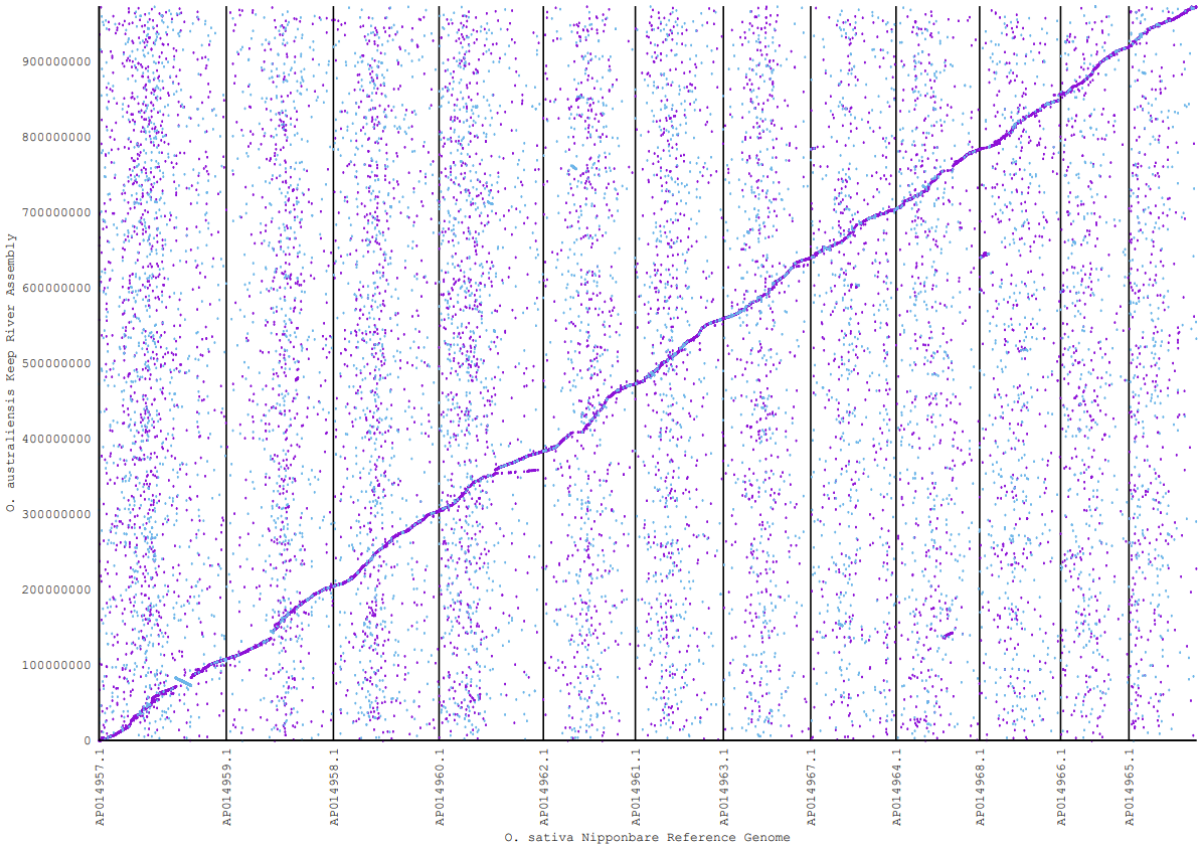
Figure 5.3. Genetic distance between *Oryza* species/accessions derived from Illumina short read libraries by *kWIP*. Samples are coloured based on the *Oryza* genome clade that they occupy (see inset Legend). *kWIP* clustered samples into the canonical *O. sativa* (AA genome) and O. officinalis (BB, CC, BBCC, CCDD, EE genomes) complexes. The *O. australiensis* lineage is divergent from the rest of the *O. officinalis* genome complex, suggesting it harbours lineage specific adaptations that could be explored for stress tolerance. Within the *O. australiensis* lineage, the four genotype samples re-sequenced in the present study (KR, CH, VR, and D) show genetic distances that are correlated with geographic distance. The sample divergence between KR, CH, VR, and D, as well as the other *O. australiensis* accessions shown here suggests that there may be within-species genomic variation that can also be explored for tolerance to stresses (e.g., genotype-specific tolerance to salt stress). Samples that appear with red borders are suspicious as they did not cluster with their corresponding genome clade. This may be due to errors in the sequencing files that were not corrected prior to running *kWIP* or may reflect human error during sample preparation for sequencing.

Supplementary Figures (please also see link in “Data Availability”)

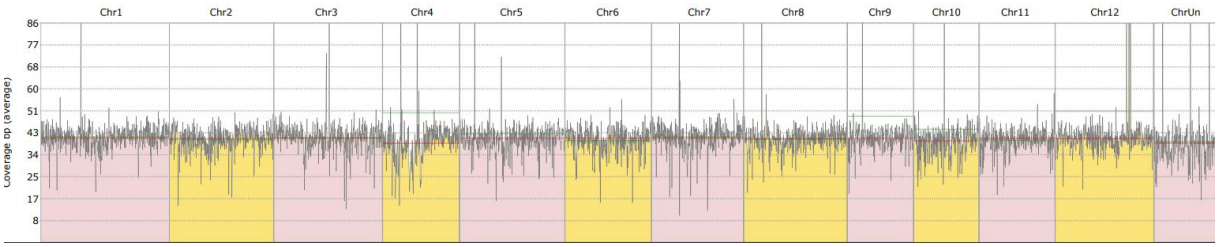


Supplementary Figure 5.4. High-quality transcripts derived from Iso-Seq data. (A) total abundance of HQ transcripts for five O. australiensis Keep River tissues (red = reproductive tissue [rachillae, developing seed, and pollen], dark blue = 7-day old leaves, light blue = 7-day old root tips, green = 6-day old coleoptiles, orange = growth zones). (B-F) tissue-specific detection of select transcripts (peroxidase [EAY89165.1], salt-stress-induced protein [LOC_Os01g24710.2], starch synthase [X65183.1], Oryza vacuolar pyrophosphatase 1 [OVP1; D45383.1], and Oryza vacuolar pyrophosphatase 3 [OVP3; XM_015771421.1], respectively) involved in response to several abiotic stressors.

Figure too large to display. Please access the figure at link provided above.
Supplementary Figure 5.5 Long-read coverage of the *O. australiensis* Keep River contigs. Long reads used to assemble the contigs were mapped to the contigs and the resulting indexed BAM file was used by JVarkit's wgscoverageplotter tool to find genome coverage. The green line represents the average coverage, and the red line represents the median coverage. Note: not all contigs are shown here.

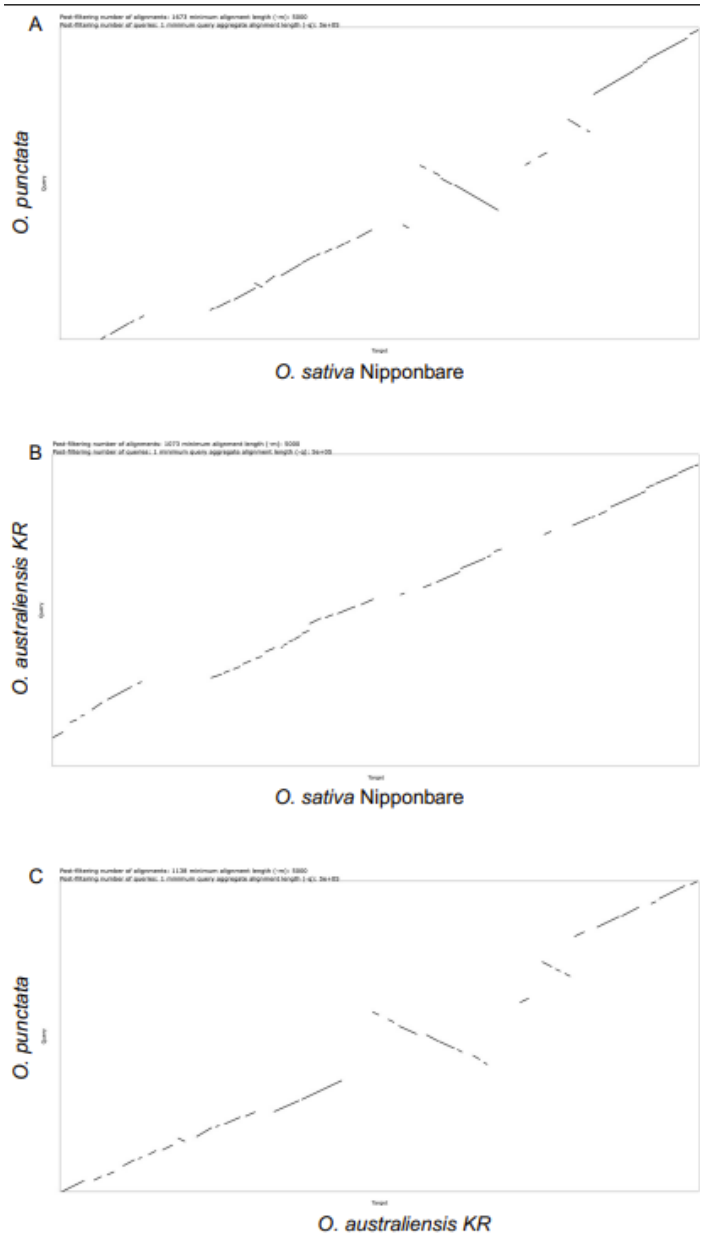


Supplementary Figure 5.6. Whole genome alignment of *O. australiensis* Keep River contigs and *O. sativa* Nipponbare chromosomes. Shows a genome-genome alignment between the *O. australiensis* Keep River contigs described in this paper and the *O. sativa* Nipponbare reference genome. Contigs were mapped to the *O. sativa* Nipponbare chromosomes using minimap2 and the resulting alignment was visualised, using MashMap. The X-axis shows the *O. sativa* chromosomes, while the Y-axis shows the *O. australiensis* contigs, expressed in Mbp.

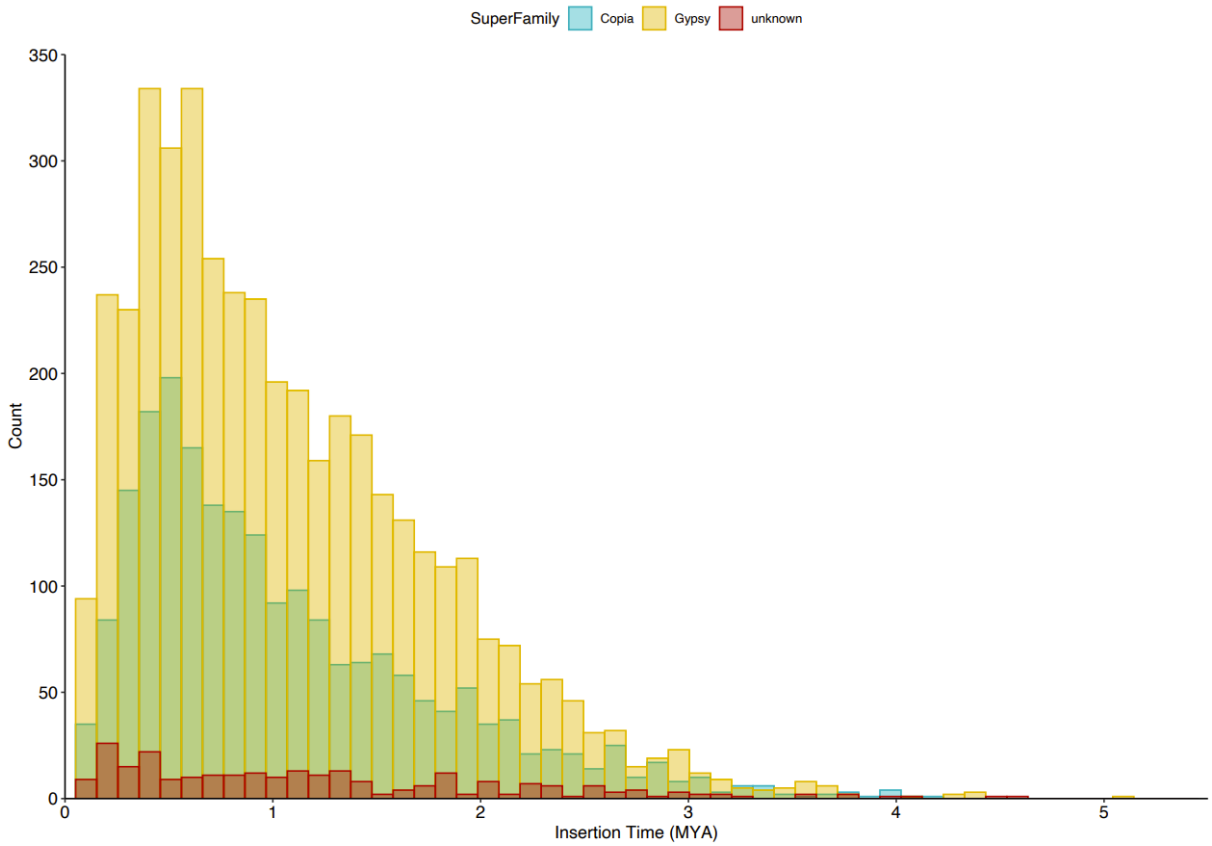


Supplementary Figure 5.7. Long-read coverage of the *O. australiensis* Keep River pseudomolecules. Shows a long-read coverage plot for the *O. australiensis* Keep River pseudomolecules. Long reads used to assemble the contigs were mapped to the pseudomolecules and the resulting indexed BAM file was used by JVarkit's wgscoverageplotter tool. The green line represents the average coverage, and the red line represents the median coverage.

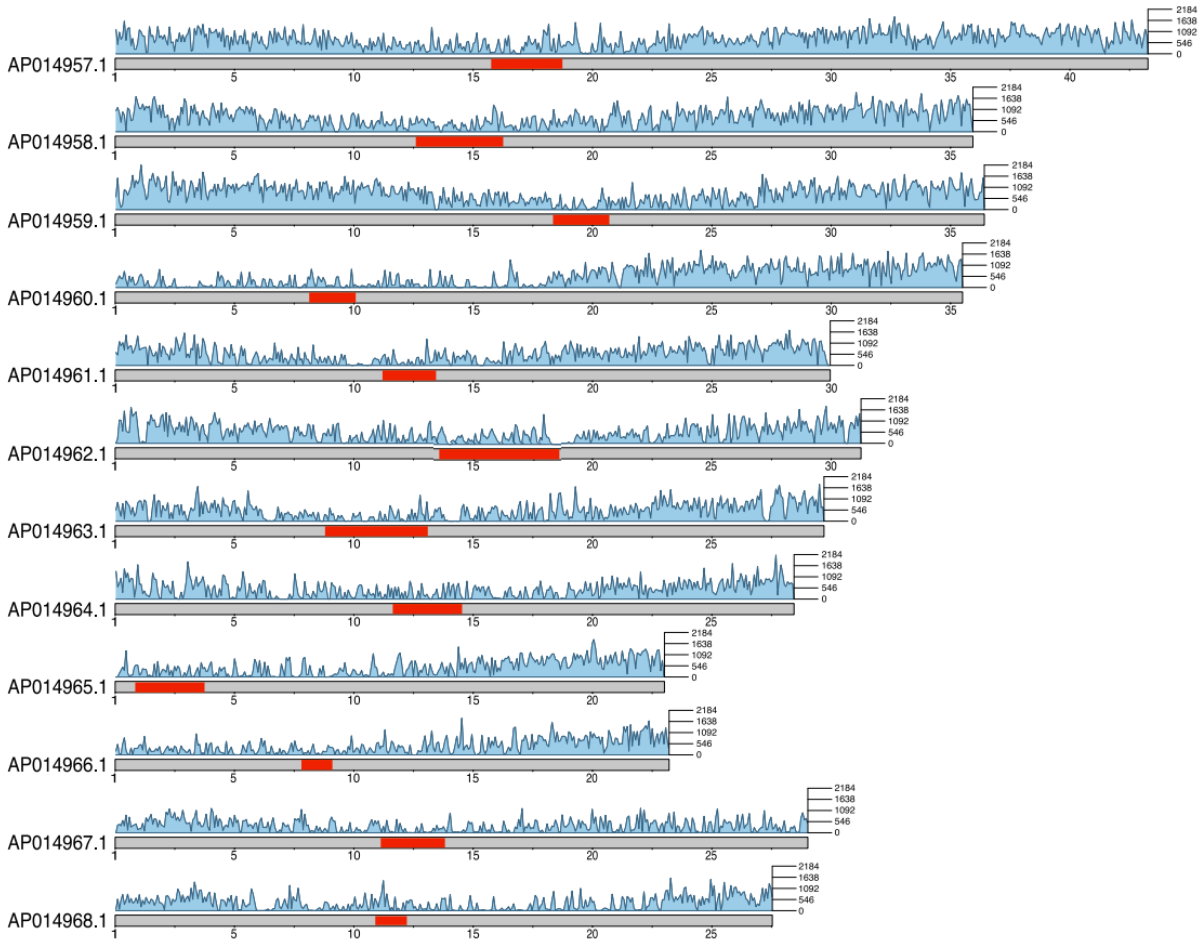
Supplementary Figure 5.8. Location of contig-contig boundaries, and assessment of, the structural variants (SVs) detected in *O. australiensis* KR pseudomolecules. Black lines indicate the boundaries between contigs, which are made up of 100 Ns. The light green bands represent the major SVs (inversions, translocations, and duplications) relative to *O. sativa* Nipponbare. SV coordinates (± 10 Kbp) were manually curated using an interactive dotPlotly plot of the whole-genome alignment between *O. australiensis* and *O. sativa* and the visualised using karyoploteR in the R environment. SVs (± 10 Kbp) that are wholly contained within a contig are likely to be real and not artefacts of incorrect reference-guided contig orientation. SVs shown in red indicate those that cross a contig-contig boundary. The manually curated list of SVs (inset) reports SVs as they occur in numerical order. The scale bars beneath each scaffold show the size of the scaffolds in Mbp.



Supplementary Figure 5.9. Inspection of Chr2 for a previously reported inversion. (A) Alignment between *O. sativa* Nipponbare Chr2:34-35 Mbp and *O. punctata* Chr2 (CM002489.2). This alignment shows an inversion on *O. punctata* Chr2 relative to *O. sativa* Chr2. (B) The alignment between *O. sativa* Nipponbare Chr2 and *O. australiensis* KR Chr2 does not show the canonical non-AA rice Chr2 inversion. (C) The alignment between *O. australiensis* KR Chr2 and *O. punctata* Chr2 shows the same inversion seen in (A). Thus, we do not observe the previously reported Chr2 inversion in the *O. australiensis* KR assembly.



Supplementary Figure 5.10. Insertion times for the major LTR-RT superfamilies identified in the scaffolded *O. australiensis* nuclear genome assembly (including ChrUn) by LTR_Retriever.



Supplementary Figure 5.11. SNVs identified between the assembled *O. australiensis* Keep River genome (blue) and *O. sativa* Nipponbare reference genome (grey rectangles). The *O. sativa* chromosomes appear in numerical order from top to bottom. Red boxes indicated location of centromeric regions on the *O. sativa* chromosomes as reported by Mizuno et al.57. SNVs were detected by mapping *O. australiensis* Keep River short reads to the *O. sativa* Nipponbare genome and using mpileup to call SNVs. Chromosomes appear in order from 1 to 12.

Figure too large to display. Please access the figure at link provided above.
Supplementary Figure 5.12. Short-read coverage plot. *O. australiensis* KR Illumina short reads were mapped to *O. sativa* Nipponbare chromosomes and the resulting indexed BAM file was used by JVarkit's wgscoverageplotter tool to plot per-base coverage. The green line represents the average coverage for each chromosome. *O. australiensis* KR short read coverage is not even across each chromosome, which may affect analyses of SNV density.

scientific reports

Check for updates

OPEN

The first long-read nuclear genome assembly of *Oryza australiensis*, a wild rice from northern Australia

Aaron L. Phillips[1,2], Scott Ferguson[3,4], Nathan S. Watson-Haigh[5,6], Ashley W. Jones[3,4], Justin O. Borevitz[3,4], Rachel A. Burton[1,2] & Brian J. Atwell[7]✉

Oryza australiensis **is a wild rice native to monsoonal northern Australia. The International *Oryza* Map Alignment Project emphasises its significance as the sole representative of the EE genome clade. Assembly of the *O. australiensis* genome has previously been challenging due to its high Long Terminal Repeat (LTR) retrotransposon (RT) content. Oxford Nanopore long reads were combined with Illumina short reads to generate a high-quality ~ 858 Mbp genome assembly within 850 contigs with 46× long read coverage. Reference-guided scaffolding increased genome contiguity, placing 88.2% of contigs into 12 pseudomolecules. After alignment to the *Oryza sativa* cv. Nipponbare genome, we observed several structural variations. PacBio Iso-Seq data were generated for five distinct tissues to improve the functional annotation of 34,587 protein-coding genes and 42,329 transcripts. We also report SNV numbers for three additional *O. australiensis* genotypes based on Illumina re-sequencing. Although genetic similarity reflected geographical separation, the density of SNVs also correlated with our previous report on variations in salinity tolerance. This genome re-confirms the genetic remoteness of the *O. australiensis* lineage within the *O. officinalis* genome complex. Assembly of a high-quality genome for *O. australiensis* provides an important resource for the discovery of critical genes involved in development and stress tolerance.**

Abbreviations

AGP	A golden path
ANU	Australian National University
BAM	Binary Alignment Map
BCF	Binary variant call format
BLAST	Basic local alignment search tool
bp	Base pairs
BUSCO	Benchmarking universal single-copy orthologs
Chr	Chromosome
ChrUn	Chromosome unassigned
DVC(R)	Deputy Vice Chancellor (Research)
EDTA	Ethylenediaminetetraacetic acid
EDTA	Extensive de novo TE annotator
GFF3	General Feature Format 3
GO	Gene Ontology
HPC	High performance computer
HQ	High quality
IGV	Integrative genomics viewer
IOMAP	International *Oryza* Map Alignment Project
Iso-Seq	Isoform sequencing
Kbp	Kilo base pairs
LAI	LTR assembly index

[1]Department of Food Science, University of Adelaide, Adelaide, SA, Australia. [2]ARC Centre of Excellence in Plant Energy Biology, Adelaide, SA, Australia. [3]Research School of Biology, Australian National University, Canberra, ACT, Australia. [4]ARC Centre of Excellence in Plant Energy Biology, Canberra, ACT, Australia. [5]South Australian Genomics Centre, University of Adelaide, Adelaide, SA, Australia. [6]Australian Genome Research Facility, Victorian Comprehensive Cancer Centre, Melbourne, VIC, Australia. [7]School of Natural Sciences, Macquarie University, Sydney, NSW, Australia. ✉email: brian.atwell@mq.edu.au

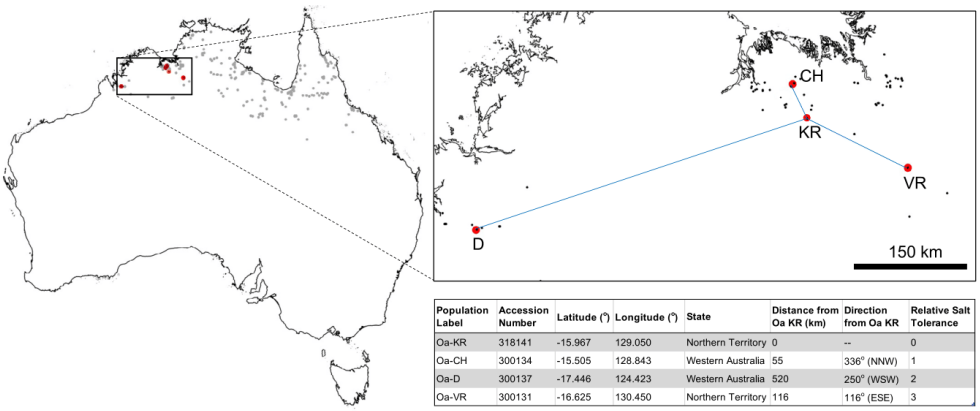
LQ	Low quality
LTR	Long terminal repeat
RT	Retrotransposon
Mbp	Mega base pairs
MYA	Million years ago
NCBI	National Center for Biotechnology Information
NSW	New South Wales
OMAP	*Oryza* Map Alignment Project
ONT	Oxford Nanopore Technologies
PAF	Pairwise mApping format
PEG	Polyethylene glycol
PIF	P Instability factor
PVP	Polyvinylpyrrolidone
RIN	RNA integrity number
RMBlast	RepeatMasker compatible NCBI BLAST suite
SAM	Sequence Alignment Map
SNV	Single nucleotide variation
SV	Structural variation
TIR	Terminal inverted repeat
VCF	Variant call format

The genus *Oryza* is made up of ~ 27 species that have been classified into 11 genome groups[1,2]. Two of these species (*O. sativa* and *Oryza glaberrima*), both of which belong to the AA genome clade, were independently domesticated in Asia and Africa, respectively. Together, these two domestic rice species serve as a staple food source for about 60% of the global human population. The remaining ~ 25 species occur in wild populations and have become specialised to occupy a diverse set of environments around the world[3].

To date, the species and cultivars that make up the AA genome clade have provided most of the genetic variation and resources that have led to the success of domesticated rice[4], while other genome clades have contributed relatively little to domestication. Encouragingly, rich diversity is apparent, even within landraces of just one *O. sativa* subspecies. For example, in a GWAS study for 14 agronomic traits (including tolerance to drought and degree of seed shattering) in 517 *indica* sub-species landraces, Huang et al.[5] revealed ~ 3.6 million SNPs. Currently nine of the 11 *Oryza* genome assemblies available on Ensembl Plants; (accessed 26/2/2021)[6] belong to the AA genome clade, one belongs to the BB genome (*Oryza punctata*), and one to the FF genome (*Oryza brachyantha*). Therefore, wild species of rice occupy a number of genomic clades that are not yet represented by an assembly and undoubtedly include traits that would be desirable in domestic cultivars. The wild species therefore represent vast reservoirs of untapped genetic variation that could be harnessed for domestic rice improvement[2–4,7]. Recently, Hiromi et al.[8] provided an update to the OryzaGenome database which includes short-read sequencing data for several non-AA genome *Oryza* species. The availability of *Oryza* '-omics' resources is increasing, yet the number of whole-genome assemblies for non-AA genome wild species remains low. Whole-genome assemblies of wild *Oryza* species will provide new sources of genes and gene variants/haplotypes, enhancing the ability of breeding programs to introduce desirable traits into domestic cultivars. However care is necessary to navigate issues such as linkage drag and suppressed recombination frequencies that may be associated with use of wild germplasm. Additionally, such projects could lead to the domestication of wild rice populations (e.g., through the targeted modification of domestication-associated traits, such as seed shattering)[9]. Overall, harnessing the genetics of wild *Oryza* species has wide application for cereal crop improvement and specifically for crop resilience. In fact, the *Oryza* Map Alignment Project (OMAP) and the International OMAP (IOMAP) recognised the value that the genome assemblies of wild *Oryza* species would provide to research in abiotic and biotic stress tolerance, and domestication, as long ago as 2003[10–12]. These projects recommended sequencing the genomes of several wild rice species but progress in this respect has been slow.

Recently, several non-AA genetic resources have become available, including genomes within the *O. officinalis* complex[13] and the genome of *Oryza granulata*[14]. *O. granulata,* an upland wild rice species, has a demonstrated ability to tolerate drought stress and bacterial blight, providing a valuable new source of *Oryza* genetic variation. However, owing to its large genome size, high repeat content, and being a short-read assembly, the *O. granulata* genome assembly is highly fragmented, reducing its utility. These same difficulties (i.e., genome size and repeat content) have been encountered during assembly of the EE genome, represented solely by *O. australiensis,* which is reported to be the largest diploid *Oryza* genome.

Oryza australiensis is a perennial wild relative of rice endemic to the tropical regions of northern Australia. Due to the extremely hot and sporadically dry environment in which *O. australiensis* exists, this species has been described as an extremophile[3]. *Oryza australiensis* is the sole member of the EE genome clade and its genome has been estimated to be 965 Mbp in size, more than double that of *O. sativa* ssp. *japonica*[15,16]. The size of the *O. australiensis* genome is the result of the expansion of long terminal repeat (LTR) retrotransposon (RT) families (e.g., *Gypsy* and *Copia*), which comprise an estimated 65% of the genome. In comparison, the genome of *O. sativa* is composed of approximately 10% LTR-RT[15]. It is uncertain what led to the rapid genome-wide expansion of these LTR-RT families. However, it has been noted that the expansion of LTR-RTs facilitated rapid adaptive genome evolution in the *Oryza* genus, leading to speciation events[17]. These same processes may also play a role in the evolution of biotic and abiotic stress tolerances. Heat tolerance has been well-characterised in *O. australiensis*[18], and transformation of domestic rice with a key gene (Rubisco activase) involved in heat tolerance fortifies grain yield during episodes of heat[19]. Further, Yichie et al.[20] identified population-specific



Population Label	Accession Number	Latitude (°)	Longitude (°)	State	Distance from Oa KR (km)	Direction from Oa KR	Relative Salt Tolerance
Oa-KR	318141	-15.967	129.050	Northern Territory	0	--	0
Oa-CH	300134	-15.505	128.843	Western Australia	55	336° (NNW)	1
Oa-D	300137	-17.446	124.423	Western Australia	520	250° (WSW)	2
Oa-VR	300131	-16.625	130.450	Northern Territory	116	116° (ESE)	3

Figure 1. Locations of *O. australiensis* seed collection sites for this study. List of wild rice accessions (obtained from the Australian Grain Genebank) used in this study reported by Yichie et al.[19] to demonstrate varying degrees of salt tolerance. Dots on the map show the occurrence of *O. australiensis* (retrieved from the Atlas of Living Australia). In-set table: accessions are ordered by salt tolerance, with Oa-KR being the most sensitive to salt, and Oa-VR being most tolerant[19].

tolerance to salinity stress in *O. australiensis* accessions, and Hamzelou et al.[21] report on drought tolerance in this species. The unique anatomy and photosynthetic efficiency of *O. australiensis* leaves also make it an important species for understanding how rice photosynthesis can be augmented[22]. In terms of food quality, the grain of *O. australiensis* has unique pigmentation and starch composition, distinct from Asian species[23,24], that could prove to be a profitable food item in and of itself.

A highly-quality reference genome for *O. australiensis* could reveal novel loci for abiotic stress tolerance and help shed light on the genome evolution of the *Oryza* genus. However, since the conception of OMAP in 2003[11], and despite international sequencing efforts since[12], no research group has successfully assembled a high-quality, chromosome-level genome for *O. australiensis*. A previous attempt to assemble the *O. australiensis* genome with short reads resulted in a highly fragmented assembly due to the genome size and high repeat content[25]. Efforts such as those by Joly-Lopez et al.[25] highlight the difficulties involved in assembling highly contiguous plant genomes with significant repeat content[26].

In this study, we report on the assembly of a reference-quality *O. australiensis* nuclear genome from Keep River (KR; Northern Territory, Australia) into 1956 contigs using Oxford Nanopore Technology (ONT) long-reads. The contigs were subsequently scaffolded into 12 pseudomolecules representing 12 chromosomes using the *O. sativa* Nipponbare genome assembly as a guide.

Methods

Growth conditions for plant material used for genome sequencing. Seed of four *O. australiensis* populations was collected from the monsoonal savannah region of northern Australia (Fig. 1), compliant with the relevant national and international guidelines and legislation for wild species collections. All seed was obtained from the Australian Tropical Crops and Forages Collection (AusTRCF) of the Australian Grains Genebank (previously the Australian Plant Genetic Resource Information System) and as part of the Australian Tropical Crops Genetic Resource Centre Collection (ATCGRC). The Keep River accession used for the reference genome assembly bears the accession number AusTRCF 318143 and was collected by Dr I. Cowie at Darwin Herbarium, Northern Territory, Australia (Voucher ID# 9583) on 2nd May 2002 from a drying creek bed in the Keep River National Park after the passing of the summer monsoon. The other three accessions used for re-sequencing (CH, D and VR) are held in the same collection with accession numbers of AusTRCF 300134, 300137 and 300131. These lines are vouchered in the Jeff Corfield Collection in Townsville (Queensland, Australia) under the accession numbers JC 2317, JC 2336 and JC 2312, respectively. CH, D and VR were collected from the sites denoted in Fig. 1 between 1995 and 1997 by Mr I. Watson. Seed was stored at 4 °C until used. Seeds were imbibed in water for one hour and then surface sterilised in a 50% sodium hypochlorite bleach solution for 30 min followed by rinsing. Sterile seeds were sown onto absorbent cotton-lined petri dishes and incubated at 34 °C for 2 days to trigger germination. Germinated seeds were sown into a fine textured krasnozem soil (sourced from Robertson, NSW, Australia) and covered in a thin layer of vermiculite. Plants were grown in glasshouses at Macquarie University (Sydney, Australia) under a 30 °C/22 °C day/night cycle in April–June 2019.

DNA extraction. Roughly 1 g of healthy, fully expanded leaf material was collected from ~ 2-month-old *O. australiensis* plants (one plant per genotype) and ground in a mortar and pestle under liquid nitrogen until a fine

powder was achieved. High molecular-weight DNA was extracted at the Australian National University (ANU; Canberra, Australia) following the method of Jones et al.[27].

Nanopore library preparation, sequencing, and base calling. The sequencing library was prepared as per the manufacturer's (Oxford Nanopore Technologies, Oxford, United Kingdom) instructions for genomic DNA by ligation method (SQK-LSK109) with small modifications. The manufacturer recommends using 0.2 pmol of DNA, however this is optimised for DNA fragments 8 kbp in length. As we were working with substantially longer (~ 20 kbp) DNA fragments, we loaded 3 µg (0.23 pmol) of DNA.

The library was run on FLO-MIN106 (R9.4.1 revD) flowcells as per the manufacturer's instructions. To obtain a good level of coverage of the *O. australiensis KR* genome, the same library preparation for this sample was loaded onto three separate flow cells (i.e., the genomic DNA from one plant was sequenced on three flowcells). The raw signal intensities obtained from the flow cells were used by *guppy* (version 3.0.3) from ONT to perform base calling.

Illumina sequencing. Aliquots (containing 2.24 ng DNA) of the same DNA samples from the four *O. australiensis* genotypes described above were prepared for Illumina short read sequencing. Short-read libraries were created using the Illumina Nextera tagment DNA enzyme (TDE1) according to Jones et al.[28]. Libraries were size selected for 350–600 bp fragments using a PippinHT (Sage Science). Sequencing was performed on an Illumina NovaSeq 6000 S4 flow cell 300 cycles (150 bp paired end), being multiplexed with other projects. Sequencing was performed at the Biomolecular Resource Facility, ANU.

PacBio Iso-Seq library preparation and sequencing. Young grains were harvested during the milky period of development from mature *O. australiensis KR* plants from the same collection as the individual used for ONT and Illumina sequencing (i.e., different plants from the same population; germinated as above and grown in a glasshouse at 30 °C with ambient light at The Plant Accelerator, University of Adelaide, Adelaide, South Australia from February to June 2020). Some developing grain still possessed their pedicel/rachilla and pollen-containing anthers and these were included in the samples to increase the diversity of transcripts in the RNA library.

The growth zones of ten 115-day-old *O. australiensis KR* plants were excised and used for RNA extraction. The growth zone is defined here as ~ 10 mm from the base of the youngest leaves emanating from the central cylinder of the plants. It was necessary to remove surrounding leaf sheathes to access the central cylinder.

A separate cohort of *O. australiensis KR* plants was germinated as above, except germination was achieved with an incubation temperature of 30 °C. Germinated seeds were transferred to a growth rack partially submerged in a ½ Hoaglands solution (pH 5.01) spiked with 25 mg/L ferrous sulphate. After seven days of growth, root tips (5 mm from the end of the root), older seminal roots with young lateral roots, and healthy fully expanded leaves were harvested, snap frozen in liquid nitrogen and stored at − 80 °C. The growth solution was changed once during this time.

An additional cohort of *O. australiensis KR* seed was germinated at 30 °C and submerged in water in the dark to promote coleoptile growth. After six days under these conditions, etiolated coleoptiles were harvested as above.

Total RNA was extracted from the six tissues (7-day-old leaves, 7-day-old root tips, 7-day-old mature root sections, 6-day-old hypoxic coleoptiles, reproductive tissue, and 118-day-old growth zones) with the Sigma Spectrum Plant Total RNA Kit according to the manufacturer's instructions, with minor modifications for the developing grain sample as per Betts et al.[29]. RNA samples were dried in RNA Stable tubes as per the manufacturer's instructions and shipped to DNA Link (Seoul, South Korea). RNA quality and concentration of the samples were checked with a Nanodrop and with a Bioanalyzer 2100 instrument (Agilent, Santa Clara, USA). Only low-quality RNA (RIN < 7) could be extracted from the older seminal roots so this sample was omitted from further analyses.

RNA samples were barcoded upon conversion to cDNA libraries (to facilitate multiplexing) following the manufacturer's instructions of the SMARTer PCR cDNA Synthesis kit. cDNA libraries were amplified by PCR and then an equal concentration of cDNA from each was pooled into a single library. The library was prepared for SMRTbell sequencing with the Express Template Prep 2.0 kit. Iso-Seq sequencing was carried out on a Sequel II PacBio instrument using a single SMRT cell.

Iso-Seq reads were processed through the IsoSeq v3 pipeline to generate high quality polished isoforms. Briefly, raw data is converted from multiple subreads per ZMW into a single consensus CCS read (minimum read quality of 0.9), demultiplexed and clustered (minimum predicted accuracy of 0.99). The aim of the clustering step is to collapse the CCS reads derived from different isoforms of the same gene or from the same transcript isoform but which differ slightly in their 5′ and 3′ ends, typically as a result of RNA degradation.

Nuclear genome assembly and assessment. A summary of all reads obtained during this study can be found in Table 1. *Jellyfish* (RRID:SCR_005491)[30] and *genomescope2* (RRID:SCR_017014)[31,32] were used to estimate the size, unique sequence content, and heterozygosity of the *O. australiensis* KR genome using 17-mers in the KR Illumina short read libraries. Raw MinION reads for *O. australiensis KR* were corrected and then de novo assembled using *Canu* v1.9 (RRID:SCR_015880)[33] as follows: canu -p oryza -d canu genomeSize = 965 m -nanopore-raw all.fastq.gz -gridOptions = "–time = 72:00:00" -obtovlThreads = 24 -batMemory = 9. The MinION long reads and the Illumina short reads were mapped to the resulting contigs using *minimap2* v2.17 (RRID:SCR_018550)[34]. Alignment files were used to polish the genome assembly with *HyPo* v1.0.3[35]. *JVarkit* was used to visualise long-read coverage of the contigs[36]. Contigs were checked for microbial contamination by querying BLAST-generated taxon ID numbers with the lineage command of *taxonkit* v0.6.0[37]. Contigs found to contain

Purpose	Sequencing platform	Sample	No. reads (Millions)	Mean Read Length (kbp)	Read N50 (kbp)	No. bases (Gbp)	Coverage (x)
Genome assembly and polishing	MinION FLO-MIN106 R9.4.1 revD	*O. australiensis*—KR 1	0.45	26.24	40.31	12	12
		O. australiensis—KR 2	0.94	23.22	35.14	22	23
		O. australiensis—KR 3	0.62	24.95	37.38	15	16
Genome polishing	Illumina NovaSeq	*O. australiensis*—KR	138.81	NA	NA	19	21
Genetic similarity	Illumina NovaSeq	*O. australiensis*—CH (300134)	143.27	NA	NA	20	22
		O. australiensis—D (300137)	87.27	NA	NA	12	13
		O. australiensis—VR (300131)	139.67	NA	NA	19	21
Genome annotation	PacBio Sequel II	*O. australiensis*—KR Leaf	0.24	2.69	NA	0.65	NA
		O. australiensis—KR Coleoptile	0.18	2.29	NA	0.42	NA
		O. australiensis—KR Root Tip	0.3	2.16	NA	0.64	NA
		O. australiensis—KR Growth Zone	0.3	2.75	NA	0.84	NA
		O. australiensis—KR Reproductive Tissue	0.24	2.46	NA	0.59	NA

Table 1. Summary of reads used for each step of genome assembly, polishing, variant detection and annotation. See Fig. 1 for details on the accessions. KR 1, KR 2, and KR 3 are the reads obtained from a single genomic DNA preparation sequenced on three different flowcells. These reads were derived from the same *O. australiensis KR* plant and were used for the assembly of the reference *O. australiensis KR* genome. The same DNA preparation was used for the KR Illumina NovaSeq library preparation. CH, D, and VR (accession numbers appear in parentheses) refer to different accessions of *O. australiensis* that have been shown to vary in their tolerance to salt[20]. Reads from these accessions were used to estimate genetic similarity between the genotypes. Multiple *O. australiensis* KR plants were used for RNA extraction for Iso-Seq analysis.

non-eukaryotic sequences were interrogated using the online genome browser IGV-Web (RRID:SCR_011793) by mapping long reads back to the assembly and checking the coverage and overlap of long reads with non-suspect sequences[38]. *Purge Haplotigs* (v1.1.2) was used to determine whether the wild rice genome assembly contained any haplotigs or junk sequences[39].

The quality of the genome assembly was assessed by analysing the quality of the LTR-RT elements of the *O. australiensis KR* genome. LTR-RT elements are difficult to assemble. As such, LTR-RT assemblies provide a proxy for the quality of the rest of the genome. LTR regions were identified with *LTR_Finder_Parallel* v1.1 (RRID:SCR_018969)[40] and *LTR_Harvest*[41]. The outputs were used as input for LTR_Retriever v2.8.7 (RRID:SCR_017623)—to generate LTR Assembly Index (LAI) scores[42]. The quality of the assembly was also assessed with *BUSCO* v4 (RRID:SCR_015008)[43] using the *Poales* database (poales_odb10.2019-11-20; contains 4896 sequences) as the lineage. Contigs were mapped to the *O. sativa* Nipponbare reference genome (GCA_001433935.1) with *minimap2* and visualised with *MashMap* v2.0[44].

Pseudomolecules for the assembly were generated using *RagTag* v1.0.0[45] using version 7 of the *O. sativa* Nipponbare reference genome[46] as a guide for the ordering and orientation of contigs into pseudomolecules (Supplementary File S1a). Contigs that could not be assigned to a pseudomolecule were concatenated (with 100 Ns at contig boundaries) into a 'chromosome' called ChrUn, though we used the unassigned contigs themselves for most of the downstream analyses. To understand why the unassigned contigs could not be placed into a pseudomolecule, the identity of a random subset of the unassigned contigs was determined by BLAST searches and the average LTR content of ChrUn was determined by *LTR_Retriever*. Contigs flagged as repeats, haplotigs or junk by *Purge Haplotigs* were kept in the assembly for completeness but were not included in scaffolding (see above). *JVarkit*'s *WGSplotter* was used to visualise long-read genome coverage[36]. BLAST was used to determine the identity of a random subset of the unassigned contigs. Subsequent annotation of the unassigned contigs (see below) also helped in determining the identity of the sequences that could not be assembled into pseudomolecules. The quality of the scaffolded assembly was assessed again with *BUSCO* and LAI scores, as above. *LTR_Retriever* was also used to produce estimates for LTR-RT insertion times. The quality of the assembly was further checked by mapping ONT long reads to the pseudomolecules to obtain a read mapping rate and to estimate read coverage across the assembly.

Comparison of *O. australiensis* KR assembly to the *O. sativa* Nipponbare assembly. The *EDTA*-masked (see below) pseudomolecules were aligned to the Nipponbare genome using *minimap2* with the 'asm20' setting. The mapping file was used as input for *dotPlotly*[47] to generate a dot plot of the whole-genome alignment with the following settings: "-m 2000 -q 500000 -k 12 -l -p 12". The resulting plot was used to interrogate suspected structural variations (SVs) in the wild rice genome. We used *O. sativa* vs *O. punctata* Chr2 (CM002489.2; a BB genome rice), *O. sativa* vs *O. australiensis* KR Chr2, and *O. australiensis* KR vs *O. punctata* Chr2 alignments to inspect a region on Chr2 in these species which has been reported as an inversion in non-AA

rice genome[1]. Alignments were made as described above, except "-m 5000" was used for *dotPlotly* visualisation. To check whether reference bias in the *O. australiensis* scaffolds was affecting the Chr2 SVs, we scaffolded the KEEP contigs using the *O. punctata* reference genome (AVCL0200000). We aligned the *O. punctata*-guided *O. australiensis* Chr2 to *O. punctata* Chr2, and then to *O. sativa* Nipponbare Chr2 as above.

To check whether the SVs were real, contig coordinates within the scaffolds were extracted from the *RagTag* AGP file and these were used to build a 'cytoband' file for *karyoploteR*[48]. The coordinates for the SVs (± 10 kbp) identified in the *O. australiensis KR* and *O. sativa* genome alignment were added to the 'cytoband' file, and these were used to visualise the boundaries between contigs and SV events using *karyoploteR*. Further, long reads were mapped to the *O. australiensis* KR scaffolds to check the level of coverage across SV junctions. SV junctions with low coverage were coloured red to identify them as suspicious.

To identify single nucleotide variations (SNVs) in the wild rice genome, *O. australiensis KR* Illumina short reads were mapped to the Nipponbare genome with *minimap2*. The resulting BAM file was used as input for the *SAMtools* (RRID:SCR_002105)[49] sub-command *mpileup,* and *BCFtools* (RRID:SCR_005227) called SNV variants. *VCFtools* (RRID:SCR_001235)[50] filtered low quality SNVs with low-read support (minQ = 30; minDP = 3). High quality SNVs were visualised using *karyoploteR* in RStudio with a window size of 70 kbp.

Genome repeat annotation. *LTR_Retriever* (see above) was used to provide an estimate of LTR-RT elements in the *O. australiensis* KR contigs and scaffolds. Repetitive DNA elements were annotated using the Extensive de novo TE Annotator (*EDTA;* Supplementary File S2b)[51]. A manually curated rice repeat-element library provided by the developers of *EDTA* was used to identify repeats in the *O. australiensis KR* genome. A library containing *O. sativa* ssp. *japonica* transcript sequences including untranslated regions (UTRs) was downloaded from the IOMAP Genomes Database[52] and used by *EDTA* to mask endogenous genes from its annotation pipelines. *EDTA* also identifies repeat sequences in the input genome that were not part of the repeat library used for interrogation and then constructs a repeat library specific to the input assembly. Total repeat content was also estimated and annotated using the *RepeatMasker*-based *Repeat Masking* tool included in OmicsBox version 2.0.10[53,54]. The custom *O. australiensis KR*-specific repeat library generated by *EDTA* was used as input for the RMBlast search engine, while other settings were left as default.

Protein-coding gene annotation. Evidence-based gene-finding was carried out with the *OmicsBox* Gene Finding tool with *O. sativa* set as the most closely related species[53,54]. A protein evidence file containing protein sequences from 12 species/accessions of rice downloaded from the IOMAP Genomes Database[52] and the Iso-Seq isoforms were used as evidence for gene finding. Genes were searched on both strands of the pseudomolecules and the unassigned contigs and alternative splice variants were allowed. The resulting gene coordinates were used as input for the built-in *Functional Gene Annotation* workflow. This workflow uses *RepeatMasker* to mask repeat elements in the genome assembly before carrying out gene prediction with *Augustus*. Then, both *Blast2GO* and *InterProScan* pipelines were applied to search for gene and protein matches for the identified genes. Annotations were further supplemented by searching orthologues of the identified genes/proteins using *EggNOG-mapper* in *OmicsBox*. Annotations for the contigs and scaffolds were exported as GFF3 files (Supplementary Files S3c, S4d). To check whether the GFF3 files contained TE gene annotations, we used a list of known plant TE families from the Atlas of Plant Transposable Elements[55].

***O. australiensis* genotype diversity and genetic similarity.** To identify SNVs between four genotypes of *O. australiensis* with demonstrated differences in salt tolerance[20], the *O. australiensis* CH, D, and VR Illumina short reads (Table 1) were mapped to the scaffolded *O. australiensis KR* genome assembly using *minimap2* and visualised as above. Keep River short reads were also mapped back against the Keep River genome to provide a baseline of SNVs. *Mpileup* was used to call variants between the *O. australiensis* KR genome assembly and the mapped reads. VCF files were filtered using bcftools such that only reads with a mapping quality > 30, and only SNVs with a depth of 5× were considered. The patterns of SNV density described here were visually inspected.

Genetic similarity between the *O. australiensis* genotypes was determined using the Illumina short reads for each genotype as input for *kWIP*[56]. *kWIP* uses a Weighted Inner Product of different k-mer hash values derived from *khmer* (RRID:SCR_001156)[57] to reduce the effect that erroneous read data has on estimates of genetic similarity. The estimate of genetic similarity was also extended to the *Oryza* genus using publicly available Illumina short read data derived from genomic DNA samples (see Supplementary File S5e for the list of accessions).

Results

Genome assembly, making pseudomolecules and genome quality. To generate long-read sequence data for assembly, DNA from a single *O. australiensis* Keep River (KR) plant was extracted, purified, and sequenced on three ONT MinIOn flow cells (KR1, KR2, and KR3). Additionally, the long-read DNA extract was used to generate Illumina NovaSeq short read data for *O. australiensis* KR. DNA extraction and Illumina short-read sequencing was also performed on *O. australiensis* genotypes CH, D, and VR (Fig. 1, Table 1). We generated 49 Gbp (2.01 million reads) of ONT MinION long read data and 19 Gb (138.1 million reads) of Illumina short read data for the *O. australiensis* Keep River accession. Further, we generated 3.14 Gbp (1.25 million reads) of PacBio Iso-Seq data from five distinct tissue types (Table 1, Supplementary Fig. S1). Finally, we generated 20 Gbp, 12 Gbp, and 19 Gbp of Illumina short read data for three additional *O. australiensis* genotypes (CH, D, VR, respectively; Table 1).

K-mer analysis of *O. australiensis* KR short reads by *jellyfish* and *genomescope2* estimated the genome size to be ~ 822 Mbp. About 41.9% of the k-mers were unique, which suggested that the genome contained 58.1% repeat

Stage	Feature (unit)	Value
Long reads	Post-QC bases (Gbp)	49.2
	Estimated coverage (x)	51
	Expected genome size (Mbp)	965
Genome assembly (contigs)	Assembled genome size (Mbp)	996
	Long read coverage (x)	38
	No. contigs	1956
	Contig N50 (Mbp)	1.9
	Contig L50	114
	Contig N90 (kbp)	186.7
	Contig L90	799
	BUSCO score (100% = 4896)	91.9
	Whole-genome LAI	15.2
Genome scaffolds	Scaffolded genome size (Mbp)	860.9
	Long read coverage (x)	46
	Placed contigs	693
	Unplaced contigs	157
	Total length of placed contigs (including Ns; Mbp)	812
	Total length of unplaced contigs (Mbp)	46.9
	Gaps (Mbp)	2.1
	BUSCO score (100% = 4896)	97.5
	Whole-genome LAI score	17.6
LTR_Retriever (scaffolds)	LTR-RT content (Mbp)	518.2
	LTR-RT content (%)	60.2

Table 2. Genome assembly statistics for *O. australiensis* keep River.

content. Heterozygosity was estimated to be 0.001%. Based on these results, we did not expect heterozygosity to have greatly affected contig assembly.

Canu assembled a genome containing 980 Mbp across 1956 contigs (Table 2). This represents 101.5% of the expected *O. australiensis* genome size. The contig N50 was 1.9 Mbp, the largest contig was 24.9 Mbp in length, while the smallest contig was 1324 bp in length. After polishing with *HyPo* the size of the assembly increased to 996 Mbp (103.2% of the expected genome size). We achieved an average coverage of 38 × and the contigs aligned well with the *O. sativa* Nipponbare reference genome (Supplementary Figs. S2, S3). *BUSCO* and LAI scores of 91.9% (complete) and 15.2, respectively, indicate that we have assembled a highly complete contigs. We identified 93 haplotigs, 124 junk contigs, 889 repeat contigs, and 850 primary contigs using *Purge Haplotigs* (Table 2). Thus, after purging the assembly of all but the primary contigs, the final assembly size was 859 Mbp (*cv*. 965 Mbp estimated by flow cytometry[15], and 822 Mbp estimated by k-mer counting—see above).

Scaffolding resulted in the placement of 693 of the 850 primary contigs into one of 12 pseudomolecules and increased assembly size to ~ 861 Mbp. Pseudomolecules contained 81.5% of all assembled contigs (~ 812 Mbp of 996 Mbp) and long read coverage increased to 46×(Table 2, Supplementary Fig. S4, Supplementary File S1a). The remaining 157 contigs (46.9 Mbp) could not be placed into a pseudomolecule. Analysis of the unassigned contigs by *LTR_Retriever* showed that these sequences were largely made up of repeat sequences (~ 80% LTR-RT content, versus ~ 60% in the pseudomolecules), explaining why they could not be placed in the scaffolds. After scaffolding with KEEP contigs only, the *BUSCO* completeness score of the assembly was 97.5% and the LAI was 17.6. To further assess the quality of our scaffolded genome, we mapped ONT long-reads to the pseudomolecules, where 99.87% of ONT long-reads aligned to our genome with high quality (average MAPQ = 47).

After aligning our assembly long-reads (i.e., KR1, KR2, and KR3) to our assembled genome we observed several positions in the *O. australiensis* KR assembly that have low long-read coverage (< 20x, *cv*. 46× scaffold coverage; Supplementary Fig. S4). These regions predominantly occur on chromosome 4 and coincide with regions that are composed of smaller contigs (Supplementary Fig. S5). Additionally, chromosome 4 was constructed from a larger number of contigs than all other pseudomolecules (76 contigs compared to an average of 56 ± 9 contigs per pseudomolecule; Supplementary Fig. S5).

Structural variations (SVs). We aligned the *O. australiensis* KR assembly to the *O. sativa* Nipponbare assembly and observed a high level of synteny between the two genomes (Fig. 2). We also observed 378 instances of non-syntenic alignment (e.g., sequence that occurs on *O. sativa* chromosome 1 aligning to *O. australiensis* KR pseudomolecule 3). Thus, the genome–genome alignment indicated the presence of SVs. We investigated 21 SVs further (8 translocations, 7 inverted translocations, 5 inversions, and 1 duplication). The largest SV detected was an inversion that occurred on Chr11 and was 2.95 Mbp long within the *O. sativa* Nipponbare genome and 7.2 Mbp long in the *O. australiensis* KR genome assembly. We show that some of the SVs are unlikely to be caused by scaffolding errors (Supplementary Fig. S5). Seven of the 21 reported SVs cross contig–contig boundaries.

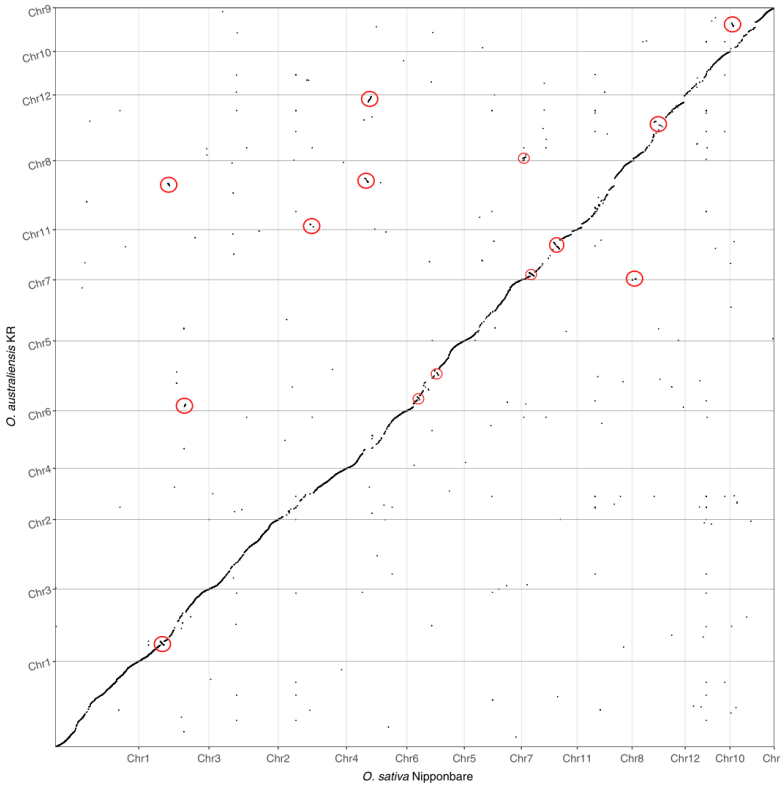


Figure 2. Alignment of scaffolded *O. australiensis* KR contigs (named Chr1–12) to the *O. sativa* Nipponbare reference genome. The wild rice genome was aligned to the domestic rice genome using *minimap2* and visualised using *dotplotly*. The 21 SVs that were investigated further are circled in red (some circles contain multiple SVs). Note: chromosomes do not appear in numerical order—*dotPlotly* orders the target sequence (*O. sativa*, here) by chromosome size; ChrUn is not included as it did not contain any large alignments.

Further, by mapping the *O. australiensis* KR long reads to the pseudomolecules and observing continuous mapping of long reads at a depth of 31–75× coverage across SV junctions, we show that the investigated SVs are likely not artefacts of errors in the assembly of the contigs. All reported SVs require further investigation before they can be confirmed.

Recently, Stein et al.[1] showed an inversion on chromosome two of non-AA rice genomes relative to AA genome rice. This inversion was not detected in the assembly presented here (Fig. 2). We verified the inversion reported by Stein et al.[1] and the lack of inversion in the *O. australiensis* KR assembly by aligning *O. sativa* Nipponbare Chr2 with *O. punctata* Chr2 (Supplementary Fig. S6A) and *O. australiensis* KR Chr2 (Supplementary Fig. S6B), respectively. When we aligned *O. australiensis* KR Chr2 to *O. punctata* Chr2 we saw an inversion similar to the *O. sativa* Nipponbare vs *O. punctata* Chr2 alignment (Supplementary Fig. S6C). To check the validity of this finding, we mapped *O. australiensis* KR long reads back to the assembly and observed 38–43× coverage at the alleged SV junction. Further, the coordinates of the alleged SV (Chr2:74,761,274–74,941,712) in the *O. australiensis* KR assembly, derived from the alignment of *O. australiensis* KR Chr2 to *O. punctata* Chr2 (Supplementary Fig. S6C), sit well within the boundaries of a contig (tig00003125 occupies Chr2: 69,779,629–76,584,758). After aligning *O. punctata*-guided *O. australiensis* Chr2 to *O. punctata* Chr2 we again observed an inversion at these coordinates, but no inversion when *O. punctata*-guided *O. australiensis* Chr2 and *O. sativa* Nipponbare Chr2 (data not shown).

Transcripts and genome annotation. Using a single PacBio Sequel II 8M SMRT Cell we were able to generate 3.3 million polymerase reads. Once these data were collapsed to Circular Consensus Sequencing (CCS) reads and demultiplexed, we had ~183–305 k CCS reads per tissue. Once the demultiplexed CCS reads were

Tissue	# CCS reads	Mean length (kbp)	# Polished isoforms	
			High quality	Low quality
Leaf	242,841	2.69	25,805	6
Coleoptile	182,785	2.29	21,013	2
Root tip	297,806	2.16	33,619	3
Growth zone	304,934	2.75	35,048	12
Developing grain	239,213	2.46	23,176	6

Table 3. Summary statistics of the Pac Bio Iso-Seq data.

Class	Superfamily	Count	Masked (Mbp)	Masked (%)
LTR-RT	*Copia*	168,433	97.5	11.4
	Gypsy	299,854	319.2	37.2
	Unknown	63,847	39.1	4.6
Non-LTR-RT	–	365,794	167.2	19.5
Total	–	897,928	623	72.7

Table 4. Repeat elements in the *O. australiensis* KR scaffolds[51].

clustered we had ~21–35 k polished high quality (HQ) isoforms per tissue (Table 3). *BUSCO* analysis indicated that the Iso-Seq transcripts from the five selected tissues accounted for ~ 40% of the 4896 sequences in the *Poales* database. Only the HQ transcripts were used for annotation (low quality transcripts were discarded). A number of transcript isoforms were detected (see Supplementary Fig. S1B–D for examples of transcript isoforms for genes involved in stress response in rice).

Following the production of pseudomolecules via reference-guided scaffolding, *LTR_Retriever* identified 608.7 Mbp (61.6%) LTR-RT content in the assembly (assigned and unassigned contigs inclusive). *EDTA* identified 730.1 Mbp of repeat sequences in the pseudomolecules (Table 4, Supplementary File S2b), including 53.07% LTR-RT sequences. A peak in insertion events for both *Gypsy* and *Copia* LTR-RT elements was observed at ~ 0.5 million years ago (MYA; Supplementary Fig. S7). *Gypsy* repeat elements are more abundant than *Copia* elements in the *O. australiensis KR* genome. A third 'unknown' LTR-RT superfamily shows a slow increase in copy number over the last four million years (Supplementary Fig. S7).

Using the HQ Iso-Seq transcripts and protein sequences from the *Oryza* genus as evidence, *OmicsBox* identified 52,090 genes and 60,325 transcripts in the assembly (Supplementary Files S3c, S4d). *OmicsBox* functionally annotated 35,620 genes and 43,362 transcripts. To check whether our annotation included TE genes, we searched the gene transcripts for known plant TE families[55]. Of our annotated genes 1033 sequences (829 *Gypsy* LTR-RTs, 81 *Copia* LTR-RTs, 4 unknown LINEs, 105 CACTA TIRs, 4 MuDR TIRs, 8 Pong TIRs, and 2 unknown non-LTR-RT repeat elements) were found to be TE genes and removed from the *O. australiensis* gene annotation. After accounting for the presence of TE genes we were left with 51,057 predicted genes, 34,587 functionally annotated genes, and 42,329 functionally annotated transcripts.

Genotypic diversity and genetic similarity across four *O. australiensis* accessions and within the genus Oryza. After mapping the *O. australiensis* KR short reads to the *O. australiensis* KR assembly, we detected ~ 2 million SNVs. In comparison, we detected ~ 2.9 million, ~ 5 million, and ~ 6.4 million SNVs when mapping the CH, D, and VR accession short reads to the *O. australiensis* KR assembly. In the samples we analysed, genome wide SNV number did not increase with increasing geographic distance (Fig. 1). On the other hand, Fig. 3 shows that the CH accession is most closely related to KR, followed by VR, and then D—that is, genetic distance (as determined by *kWIP*) increases with increasing geographic distance (Figs. 1, 3).

Genetic distance between 15 species/accessions (representing the AA, BB, CC, BBCC, CCDD, EE, FF, GG, HHJJ, and HHKK genomes) within the *Oryza* genus was estimated with *kWIP*. A full list of the short read accessions used can be found in Supplementary File S5e. The genetic distances estimated by *kWIP* allowed for the reconstruction of the *Oryza* phylogeny except for two 'misplaced' samples (10. *O. ridleyi*, and 13. *O. australiensis*; Fig. 3). We were able to re-construct the canonical *O. sativa* (AA) and the *O. officinalis* (BB, CC, CCDD, EE) genome complexes, and the remaining *Oryza* genomes (FF, GG, HHJJ, HHKK) form a third complex. The *O. australiensis* lineage diverges from the rest of the *O. officinalis* complex members (Fig. 3).

Mapping the *O. australiensis* KR short reads to the *O. sativa* Nipponbare reference genome revealed ~ 2.9 million SNVs between *O. australiensis KR* and *O. sativa* Nipponbare (Supplementary Fig. S8). There are regions on all *O. sativa* chromosomes that, from visual inspection, have low SNV density. Also, the density of SNVs appears to increase at both terminal ends of the *O. sativa* chromosomes, though this may be due to uneven short read mapping density across the *O. sativa* Nipponbare genome (Supplementary Fig. S9).

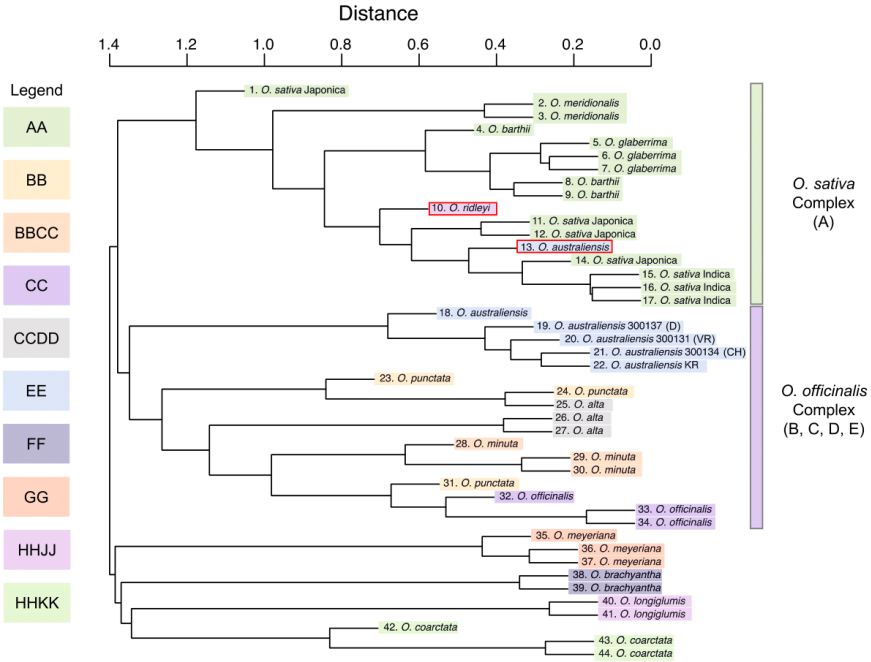


Figure 3. Genetic distance between *Oryza* species/accessions derived from Illumina short read libraries by *kWIP*. Samples are coloured based on the *Oryza* genome clade that they occupy (see inset Legend). *kWIP* clustered samples into the canonical *O. sativa* (AA genome) and *O. officinalis* (BB, CC, BBCC, CCDD, EE genomes) complexes. The *O. australiensis* lineage is divergent from the rest of the *O. officinalis* genome complex, suggesting it harbours lineage specific adaptations that could be explored for stress tolerance. Within the *O. australiensis* lineage, the four genotype samples re-sequenced in the present study (KR, CH, VR, and D) show genetic distances that are correlated with geographic distance. The sample divergence between KR, CH, VR, and D, as well as the other *O. australiensis* accessions shown here suggests that there may be within-species genomic variation that can also be explored for tolerance to stresses (e.g., genotype-specific tolerance to salt stress). Samples that appear with red borders are suspicious as they did not cluster with their corresponding genome clade. This may be due to errors in the sequencing files that were not corrected prior to running *kWIP* or may reflect human error during sample preparation for sequencing.

Discussion

In this study we report on the first long-read assembly of an *O. australiensis* nuclear genome. Our results indicate that the assembled genome is of high quality and completeness. The assembled pseudomolecules (excluding ChrUn) are, on average, 2.2 times larger than the *O. sativa* Nipponbare reference chromosomes, which is in line with previous observations that *O. australiensis* chromosomes are larger than other *Oryza* species when seen during mitosis[58]. Our scaffolded *O. australiensis* genome assembly (858 Mbp) is smaller than estimated by flow cytometry (965 Mbp[16,58]), however our k-mer based estimation of genome size using *O. australiensis* short reads predicted a genome of 822 Mbp. The discrepancy in size between the genome reported here and the estimated genome size for *O. australiensis*[15,16] may come about due to irresolvable LTR-RT elements.

After aligning the ONT long reads used to assemble the *O. australiensis* KR contigs to the reference-guided pseudomolecules, we observed unequal coverage distribution on some regions of the pseudomolecules (Supplementary Fig. S4). Most often, coverage was halved in these regions (for example, see Chr4). Having purged the assembly of haplotigs it is unlikely that low coverage results from the presence of homologous maternal and paternal sequences being present in the assembly. Instead, it appears that the low coverage came about due to the relatively high number of small contigs that make up some parts of the pseudomolecules. For example, Chr4, which has the lowest median long read coverage, is composed of the most contigs compared to all other pseudomolecules (apart from ChrUn; Supplementary Fig. S5). The fragmented nature of the contigs that have been used to build these regions of Chr4, and the other pseudomolecules, is characteristic of repetitive genome regions that are typically difficult to assemble. To overcome the limitations imposed by highly repetitive genome

regions and, thus, to resolve the sequences therein, it is recommended that future studies employ long reads sufficient for a higher genome coverage.

We found that ~ 74% of the assembly is repeat sequence, of which ~ 55.5–61.6% was identified as LTR-RTs. These estimates are in line with previous estimates (~ 65%) of *O. australiensis* genome repeat content[15]. Stein et al.[1] report a maximum total repeat content of 49.6% for *O. punctata* (BB genome), and a minimum of 27.2% for *O. meridionalis* (AA genome), thus *O. australiensis* has the highest total repeat content among the few *Oryza* species sequenced to date. We show that there was a relatively recent (~ 0.5 MYA) increase in LTR-RT (*Gypsy* and *Copia*) composition of the *O. australiensis KR* genome, in line with previous observations of *O. australiensis* Bacterial Artificial Chromosome (BAC) library sequences[15]. Similar trends of expansion for these LTR-RT types have been documented for both *O. sativa* indica and *O. sativa* ssp. *japonica* varieties[1]. However, along with a recent expansion of these LTR-RTs, the *O. sativa* genomes have undergone subsequent substantial gene loss, which has played a role in determining the genome size[59]. We also report on the expansion of an unknown LTR-RT family that appears to be in progress in modern-day *O. australiensis* plants, similar to the two *O. sativa* subspecies mentioned above and other *O. australiensis* accessions[15]. Piegu et al.[15] showed that the LTR-RT element called '*Kangourou*', a low copy number LTR-RT relative to *Gypsy* or *Copia*, had a peak burst time of ~ 2.3 MYA. The magnitude of the copy number burst in the *O. australiensis* genome of *Kangourou* in previous studies using BAC libraries[15] matches that of the 'unknown' LTR-RT element we report here, however our results suggest a much more recent (~ 0.2 MYA) but slower increase in copy number. Other species of wild rice appear to have more ancient LTR-RT family (*Gypsy*, *Copia*, and unknown) expansions. For example, the *Gypsy*, *Copia*, and unknown LTR-RT families of *O. brachyantha* peaked in abundance around 1.8, 2, and 2.5 MYA, respectively[1]. Shenton et al.[13] show a burst in the *Mutator* MuDR2 family composition of the *O. officinalis* genome complex. Further, *O. meridionalis* seems to have more dated (~ 1–1.5 MYA) expansions and then contractions of the same LTR-RTs relative to *O. australiensis* KR[1]. However, given the differences in the assembly methods and qualities of the various *Oryza* genomes, it is difficult to make direct comparisons of LTR-RT family ages and so these speculations should be viewed with caution.

After accounting for TE genes in the annotation file, we predicted 51,057 genes and functionally annotated 34,587 of those genes. This is in line with previously reported *Oryza* genome annotations. For example, Stein et al.[1] report a maximum of 38,550 annotated loci for O. *sativa* ssp. *japonica*, and a minimum of 24,208 for *O. brachyantha*. In comparison, Release 7 of the *O. sativa* Nipponbare genome contains 39,045 non-TE annotated genes and 49,066 non-TE gene models[46]. Thus, the number of genes/transcripts annotated in this version of the *O. australiensis KR* genome assembly is comparable to that of the *O. sativa* Nipponbare version 7 genome. This is consistent with the hypothesis that the size of the *O. australiensis* genome is the result of the expansion of LTR-RT elements, rather than gene duplication events[15]. However, similar numbers of coding genes invite speculation about a functional role for expansion of LTR-RT elements. For example, it is conceivable that the especially adverse environments where *O. australiensis* evolved in northern Australia have selected for expansion of particular gene families involved in tolerance to stress.

Comparing the *O. australiensis* KR genome assembly to the *O. sativa* Nipponbare assembly, we identified a number of large structural variations. Recently, Stein et al.[1] showed an inversion on chromosome two of non-AA rice genomes relative to AA genome rice. The same inversion appears in *Leersia perrieri* and *Brachypodium distachyon* and so it may be a conserved feature in grass genomes, except in rice with the AA genome[1]. This inversion was not detected by reference-guided construction of the *O. australiensis* assembly (Fig. 2), suggesting a need for deeper coverage of this region in future studies. However, we can report 36–45× coverage at the putative inversion junction in the *O. australiensis* KR assembly, and we show that the putative SV sits well within the boundaries of a single contig. Hence, we suggest that this region of the *O. australiensis* KR assembly is not the result of misassembly. The possibility remains that it is an artefact of reference-biased scaffolding, however we found that the inversion is missing in the *O. australiensis* assembly even after using a BB rice genome (*O. punctata*) to order and orient the contigs (data not shown), so this is unlikely. The SV reported by Stein et al.[1] in non-AA rice genomes may truly be absent from the *O. australiensis* KR genome. One hypothesis for its absence is hybridisation between the co-occurring AA-genome species (*O. meridionalis* and *O. rufipogon*) and *O. australiensis*, thus eliminating the Chr2 SV in the *O. australiensis* KR genome. However, given the difficulty in producing viable progeny from artificial crosses between the AA genome species and *O. australiensis*[60], we find this to be an improbable explanation for the observed results. Thus, in the absence of further evidence, all structural variations should be viewed with caution.

Given that the 21 large SVs we observed do not appear to be the result of either scaffolding or assembly artefacts, if these SVs prove to be real, they highlight considerable differences in the structure of the *O. australiensis KR* assembly and *O. sativa*. These SVs, and the substantial repeat content of the wild rice genome, would offer insight into why crossing the AA genome *O. sativa* and the EE genome *O. australiensis* is difficult and why successful crosses have abnormal chromosomal arrangements and the resulting progeny are sterile[61,62]. To validate the putative SVs, future scaffolding attempts could make use of other *Oryza* species for reference-guided scaffolding (e.g., the *O. officinalis* assembly may be a better candidate for scaffolding than the *O. sativa* Nipponbare assembly as *O. australiensis* is a member of the *O. officinalis* complex). However, this would still present biases in contig ordering and orientation. Therefore, we recommend that future scaffolding attempts either make use of pre-existing *O. australiensis* BAC libraries and FingerPrinted Contigs (FPC) to order and orient the contigs[1,13] or use Hi-C technology to generate new data for *O. australiensis*. Given the still experimental nature of Hi-C and difficulties in obtaining high coverage of linked reads these efforts may require significant investment.

Along with structural variations in the *O. australiensis KR* genome relative to *O. sativa* Nipponbare, we also uncovered many SNVs (Supplementary Fig. S8). Interestingly, there were more SNVs between some *O. australiensis* genotypes and *O. australiensis* KR than there were between *O. sativa* and *O. australiensis* KR based on short read mapping. The density of SNVs increases at the terminal regions of the *O. sativa* chromosomes. Conversely,

there is a region on each chromosome where the density of SNVs is low. These regions may represent the centromeric repeat elements of the chromosome (Supplementary Fig. S8). While centromeric regions have a common functional role in organising cell division, their sequences can be diverse between species. However, in the case of rice and other cereals, they are highly conserved[63,64]. For example, Gao et al.[65] showed that *O. australiensis* and *O. sativa* share similar centromeric retrotransposon repeat (CRR) sequences even though other rice species such as *O. brachyantha* have unique CRR profiles. This helps explain why sequence similarity between *O. australiensis KR* and *O. sativa* at centromeric regions is higher than at other regions of the genome.

Interestingly, when comparing four genotypes of *O. australiensis* that had been documented for their contrasting salt tolerance[20], we found that SNV density was not well correlated with the geographical distance between each of the populations (Fig. 1). For example, VR had at least 1.5 million more SNVs than the other two genotypes (when compared with KR), yet it was collected from a site far closer to Keep River than the D genotype. We speculate that SNV densities may be more related to the diversity of alleles required for adaptation to stresses than physical separation between populations. For example 'relative salt tolerance'[20] (Fig. 1) appeared to correlate better with the abundance of SNV abundance than simple distance between populations.

Contrary to the SNV data, the genetic distances estimated by *kWIP do* reflect the geographic distance between the *O. australiensis* genotypes (Fig. 3). That is, populations of *O. australiensis* genotypes that occur closer to each other in space are more genetically similar than genotypes that occur further apart. This is expected as *O. australiensis* is a mostly inbreeding species (~3% outcrossing), so gene flow and introgressions between genotypes should be low. Thus, we assume that these populations of *O. australiensis* are reproductively isolated and, as such, their genomes are evolving independently. If this is true, we expect to see local adaptation in these genotypes[66], a hypothesis consistent with the findings of Yichie et al.[20].

The genetic-distance tree generated by *kWIP* follows previously published *Oryza* phylogenies[13]. Importantly, the *O. australiensis* lineage diverges from the rest of the *O. officinalis* complex members (the B, C, and D genomes) suggesting that we can expect to see lineage specific adaptation in the *O. australiensis* genome. Given that this species occupies a diverse yet adverse suite of environmental niches, the genomic variation that underpins the divergence of the *O. australiensis* lineage may be associated with tolerance to multiple abiotic and biotic stresses. The functionally annotated genome that we present for *O. australiensis* KR here is therefore a resource that can be used to explore this genomic variation to uncover genes involved in stress tolerance.

Conclusions

Here we present the first reference-quality genome of *O. australiensis KR*—a wild species of rice native to the northern regions of Australia. While *O. australiensis* co-occurs with other *Oryza* species, it is genetically and phenotypically unique among the *Oryza* genus and is the sole member of the EE genome clade[67]. It demonstrates unrivalled tolerance to extreme environmental conditions, including heat, drought and soil salinity[3,18,21,68]. The leaf anatomy and associated photosynthetic efficiency of *O. australiensis* are also unique traits[22] and its grain has distinct pigmentation and starch composition[23,24]. Such traits make it an important genetic resource in the *Oryza* genus, as recognised formally by the OMAP and IOMAP[11,12]. However, no group has successfully assembled the nuclear genome for this species until now. It is expected that this genomic resource will enable phenomics in wild cereal relatives and lead to the selection of useful traits for the improvement of domestic rice cultivars, and in the breeding and domestication of *O. australiensis* itself. For example, given that we now know the positions of genes in the *O. australiensis* KR assembly, one can use the assembly alongside short reads from diverse *O. australiensis* accessions to identify SNVs in genes associated with tolerance in accessions from adverse environments. Similarly, with further Iso-Seq data one could identify novel transcript isoforms across different *O. australiensis* accessions under multiple stresses. Furthermore, given the role that gene regulation plays in response to stress, the *O. australiensis* KR assembly will uncover promoters of genes of interest. This work could extend to the use of long read data (either PacBio or ONT reads) to investigate the response of the epigenome to abiotic stresses. Finally, an expanded collection of *O. australiensis* accessions with contrasting phenotypes under extreme conditions will facilitate the identification of markers for stress tolerance via strategic inclusion of tolerant accessions to make mapping populations.

Data availability

This Whole Genome project (i.e., the *Canu*-generated purged contigs reported here) has been deposited at DDBJ/ENA/GenBank under the accession JAIFGZ000000000. The version described in this paper is version JAIFGZ010000000. All long- and short-read data associated with the reported genome assembly have been deposited at the Sequence Read Archive and GenBank in NCBI under BioProject PRJNA743927. The AGP file required to construct the pseudomolecules from the *Canu* contigs, the annotation for protein-coding genes in the contigs, the annotation for repeat elements in the scaffolds, and the annotation for protein-coding genes in the scaffolds (all GFF3 files) have been provided as Supplementary Files and can be accessed via figshare: https://doi.org/10.6084/m9.figshare.c.5875592.v2.

Received: 17 November 2021; Accepted: 14 June 2022
Published online: 25 June 2022

References

1. Stein, J. C. *et al.* Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. *Nat. Genet.* **50**, 285–296 (2018).
2. Brar, D. S. & Khush, G. S. Wild relatives of rice: A valuable genetic resource for genomics and breeding research. In *The Wild Oryza Genomes* (eds Mondal, T. K. & Henry, R. J.) 1–25 (Springer, 2018).

3. Atwell, B. J., Wang, H. & Scafaro, A. P. Could abiotic stress tolerance in wild relatives of rice be used to improve *Oryza sativa*? *Plant Sci.* **215–216**, 48–58 (2014).
4. Bisht, D. S., Solanke, A. U. & Mondal, T. K. Informatics of wild relatives of rice. In *The Wild Oryza Genomes* (eds Mondal, T. K. & Henry, R. J.) 27–40 (Springer, 2018).
5. Huang, X. *et al.* Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* **42**, 961–967 (2010).
6. Bolser, D., Staines, D. M., Pritchard, E. & Kersey, P. Ensembl plants: Integrating tools for visualizing, mining, and analyzing plant genomics data. *Methods Mol. Biol.* **1374**, 115–140 (2016).
7. Henry, R. J. *et al.* Australian Oryza: Utility and conservation. *Rice* **3**, 235–241 (2010).
8. Hiromi, K.-K. *et al.* OryzaGenome2.1: Database of diverse genotypes in wild Oryza species. *Rice* **14**, 00468 (2021).
9. Yu, H. *et al.* A route to de novo domestication of wild allotetraploid rice. *Cell* **184**, 1156–1170 (2021).
10. Song, S., Tian, D., Zhang, Z., Hu, S. & Yu, J. Rice genomics: Over the past two decades and into the future. *Genomics Proteomics Bioinform.* **16**, 397–404 (2018).
11. Wing, R. A. *et al.* The Oryza map alignment project: The golden path to unlocking the genetic potential of wild rice species. *Plant Mol. Biol.* **59**, 53–62 (2005).
12. Jacquemin, J., Bhatia, D., Singh, K. & Wing, R. A. The International Oryza Map Alignment Project: Development of a genus-wide comparative genomics platform to help solve the 9 billion-people question. *Curr. Opin. Plant Biol.* **16**, 147–156 (2013).
13. Shenton, M. *et al.* Evolution and diversity of the wild rice *Oryza officinalis* complex, across continents, genome types, and ploidy levels. *Genome Biol. Evol.* **12**, 413–428 (2020).
14. Shi, C. *et al.* The draft genome sequence of an upland wild rice species, *Oryza granulata*. *Sci. Data* **7**, 131 (2020).
15. Piegu, B. *et al.* Doubling genome size without polyploidization: Dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. *Genome Res.* **16**, 1262–1269 (2006).
16. Ammiraju, J. S. S. *et al.* The Oryza bacterial artificial chromosome library resource: Construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genome types of the genus Oryza. *Genome Res.* **16**, 140–147 (2006).
17. Zhang, Q. J. & Gao, L. Z. Rapid and recent evolution of LTR retrotransposons drives rice genome evolution during the speciation of AA-genome Oryza species. *G3 Genes Genomes Genet.* **7**, 1875–1885 (2017).
18. Phillips, A. L., Scafaro, A. P. & Atwell, B. J. Photosynthetic traits of Australian wild rice (*Oryza australiensis*) confer tolerance to extreme daytime temperatures. *Plant Mol. Biol.* https://doi.org/10.1007/s11103-021-01210-3 (2022).
19. Scafaro, A. P. *et al.* A thermotolerant variant of Rubisco activase from a wild relative improves growth and seed yield in rice under heat stress. *Front. Plant Sci.* **9**, 1–11 (2018).
20. Yichie, Y., Brien, C., Berger, B., Roberts, T. H. & Atwell, B. J. Salinity tolerance in Australian wild Oryza species varies widely and matches that observed in *O. sativa*. *Rice* **11**, 1–14 (2018).
21. Hamzelou, S. *et al.* Wild and cultivated species of rice have distinctive proteomic responses to drought. *Int. J. Mol. Sci.* **21**, 5980 (2020).
22. Mathan, J., Singh, A., Jathar, V. & Ranjan, A. High photosynthesis rate in two wild rice species is driven by leaf anatomy mediating high Rubisco activity and electron transport rate. *J. Exp. Bot.* https://doi.org/10.1093/jxb/erab313 (2021).
23. Tikapunya, T., Fox, G., Furtado, A. & Henry, R. Grain physical characteristic of the Australian wild rices. *Plant Genet. Resour.* **15**, 409–420 (2016).
24. Tikapunya, T. *et al.* Molecular structures and properties of starches of Australian wild rice. *Carbohydr. Polym.* **172**, 213–222 (2017).
25. Joly-Lopez, Z. *et al.* An inferred fitness consequence map of the rice genome. *Nat. Plants* **6**, 119–130 (2020).
26. Michael, T. P. & VanBuren, R. Building near-complete plant genomes. *Curr. Opin. Plant Biol.* **54**, 26–33 (2020).
27. Jones, A. *et al.* High-molecular weight DNA extraction, clean-up and size selection for long-read sequencing. *PLoS ONE* **16**, e0253830 (2021).
28. Jones, A., Borevitz, J., Warthmann, N. & Jones, A. Cost-conscious generation of multiplexed short-read DNA libraries for whole genome sequencing. *Protocols.io.* https://doi.org/10.17504/protocols.io.unbevan (2018).
29. Betts, N. S. *et al.* Isolation of tissues and preservation of RNA from intact, germinated barley grain. *Plant J.* **91**, 754–765 (2017).
30. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764–770 (2011).
31. Vurture, G. W. *et al.* GenomeScope: Fast reference-free genome profiling from short reads. *Bioinformatics* **33**, 2202–2204 (2017).
32. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020).
33. Koren, S. *et al.* Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
34. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
35. Kundu, R., Casey, J. & Sung, W.-K. Hypo: Super fast & accurate polisher for long read genome assemblies. *BioRxiv.* https://doi.org/10.1101/2019.12.19.882506 (2019).
36. Lindenbaum, P. *JVarkit: Java-Based Utilities for Bioinformatics*. https://doi.org/10.6084/m9.figshare.1425030.v1 (2015).
37. Shen, W. & Xiong, J. TaxonKit: A cross-platform and efficient NCBI taxonomy toolkit. *BioRxiv*. https://doi.org/10.1101/513523 (2019).
38. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative genomics viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2012).
39. Roach, M. J., Schmidt, S. A. & Borneman, A. R. Purge Haplotigs: Allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinform.* **19**, 460 (2018).
40. Ou, S. & Jiang, N. LTR_FINDER_parallel: Parallelization of LTR_FINDER enabling rapid identification of long terminal repeat retrotransposons. *Mob. DNA* **10**, 48 (2019).
41. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinform.* **9**, 18 (2008).
42. Ou, S., Chen, J. & Jiang, N. Assessing genome assembly quality using the LTR Assembly Index (LAI). *Nucleic Acids Res.* **46**, e126 (2018).
43. Seppey, M., Manni, M. & Zdobnov, E. M. BUSCO: Assessing genome assembly and annotation completeness. In *Gene Prediction* (ed. Kollmar, M.) 227–245 (Springer, 2019).
44. Jain, C., Koren, S., Dilthey, A., Phillippy, A. M. & Aluru, S. A fast adaptive algorithm for computing whole-genome homology maps. *Bioinformatics* **34**, i748–i756 (2018).
45. Alonge, M. *et al.* RaGOO: Fast and accurate reference-guided scaffolding of draft genomes. *Genome Biol.* **20**, 224 (2019).
46. Kawahara, Y. *et al.* Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* **6**, 4 (2013).
47. Poorten, T. *dotPlotly* (2018).
48. Gel, B. & Serra, E. karyoploteR: An R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **33**, 3088–3090 (2017).
49. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
50. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

51. Ou, S. *et al.* Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, 275 (2019).
52. Wing, R. *International Oryza Map Alignment Project*. https://doi.org/10.7946/P2FC9Z (2017).
53. Götz, S. *et al.* High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* **36**, 3420–3435 (2008).
54. BioBam. *OmicsBox—Bioinformatics Made Easy* (2019).
55. Pedro, D. L. F. *et al.* An atlas of plant transposable elements [version 1; peer review: 2 approved]. *F1000Research* **10**, 1194 (2021).
56. Murray, K. D., Webers, C., Ong, C. S., Borevitz, J. & Warthmann, N. kWIP: The k-mer weighted inner product, a de novo estimator of genetic similarity. *PLoS Comput. Biol.* **13**, e1005727 (2017).
57. Crusoe, M. R. *et al.* The khmer software package: Enabling efficient nucleotide sequence analysis [version 1; peer review: 2 approved, 1 approved with reservations]. *F1000Research* **4**, 900 (2015).
58. Uozu, S. *et al.* Repetitive sequences: cause for variation in genome size and chromosome morphology in the genus Oryza. *Plant Mol. Biol.* **35**, 791–799 (1997).
59. Vitte, C., Panaud, O. & Quesneville, H. LTR retrotransposons in rice (*Oryza sativa* L.): Recent burst amplifications followed by rapid DNA loss. *BMC Genomics* **8**, 218 (2007).
60. Multani, D. S. *et al.* Development of monosomic alien addition lines and introgression of genes from *Oryza australiensis* Domin. to cultivated rice *O. sativa* L.. *Theor. Appl. Genet.* **88**, 102–109 (1994).
61. Abbasi, F. M., Ahmad, H., Sajid, M., Inamullah, & Brar, D. S. Detecting nature of chromosome pairing in A and E genomes of Oryza. *Cytologia (Tokyo)* **74**, 223–227 (2009).
62. Abbasi, F. M. *et al.* Assessment of genomic relationship between *Oryza sativa* and *Oryza australinesis*. *Afr. J. Biotechnol.* **9**, 1312–1316 (2010).
63. Aragón-Alcaide, L., Miller, T., Schwarzacher, T., Reader, S. & Moore, G. A cereal centromeric sequence. *Chromosoma* **105**, 261–268 (1996).
64. Jiang, J. *et al.* A conserved repetitive DNA element located in the centromeres of cereal chromosomes. *Proc. Natl. Acad. Sci.* **93**, 14210–14213 (1996).
65. Gao, D. *et al.* A lineage-specific centromere retrotransposon in *Oryza brachyantha*. *Plant J.* **60**, 820–831 (2009).
66. Lascoux, M., Glémin, S. & Savolainen, O. Local adaptation in plants. *eLS*. https://doi.org/10.1002/9780470015902.a0025270 (2016).
67. Henry, R. J. *Oryza australiensis* Domin. In *The Wild Oryza Genomes* (eds Mondal, T. K. & Henry, R. J.) (Springer, 2018).
68. Yichie, Y. *et al.* Salt-treated roots of Oryza australiensis seedlings are enriched with proteins involved in energetics and transport. *Proteomics* **19**, 1900175 (2019).

Acknowledgements
The authors would like to acknowledge Ming-Dao Chia for providing initial bioinformatics support. We thank Scott Ferguson, Lina Herliana, Chelsea Matthews, and Farhad Masoomi-Aladizgeh for the discussions about bioinformatics analyses. We thank Dr Sally Norton from the Australian Grains Genebank in Horsham Victoria, Australia for providing the seed a decade ago that led to the present study. We thank Dr. James Cowley (Founder of 'The Figure Legend') for his advice on designing figures. We would like to deeply thank the reviewers, whose thoughtful comments helped substantially improve the manuscript.

Author contributions
A.P., S.F., N.W.H., A.J., J.B., R.B., and B.A. contributed to technical planning after B.A. initiated the project with a successful application for a Macquarie University DVC(R) Discretionary grant. J.B. and A.J. provided technical expertise for execution of the sequencing. A.P. and A.J. performed all DNA and RNA preparations and performed all ONT MinION sequencing steps. A.J. performed Illumina sequencing steps, gave initial bioinformatics support, and supplied FASTQ files. N.W.H. performed Iso-Seq analysis and provided bioinformatics support. A.P. and S.F. performed all data analysis, genome assembly, and subsequent analyses. N.W.H. and R.B. provided access to the computing resources required. N.W.H., R.B., and B.A. provided guidance throughout. B.A. grew the plants for sequencing prior to extractions in Canberra. A.P. created all figures and wrote the initial draft of the manuscript. All authors contributed to the final manuscript.

Funding
This work was funded by the Macquarie University DVC(R) Fund with contributions from the Department of Biological Sciences. This project was supported by the ARC Centre of Excellence in Plant Energy Biology (CE140100008). AP acknowledges support from an Australian Research Training Program Scholarship and the FJ Sandoz Scholarship from the University of Adelaide. The funding was also provided by Deputy Vice-Chancellor (Research) Discretionary Fund (51277/20).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-14893-5.

Correspondence and requests for materials should be addressed to B.J.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© Crown 2022

Chapter 6

CLAW: An automated Snakemake workflow for the assembly of chloroplast genomes from long-read data

Wombo Prompt

“Chloroplast assembly”

Wombo Theme

“Comic”

Statement of Authorship

Title of Paper	*CLAW*: An automated Snakemake workflow for the assembly of chloroplast genomes from long-read data
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Aaron L. Phillips		
Contribution to the Paper	Conceived the study, analysed data, wrote the manuscript		
Overall percentage (%)	65%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	7th September 2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Scott Ferguson		
Contribution to the Paper	Conceived the study, analysed data, edited the manuscript		
Signature		Date	16th September 2022

Name of Co-Author	Nathan S. Watson-Haigh

Contribution to the Paper	Conceived the study, assisted analyses, edited the manuscript		
Signature		Date	7th September 2022

Name of Co-Author	Rachel A. Burton		
Contribution to the Paper	Assisted project logistics, edited the manuscript		
Signature		Date	12th October 2022

Target Journal:
PLoS Computational Biology Methods

Title:
CLAW: An automated Snakemake workflow for the assembly of chloroplast genomes from long-read data

Short Title:
CLAW: automated chloroplast genome assembly

Authors:
Aaron Phillips[1*], Scott Ferguson[2], Rachel Burton[1], Nathan Watson-Haigh[3,4]

Affiliations:
[1]Department of Food Science, University of Adelaide, Adelaide, South Australia, Australia
[2]Research School of Biology, Australian National University, Canberra, Australian Capital Territory, Australia
[3]South Australian Genomics Centre, University of Adelaide, Adelaide, South Australia, Australia
[4]Australian Genome Research Facility, Victorian Comprehensive Cancer Centre, Melbourne, Victoria, Australia

***Corresponding author:**
Email: aaron.phillips@adelaide.edu.au

We envision part of our tool behaving much like the claw in a vending machine – it reaches into WGS read libraries and extracts chloroplast reads – thus, we have named it Chloroplast Long-read Assembly Workflow (CLAW).

Abstract

Chloroplasts are photosynthetic organelles in plant cells that, due to an ancient endosymbiosis event(s), contain their own genome. Chloroplast genomes are commonly used in evolutionary studies and barcoding and are increasingly becoming a target for crop improvement studies. As DNA sequencing becomes more affordable, researchers are collecting vast swathes of high-quality whole-genome sequence data from laboratory and field settings alike. Whole-tissue read libraries sequenced with the primary goal of understanding the nuclear genome will inadvertently contain many reads derived from the chloroplast genome. These whole-genome, whole-tissue read libraries can additionally be used to assemble chloroplast genomes with little to no extra cost. While several tools exist that make use of short-read second generation sequencing data, to assemble chloroplast genomes, currently no long-read, third generation tools exist. Here, we present *CLAW* (Chloroplast Long-read Assembly Workflow), an easy-to-install and easy-to-use Snakemake tool to assemble chloroplast genomes from chloroplast long reads found in whole-genome read libraries (https://github.com/aaronphillips7493/CLAW). Using 19 publicly available reference chloroplast genome assemblies and long-read libraries from algal, monocot and eudicot species, we show that *CLAW* can rapidly produce chloroplast genome assemblies with high similarity to the reference assemblies. Using genes extracted from the assembled genomes we show that the output of *CLAW* can be used to generate single-gene trees that are identical to those derived from the reference assemblies. *CLAW* was designed such that users have complete control over parameterisation, allowing individuals to optimise *CLAW* to their specific use cases. We expect that *CLAW* will provide researchers with an additional resource that they can contribute to the growing number of publicly available chloroplast genome assemblies.

Author Summary

Chloroplast genomes are important resources as they can be used to help resolve phylogenies and aid in species identification. The importance of chloroplasts and their genes in plant stress responses is a field of research in its infancy that also stands to benefit from publicly available chloroplast sequence data. As long read sequencing technology becomes more accessible, researchers can generate, and access troves of data contained in long read libraries. Often embedded in these libraries are chloroplast reads. With the right tools, these reads can be extracted and use for chloroplast genome assembly. Existing tools are often hard to install and/or are made up multiple manual steps that reduce accessibility to non-expert users. Here, we present *CLAW* (Chloroplast Long-read Assembly Workflow) – an easy to install, easy to use workflow for the assembly of chloroplast genomes from long read data. We anticipate that this new tool will remove many of the barriers to entry that might dissuade novice users from participating in the field of bioinformatics and encourage the de novo assembly of chloroplast genomes from diverse plant species.

Introduction

Chloroplasts are organelles that perform photosynthesis in photosynthetic cells. They convert light energy into a stable form of chemical energy, a process essential to life on earth. In addition to their importance as the organelle of photosynthesis, chloroplasts have been extensively used throughout biological research in part because they contain their own DNA genome.

The typical chloroplast genome is circular with a quadripartite structure comprised of a large single copy (LSC) region, a small single copy (SSC) region, and two inverted repeats (IR) (e.g., 1). Chloroplast genomes have low mutation rates and are highly

conserved, usually maternally inherited, do not undergo recombination, and are typically 120-160 kbp in length (2). Each plant cell contains many chloroplast organelles and each chloroplast can contain 10,000 or more copies of its genome (3–6). All of these features make chloroplast genomes well-suited to studies in phylogenetics (7,8) and species identification (9,10). There is also growing interest in the role(s) that chloroplasts play in plant responses to stress and thus how they can contribute to bolstering crop performance in a changing world (11,12).

Most chloroplast genome assembly efforts to date have relied on short (100-150 bp) read sequencing technologies, such as Illumina. As such, several short-read specific chloroplast assembly tools have been developed (13). While short reads are highly accurate, they often fail to assemble repetitive genomic regions (14) and can fail to detect the structural variations that are now known to be pervasive in genomic sequences (e.g., 15). Assembly of chloroplast genomes, despite their small size, also suffer from this phenomenon because of the IR regions (16,17). Additionally, the LSC and SSC regions of the chloroplast can be challenging for short-read assemblers to resolve because these regions can exist in a ‘flip-flop’ state – a term used to describe the tendency of different regions (usually the SSC or IRs) of the genome to invert, leading to the possibility of multiple chloroplast genome assemblies from one cell (17–19).

With the advent of third generation long-read sequencing (e.g., Oxford Nanopore Technology (ONT), Pacific Biosciences), we can now affordably generate reads that span or can assemble across large repeats. Long-read sequencing allows for computationally easier and more structurally accurate assembly of genomes (20). It is becoming increasingly feasible for research groups of any size to collect large

volumes of whole-genome sequencing data for their species of interest. In addition to nuclear sequences, whole-genome sequencing projects will unavoidably generate many chloroplast sequencing reads, as the chloroplast genome exists at high copy numbers within tissues (5). Thus, if photosynthetic tissue is used to generate reads, chloroplast genomes can be assembled without any additional sequencing.

Different methods to assemble chloroplast genomes have been developed (13). All methods share two key steps: 1) identification and extraction of chloroplast reads; and 2) assembly of the chloroplast genome. The identification of chloroplast reads from whole-genome read libraries can be achieved by k-mer analysis, where the most frequent reads are assumed to be plastid (21); or by baiting. Here, we present *CLAW*, which uses the read baiting method. Whole-genome read libraries are aligned to a related pre-existing chloroplast genome assembly, and aligned reads are extracted, creating a chloroplast enriched dataset. Enrichment/baiting is required as nuclear genomes contain genes that have high sequence similarity with chloroplast genes, thus not all aligned reads will originate from the chloroplast genome. After baiting, the chloroplast enriched read library is assembled with two long-read assemblers, *Flye* (22) and *Unicycler* (23), that are circular sequence aware. While this library is enriched for chloroplast reads, there is a possibility that the library will also contain mitochondrial reads because of high sequence similarity between mitochondrial and chloroplast genomes (24,25). Thus, it is possible for *CLAW* to assemble mitochondrial (likely fragmented) contigs too. *CLAW* makes no attempt to resolve this, so users will have to manually check the assembled contigs for similarity to chloroplast and/or mitochondrial sequences.

Despite the increasing use of long-read sequencing technologies, an easy to use, automated, and reliable method to assemble long-reads into chloroplast genomes remains unavailable. *CLAW* is an easy to install and easy to use tool for the reference-guided long-read assembly of chloroplast genomes. This tool provides users having little bioinformatics experience a fast, easy, and reproducible way of assembling chloroplast genomes from long reads. Our method is automated, requiring only minimal user input and makes use of freely available and/or published tools.

Methods

Overview of *CLAW*

CLAW begins with a long-read whole genome sequence library, stored in either the FASTA or FASTQ format. FASTA/Q files can optionally be compressed with *gzip*. If using ONT technology, the raw sequencer output (fast5) must be base-called by base-calling software (e.g., *Guppy* or another method). As *CLAW* makes use of a reference chloroplast genome (RCG) to bait reads for assembly, it is a requirement to provide a RCG of the focal species or a closely related species. *CLAW* will download this reference - the user only has to provide a National Center for Biotechnology Information (NCBI) reference sequence identifier (e.g., NC_031333.1). Some HPC configurations will not allow for external data to be downloaded via submitted jobs, which may interfere with this step. Additional parameters (See ‘User Specifications’ below) can be set to specify: whether the user is working with FASTA or FASTQ file(s), the kind of long reads the user is working with (e.g., ONT vs PacBio), the number of aligned reads to subsample for assembly, the expected chloroplast genome size, and how many CPUs *CLAW* should use.

After downloading the RCG, *CLAW* circularises (i.e., duplicates the linear reference sequence and then joins the 3’ end of one to the 5’ end of the other) the RCG to facilitate read mapping (Fig. 6.1). This step is required as chloroplast genomes, which are circular in nature, are stored as linear sequences. Attempting to map reads across the break point may result in poor read alignment (Fig. 6.1). The read library is mapped to the RCG, and all chloroplast mapping reads (CMRs) are extracted. The extracted CMRs are filtered to remove reads shorter than 5 kbp in length and reads that are larger than the expected chloroplast genome size. Extremely high coverage can confound assemblers and worsen results, as such a reduced subset of reads is randomly selected for assembly. For repeatability, a seed for read sub-setting can be specified or, alternatively, a random seed will be generated by *CLAW*. The number of randomly selected assembly reads can be user-specified, enabling easy coverage adjustment for tuning of assembly time and success (e.g., more reads will give higher genome coverage but will increase the time required for assembly).

Assembly is performed using *Flye* and/or *Unicycler,* resulting in the generation of a fasta file (containing the assembly) and a genome graph (showing paths through the genome). The resulting assemblies will be rotated to match the breakpoint of the reference chloroplast genome and a dot plot produced using the MUMmer3 suite (27) for assessment of chloroplast assembly success. *CLAW* will also produce BAM and bigwig files of read alignments to RCG and the candidate chloroplast genome assembly such that read depth can be investigated. If a complete chloroplast genome is not produced, *CLAW* can be rerun with a different random seed, and different number of assembly reads.

User Specifications

For *CLAW* to operate successfully, the user must edit the “config.yml” file located in the same directory as the Snakefile. Within “config.yml” the **mandatory** user-defined fields are:

- **NCBI_reference_accession**, where users must specify the NCBI accession number for the RCG of interest. We recommend that users select a chloroplast genome from a closely related species. The accession number for an *Oryza sativa* chloroplast genome assembly is provided as default (i.e., NC_008155.1).
- **my_email**, where users must enter an email address; this will be used by NCBI to ensure that users are not abusing the system by downloading large numbers of genome files.
- **fast_file**, where users must specify whether their read file(s) is in the FASTA or FASTQ format and whether it is gzip compressed or not (‘fasta’ vs ‘fastq’ vs ‘fasta.gz’ vs ‘fastq.gz’).
- **flye_parameter**, where users must specify if they are working with ONT long reads, PacBio reads, or PacBio HiFi reads.
- **minimap2_parameter**, where users must specify if they are working with ONT long reads, PacBio reads, or PacBio HiFi reads.
- **chloroplast_size**, where users must provide an estimated chloroplast genome size (in units of bp). For simplicity, we recommend that this is obtained from the NCBI page for the RCG selected above.

There are also some editable parameters that users need not change. While these parameters need not be edited, the assembly process can be optimised on a per sample basis by fine tuning them:

- **rand_seed**, which provides a seed number to *seqtk* for the random sub-sampling of long reads prior to genome assembly. If the user does not specify a seed, *CLAW* will generate a random seed each time it is run. Default = blank.
- **number_reads**, which specifies the number of reads that *seqtk* should sub-sample from the total read library that will be used for chloroplast genome assembly. Read number can be increased to increase genome coverage. Note: high coverage (~>300x) can affect assembly performance. Default = 3000.
- **read_min_length**, which specifies the minimum read length that will be carried forward for assembly. Reads below this threshold will be removed from the aligned read pool and thus will not be used for assembly. Default = 5000 bp.
- **cpus**, which specifies the number of CPUs the pipeline will call upon for the execution of the workflow. We recommend that users check their system before modifying this parameter. Default = 6.

Testing *CLAW*

We used publicly available ONT long read libraries generated from 19 species (4 algal, 4 monocot, and 11 eudicots) with previously assembled chloroplast genome sequences to test *CLAW* (Table 1). Long-read libraries were downloaded from NCBI using *fasterq-dump* and saved into a directory called “chloro_assembly/reads” within the cloned git repository. Reference genomes were downloaded from NCBI using the *entrez-direct* ‘esearch’ function and saved into a directory called “chloro_assembly/reference” within the cloned git repository. Resources for each job in *CLAW* (e.g., time and memory) were defined in “cluster-configs/default.yaml” and

tracked using Snakemake's internal "benchmark" feature. *CLAW* was executed on a Slurm HPC by running the following command for each sample: "snakemake --profile profiles/slurm". Each time *CLAW* was run, the reference genome for each respective read file was specified in the "config.yaml" file. *DNAdiff* (27) was used to calculate percent identity between the assembled chloroplast genomes and their corresponding reference genome assemblies.

Exploring the Assemblies

GFA ('Graphical Fragment Assembly') files generated by *Flye* were used as input for *Bandage* to visualise genome structure and confirm the assemblies as chloroplast sequence as follows. The coding sequences of all chloroplast and mitochondrial-encoded genes for the 19 test samples (Table 1) were downloaded from NCBI as FASTA files. After building a BLAST database for each chloroplast assembly in *Bandage*, the chloroplast and mitochondrial coding sequences were used to annotate each assembly using the BLAST annotation feature within *Bandage*. Mitochondrial coding sequences were used to test whether the assembled contigs were of mitochondrial origin.

The Rubisco large subunit (RbcL) genomic sequences were identified in each of *CLAW*'s assemblies by using the *BLAST+ suite* (28) and publicly available RbcL sequences from each of the 19 reference chloroplast genomes as the query. The extracted RbcL genes were aligned to their respective reference RbcL genes in *UGENE* to assess sequence differences (manual inspection).

Paired two-sample *t*-tests were performed using the RealStats add-in for Excel to test for differences in the performance (genome contiguousness, genome percent identity to reference, time to completion, and RAM usage) of *Flye* and *Unicycler*-based workflows.

Table 6.1. Information on long reads used as input for *CLAW* and the *Flye*-generated chloroplast genome assembly statistics.

Genome graph in Figure 6.3	Taxonomic group	Species	Long read accession no.	Long reads used as input (Mbp)	Reference chloroplast genome accession no.	Reference chloroplast size (kbp)	Assembly size (kbp)		No. contigs		Mean coverage (x)		Similarity (%)	Time to completion (min)	RAM used (Gb)
							Chl	Mit	Chl	Mit	Chl	Mit			
A	Algae	*Chlamydomonas reinhardtii*	ERR3237140	60	NC_005353	204	204	-	2	-	240	-	98.2	15.2	15.9
B	Algae	*Chlorella variabilis*	DRR149372	21	NC_015359	124	120	-	1	-	161	-	96.2	5.2	10.9
C	Algae	*Ostreococcus tauri*	ERR5421724	8.5	NC_008289	72	75	62	1	1	92	35	99.5	3.3	9.8
D	Algae	*Pycnococcus provasolii*	DRR252953	12	NC_012097	80	80	-	1	-	136	-	99.5	4.9	9.4
E	Monocot	*Asparagus officinalis*	SRR9643837	22	NC_034777	157	136	-	2	-	107	-	99.1	8.1	12.0
F	Monocot	*Deschampsia antarctica*	SRR13908657	33	NC_023533	135	139	-	1	-	133	-	99.1	5.0	10.8
G	Monocot	*Oryza sativa*	DRR196880	42	NC_008155	135	163	-	1	-	230	-	99.4	12.9	15.8
H	Monocot	*Spirodela polyrhiza*	SRR11472010	67	NC_015891	169	168	-	1	-	249	-	99.2	19.8	17.5
I	Dicot	*Aquilaria sinensis*	SRR9858982	54	NC_029243	160	180	-	1	-	233	-	98.4	9.5	15.7
J	Dicot	*Cannabis sativa*	ERR3850904	42	NC_027223	154	189	-	1	-	199	-	99	8.9	14.7
K	Dicot	*Corylus avellana*	ERR4852503	26	NC_031855	160	189	-	3	-	109	-	99.6	5.9	12.4
L	Dicot	*Eucalyptus polybractea*	SRR8692273	36	NC_022393	160	158	-	2	-	176	-	98.8	7.0	13.8
M	Dicot	*Gossypium longicalyx*	SRR10377593	41	NC_023216	160	175	-	2	-	171	-	98	6.3	13.5
N	Dicot	*Lathyrus sativus*	ERR3374012	67	NC_014063	121	120	-	1	-	209	-	99.2	7.5	13.7
O	Dicot	*Medicago truncatula*	SRR10194526	45	NC_003119	124	123	-	1	-	287	-	99	15.4	15.9
P	Dicot	*Panax ginseng*	SRR13070229	29	NC_006290	156	201	-	3	-	118	-	99.6	6.3	11.9
Q	Dicot	*Prunus dulcis*	ERR3430399	31	NC_034696	158	162	-	1	-	142	-	99.1	6.7	14.5
R	Dicot	*Solanum commersonii*	SRR12407219	46	NC_028069	156	175	-	2	-	221	-	99.4	9.3	11.3
S	Dicot	*Vigna radiata*	SRR12549534	76	NC_013843	151	150	-	1	-	316	-	98.8	11.3	16.9

Results

Using 19 publicly available long read libraries and their corresponding reference chloroplast genome assemblies, we show that *CLAW* can be used to assemble chloroplast genomes from long reads of chloroplast origin contained within larger long read libraries (Fig. 6.2; Table 1). Unicycler workflows completed significantly faster than Flye workflows ($p < 0.0001$) and used significantly less RAM than Flye workflows ($p < 0.0001$; Table 1 & 2, and Supplementary Table 6.3). *Flye*-assembled genomes were more similar o their reference genomes than *Unicycler*-assembled genomes ($p < 0.01$), and *Flye* assemblies were significantly more contiguous than those generated by *Unicycler* ($p < 0.01$). Thus, we focus on the results from the *Flye*-generated assemblies only. The mean finish time of all 19 *Flye* tests of *CLAW* was 8.9 mins, with a range of 3.3-19.8 mins (Table 1). The mean memory usage for all 19 tests of *CLAW* was 13.5 Gb, with a range of 9.4-16.9 Gb (Table 1). The read alignment steps using *minimap2* and the chloroplast genome assembly by *Flye* consumed the most resources out of all jobs submitted as part of *CLAW*, with assembly requiring an average time of 5.3 mins and 8.3 Gb (RAM) to complete, and initial read alignment an average of 1.9 mins and 2.6 Gb.

Table 6.2. Mean (± SE) time to completion, RAM used, percent identity to the reference genomes, and number of contigs generated by *CLAW* following the Flye or Unicycler workflows.

Assembler	Time (min)	RAM (Gb)	Percent identity to reference (%)	No. contigs
Flye	8.87 ± 0.99	13.5 ± 0.56	98.9 ± 0.18	1.5 ± 0.16
Unicycler	6.19 ± 0.73	6.1 ± 0.62	98.6 ± 0.18	3 ± 0.44

When using *Flye*, *CLAW* was able to assemble 11 of the 19 (~58%) chloroplast genomes in a single contig (Table 1). Alignment of the assembled chloroplast genomes to their corresponding reference genomes produced the canonical

chloroplast-chloroplast genome alignment pattern (Fig. 6.2). Table 1 shows high sequence similarity (mean = 98.9%) between the genomes assembled by *CLAW* and the reference chloroplast genomes. Assembled genome size was on average ±15.5 kbp different from the reference chloroplast genome size, with a range of 0-65 kbp (Table 1). Seven of the assembled genomes were larger than the expected reference genome size, and 12 were smaller than the expected reference genome size. *Bandage* annotation using publicly available coding-gene information for each of the reference genomes shows that all of the assemblies are of chloroplast genomes (except for two additional sequences; see below). The Bandage plots also show a mixture of genome structure representations (Fig. 6.3). Circular representations of the assemblies indicate that the inverted repeats (IRs) of the genomes are not identical, while the “lasso” representations (which highlight the quadripartite structure of chloroplast genomes) indicate that the IRs are identical. For the algal samples, 50% (2/4) of the assemblies are represented as closed circles, and the remaining 50% are represented as “lasso’s”. As expected, *CLAW* generated part of a mitochondrial genome for some of the tested samples (in this case only the Flye assembly of *O. taurii*; Fig. 6.3C). Users will need to be cognisant that this could occur for their samples too and may inflate assembly size. For example, the *O. tauri* assembly was 137 kbp while the expected size was 72 kbp (Table 1). In the absence of the assembled mitochondrial contig, the chloroplast assembly size for *O. tauri* is 75 kbp, much more in line with the expected genome size of 72 kbp (Table 1). For the monocot samples, all four of the assemblies are represented by the “lasso” structure. For the eudicot samples, ~45% (5/11) of the assemblies are represented as closed circles, and the remaining ~55% are represented as “lasso’s”. For *C. avellana*, a second sequence that represents the IR for this species is present in the bandage plot (Fig. 6.3K).

We were able to identify and extract the genomic sequence of the RbcL genes from the 19 assemblies. On average, these genes were 99.3% identical to their corresponding reference RbcL gene sequences (including gaps), with a range of 98 - 100% identity. In total, we identified 96 gaps in the 19 assembled RbcL genes relative to the 19 reference RbcL genes. Of these, 95 occur in the assembled sequences and one appears in the reference for C. variabilis (i.e., an 'insertion' in the assembled C. variabilis gene). Variation appears to be driven in part by systematic ONT long-read errors, as ~54.9% (67/122) of identified variations occur as gaps in the assembled genomes where there are homopolymer runs of ~4-6 bp. The remaining 29 gaps (96 – 67 = 29) are not suspected to be caused by ONT systematic errors as their genomic contexts are random. Excluding the 96 gaps, 14 of the 19 assembled RbcL genes were identical to their reference sequences (RbcL genes from *O. taurii*, *C. variabilis*, *M. truncatula*, *V. radiata*, and *P. dulcis* were 99% similar to their respective reference sequences). Some substitutions (~13% = 16/122) occur between reference and assembled sequences, though these only occur in two of the algal RbcL genes (*O. taurii* and *C. variabilis*). The remaining variation is an 'insertion' in the assembled RbcL gene of *C. variabilis* relative to its reference gene.

Discussion

Here, we present *CLAW* - an easy to install, easy to use tool for the assembly of chloroplast genomes from chloroplast long reads embedded within Whole-Genome Sequencing (WGS) data sets. Pre-existing chloroplast genome assembly tools largely make use of short read data (13). To the best of our knowledge, this is the first report of a publicly available automated pipeline that assembles chloroplast genomes from long read data. Jin et al. (29) report on the possibility of using long

read data to assemble chloroplast genomes with their bespoke tool, *GetOrganelle*. However, at the time of writing, chloroplast genome assembly using long read data is not possible with the *GetOrganelle* toolkit. We show that *CLAW* can be used to assemble high quality chloroplast genomes from chloroplast reads embedded in larger read files. Additionally, *CLAW* can be used to glean additional value from any WGS project targeting photosynthetic tissues/cells (e.g., leaves, or algal cell suspensions). Even though we present results using ONT long read data alone here, *CLAW* can be easily modified to work with PacBio reads (continuous long reads and HiFi reads) via the modification of the “config.yml” file.

A recent review of chloroplast genome assembly tools making use of short read data showed a large range of completion times for genome assembly (~3 to ~835 mins; ,13). Further, the peak memory required for the analysed tools/parameters was ~2.6-25.2 Gb. Jin et al. (29) compared *GetOrganelle* to another plastid assembly pipeline, *NOVOPlasty*, and found that runtimes varied substantially (~4 to ~1024 mins), with a peak memory consumption of ~2-4 Gb. It took *CLAW* an average of 8.9 mins and 13.5 Gb to assemble each of the 19 chloroplast genomes. Thus, the pipeline reported here is time efficient, though more memory intensive than some assembly tools making use of short read data. Given that *CLAW* makes use of long-read data, this is to be expected.

CLAW was able to assemble complete and accurate (average identity = 98.9%) chloroplast genomes for all 19 test datasets. The accuracy of the assemblies relative to the reference genomes was not as high as the accuracy obtained from short read assemblers (which achieved an average identity of 99.9%, 29). This is to be expected, as short reads typically have higher per base accuracy than ONT long

reads (cv. 0.09-0.47% error rate in Illumina short reads vs. 5-22% error rate for ONT long reads depending on the base caller used; ,30–33). However, high coverage, such as those achieved by *CLAW* for the assembled genomes (mean coverage = 183x), can reduce the impact of random (though not systematic) ONT errors on the assembly.

While we report high accuracy of the assembled chloroplast genomes, *CLAW* produced incomplete, fragmented assemblies of some of the mitochondrial genome as well (Fig. 6.3). The in-built BLAST-based annotation feature of *Bandage* confirmed that the assembled chloroplast genomes were indeed chloroplast genomes and that the other sequences sometimes assembled were of mitochondrial origin. *CLAW* probably attempted to assemble mitochondrial sequences because some mitochondrial reads were likely included in the assembly read pool due to high sequence similarity with the reference chloroplast genomes. For example, Zhang et al. (26) report high (up to 72%) sequence similarity between plastid and mitochondrial sequences. Further, maize (*Zea mays*) mitochondrial and chloroplast genomes contain a 12kbp stretch of sequence that is identical (25). Sequence similarity between mitochondrial and chloroplastic genomes comes about in part due to ancient chloroplast-to-mitochondria gene transfer events (24). Thus, the read-enrichment strategy employed by *CLAW* is capable of baiting mitochondrial reads as well as chloroplastic reads. Future work should aim to either reduce or increase the representation of mitochondrial reads in the baited read set to avoid aberrant mitochondrial genome assemblies, or to make it possible to assemble more complete mitochondrial genome sequences, respectively. Reducing the representation of mitochondrial reads in the baited read library may be possible by implementing a filtering rule that uses sequence similarity with publicly available mitochondrial

genome sequences. A k-mer approach could also be used to delineate chloroplastic and mitochondrial reads in the future.

We were able to identify and isolate RbcL genes from all of the chloroplast genomes assembled using *CLAW*. The RbcL genes derived from our assemblies had 98-99.3% sequence identity (0.7-2% diversity) compared with the published gene sequences. This is not an unexpected level of diversity since RbcL genes within the Diospyros genus have ~98.2% sequence identity (34). Although much of the "diversity" observed in the RbcL genes we extracted from our assemblies is likely to be due to systematic errors known to exist in ONT data. The most obvious being the inability of ONT to accurately call the correct number of bases in homopolymer runs. Delahaye and Nicholis (35) analysed systematic issues associated with ONT long reads derived from bacterial and human samples and found that deletions occur at a frequency of 1.6-2.7% and are more likely to occur in GC rich regions. This may help to explain why we see up to 2% divergence in some of the RbcL genes, and indeed in the genomes themselves, here (i.e., the level of diversity is within the margin of error remaining in ONT-only assemblies). We are confident that *CLAW* can therefore be used to help answer questions regarding plastid sequence evolution.

While we report on the successful runs of *CLAW* here, there were several instances in which *CLAW* failed to complete or when chloroplast genome assembly failed. Typically, *CLAW* failed to complete due to insufficient memory allocation for certain steps of *CLAW*. These steps were commonly the read mapping and genome assembly steps. In this case, we were able to overcome this issue by modifying the memory allocation in the "cluster-configs/default.yaml" file. Future releases should aim to integrate dynamic memory allocations to further minimise user input for the

successful completion of *CLAW*. When genome assembly failed, we found that re-running *CLAW* with a different seed (randomly generated as part of the *CLAW*, unless specified by the user in “config.yaml”) resulted in successful completion of genome assembly. We also found it necessary to, on occasion, adjust the number of chloroplast reads used for genome assembly as too many reads (and therefore too high genome coverage) could cause the assemblers to fail. Thus, by modifying the random seed and the number of reads used for assembly we were able to generate assemblies for all 19 test datasets.

Figures

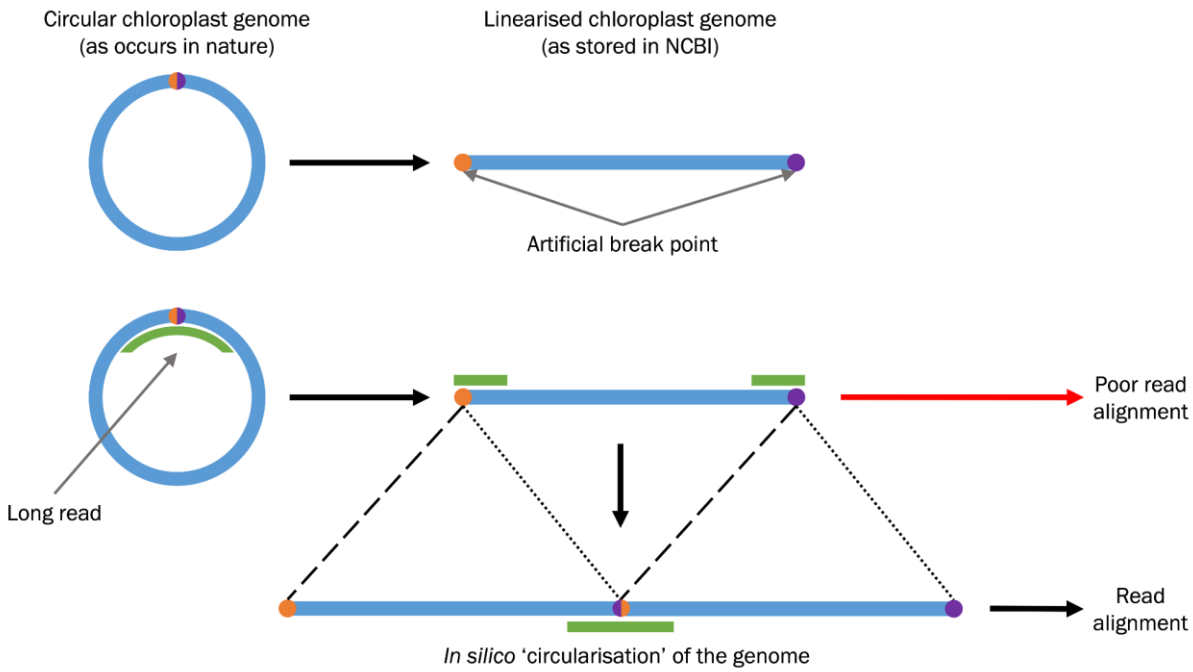


Figure 6.1. Demonstration of the need to 'circularise', *in silico*, linear reference chloroplast genome sequences after download from online databases. Linearised chloroplast genome sequences introduce artificial breakpoints in the sequence (orange and purple circles). Artificial breaks lead to poor long read (green lines) alignment, which affects chloroplast read enrichment/baiting. *In silico* re-circularisation of the linear sequence allows long reads to map across the artificial break points.

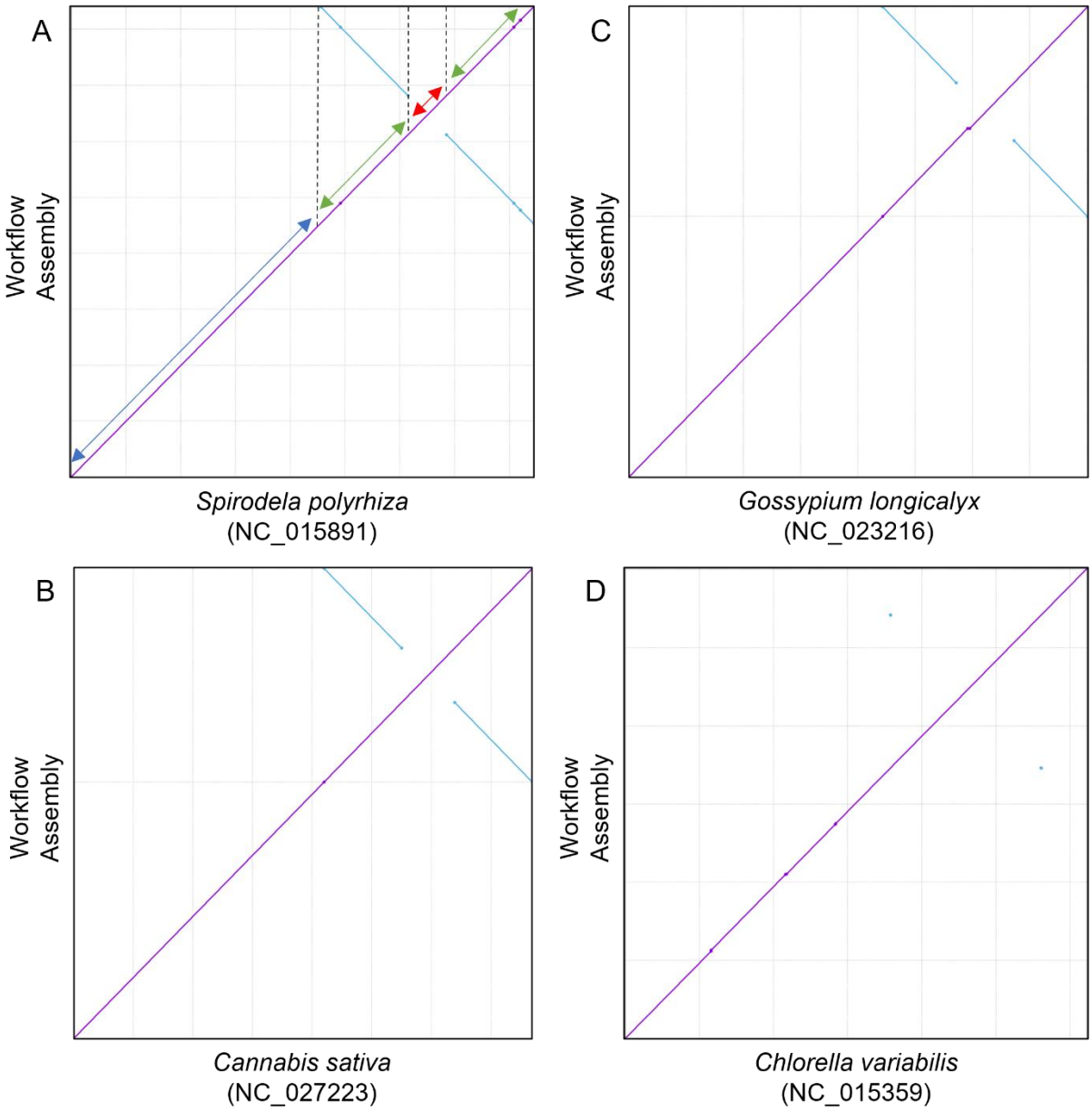


Figure 6.2. Representative reference-sample chloroplast genome alignments for (A) a monocot species, (B & C) two dicot species, and (D) an algal species. The reference genome is represented on the X-axis, and the genome assembled by *CLAW* is represented on the Y-axis. Species names and reference genome NCBI accession numbers appear on the X-axis. Please refer to Table 1 for the long read accession numbers used by *CLAW* for genome assembly. Each alignment presented here follows the canonical chloroplast genome-genome alignment patterns and the LSC (dark blue line in A), SSC (red line in A), and the two IR (green lines in A) regions are clearly identifiable (broken black line in A indicates the boundaries of each region).

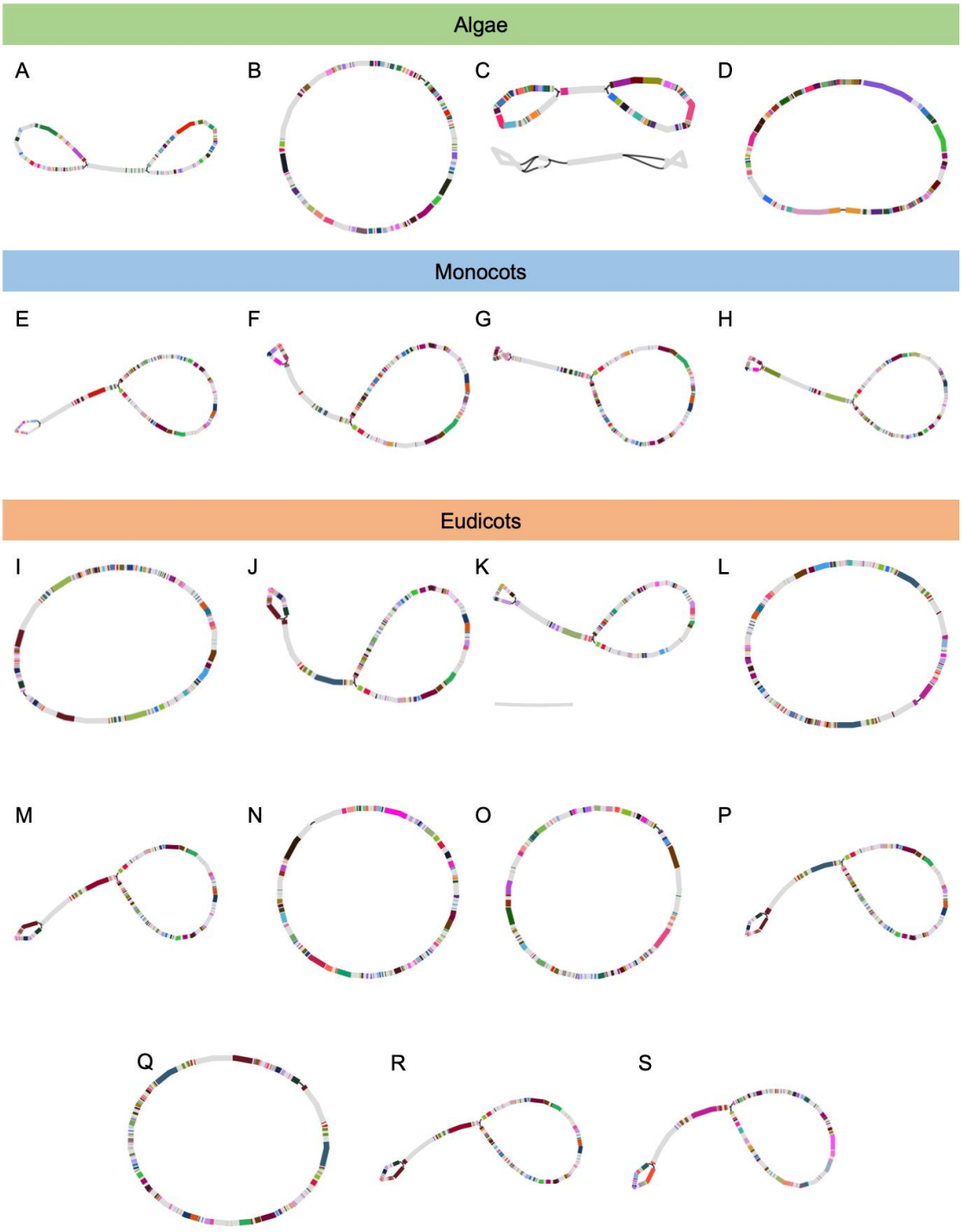


Figure 6.3. Repeat graphs for the 19 chloroplast genomes (4 algal species, 4 monocot species, and 11 eudicot species) assembled by *CLAW*. The 'lasso' style genome plots represent assemblies in which the IR regions (the flat lines connecting two circular pieces) are perfectly palindromic in the assembly, while the circular style genome plots represent assemblies in which the IR regions are not identical. The order of species in this plot follows the order of species in Table 1. The colouring of segments of each genome represents genome annotations assigned using the BLAST-based genome annotation feature of *Bandage*. Genomes were annotated using publicly available coding region annotations from each of the reference chloroplast genomes. (C) and (K) have additional contigs that could not be annotated using chloroplast coding regions as they are mitochondrial genome fragments.

Supplementary Material

Supplementary Table 6.3. Information on long reads used as input for *CLAW* and the *Unicycler*-generated chloroplast genome assembly statistics.

Genome graph in Figure 3	Taxonomic group	Species	Long read accession no.	Long reads used as input (Mbp)	Reference chloroplast genome accession no.	Reference chloroplast size (kbp)	Assembly size (kbp)			No. contigs			Similarity (%)	Time to completion (min)	RAM used (Gb)
							Chl	Mit	Oth	Chl	Mit	Oth			
A	Algae	*Chlamydomonas reinhardtii*	ERR3237140	60	NC_005353	204	467	-	59	3	-	1	98.1	7.0	8.9
B	Algae	*Chlorella variabilis*	DRR149372	21	NC_015359	124	120	-	-	1	-	-	96.1	3.8	2.8
C	Algae	*Ostreococcus tauri*	ERR5421724	8.5	NC_008289	72	184	44	-	3	1	-	98.4	1.5	1.9
D	Algae	*Pycnococcus provasolii*	DRR252953	12	NC_012097	80	80	49	-	1	1	-	99.5	2.2	1.8
E	Monocot	*Asparagus officinalis*	SRR9643837	22	NC_034777	157	159	10	-	1	1	-	98.7	6.7	4.2
F	Monocot	*Deschampsia antarctica*	SRR13908657	33	NC_023533	135	152	26	17	2	1	2	98.9	3.3	3.5
G	Monocot	*Oryza sativa*	DRR196880	42	NC_008155	135	516	37	-	6	1	-	98.5	10.4	8.5
H	Monocot	*Spirodela polyrhiza*	SRR11472010	67	NC_015891	169	512	96	8	2	1	1	99.4	11.8	10.0
I	Dicot	*Aquilaria sinensis*	SRR9858982	54	NC_029243	160	173	197	-	1	1	-	98.3	6.9	8.8
J	Dicot	*Cannabis sativa*	ERR3850904	42	NC_027223	154	153	-	7	1	-	1	98.6	7.2	7.3
K	Dicot	*Corylus avellana*	ERR4852503	26	NC_031855	160	244	-	47	5	-	2	98.5	4.1	4.7
L	Dicot	*Eucalyptus polybractea*	SRR8692273	36	NC_022393	160	159	-	-	1	-	-	98.8	5.0	6.2
M	Dicot	*Gossypium longicalyx*	SRR10377593	41	NC_023216	160	115	-	-	1	-	-	97.9	3.8	5.4
N	Dicot	*Lathyrus sativus*	ERR3374012	67	NC_014063	121	121	78	-	1	3	-	99.2	5.7	7.3
O	Dicot	*Medicago truncatula*	SRR10194526	45	NC_003119	124	124	-	-	1	-	-	99	12.1	8.7
P	Dicot	*Panax ginseng*	SRR13070229	29	NC_006290	156	136	-	14	1	-	2	99.6	3.6	4.3
Q	Dicot	*Prunus dulcis*	ERR3430399	31	NC_034696	158	159	-	-	1	-	-	98.1	7.0	7.2
R	Dicot	*Solanum commersonii*	SRR12407219	46	NC_028069	156	186	-	60	1	-	2	99.4	4.5	3.6
S	Dicot	*Vigna radiata*	SRR12549534	76	NC_013843	151	375	-	26	2	-	1	98.5	11.1	9.8

Acknowledgements

We thank Chelsea Matthews for discussions about Snakemake workflows, which helped us improve on the quality of this work. We thank Professor Brian Atwell, and Dr. James Cowley for helping to review and improve this manuscript. AP acknowledges support from an Australian Research Training Program Scholarship and the FJ Sandoz Scholarship from the University of Adelaide.

References

1. Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S. Complete Structure of the Chloroplast Genome of Arabidopsis thaliana. DNA Res. 1999 Jan 1;6(5):283–90.

2. Teske D, Peters A, Möllers A, Fischer M. Genomic Profiling: The Strengths and Limitations of Chloroplast Genome-Based Plant Variety Authentication. J Agric Food Chem. 2020 Dec 9;68(49):14323–33.

3. Palmer JD, Stein DB. Conservation of chloroplast genome structure among vascular plants. Curr Genet. 1986 Jul 1;10(11):823–33.

4. Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci. 1987 Dec;84(24):9054–8.

5. Morley SA, Nielsen BL. Chloroplast DNA Copy Number Changes during Plant Development in Organelle DNA Polymerase Mutants. Front Plant Sci. 2016 Feb 4;7:57.

6. Dobrogojski J, Adamiec M, Luciński R. The chloroplast genome: a review. Acta Physiol Plant. 2020 May 18;42(6):98.

7. Martín M, Sabater B. Plastid ndh genes in plant evolution. Plant Physiol Biochem. 2010 Aug 1;48(8):636–45.

8. Simmonds SE, Smith JF, Davidson C, Buerki S. Phylogenetics and comparative plastome genomics of two of the largest genera of angiosperms, Piper and Peperomia (Piperaceae). Mol Phylogenet Evol. 2021 Oct 1;163:107229.

9. Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. Use of DNA barcodes to identify flowering plants. Proc Natl Acad Sci. 2005 Jun 7;102(23):8369–74.

10. CBOL Plant Working Group, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, et al. A DNA barcode for land plants. Proc Natl Acad Sci. 2009 Aug 4;106(31):12794–7.

11. Simkin AJ, López-Calcagno PE, Raines CA. Feeding the world: improving photosynthetic efficiency for sustainable crop production. J Exp Bot. 2019 Feb 20;70(4):1119–40.

12. De-la-Peña C, León P, Sharkey TD. Editorial: Chloroplast Biotechnology for Crop Improvement. Front Plant Sci [Internet]. 2022 [cited 2022 Jul 27];13. Available from: https://www.frontiersin.org/articles/10.3389/fpls.2022.848034

13. Freudenthal JA, Pfaff S, Terhoeven N, Korte A, Ankenbrand MJ, Förster F. A systematic comparison of chloroplast genome assembly tools. Genome Biol. 2020 Sep 28;21(1):254.

14. Wang P, Meng F, Moore BM, Shiu SH. Impact of short-read sequencing on the misassembly of a plant genome. BMC Genomics. 2021 Feb 2;22(1):99.

15. Sethi R, Becker J, Graaf J de, Löwer M, Suchan M, Sahin U, et al. Integrative analysis of structural variations using short-reads and linked-reads yields highly specific and sensitive predictions. PLOS Comput Biol. 2020 Nov 23;16(11):e1008397.

16. Wang W, Schalamun M, Morales-Suarez A, Kainer D, Schwessinger B, Lanfear R. Assembly of chloroplast genomes with long- and short-read data: a comparison of approaches using Eucalyptus pauciflora as a test case. BMC Genomics. 2018 Dec 29;19(1):977.

17. Wang W, Lanfear R. Long-Reads Reveal That the Chloroplast Genome Exists in Two Distinct Versions in Most Plants. Genome Biol Evol. 2019 Dec 1;11(12):3372–81.

18. Stein DB, Palmer JD, Thompson WF. Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus Osmunda. Curr Genet. 1986 Jul 1;10(11):835–41.

19. Kim D, Lee J, Choi JW, Yang JH, Hwang IK, Yoon HS. Flip-flop organization in the chloroplast genome of Capsosiphon fulvescens (Ulvophyceae, Chlorophyta). J Phycol. 2019;55(1):214–23.

20. Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, Gouil Q. Opportunities and challenges in long-read sequencing data analysis. Genome Biol. 2020 Feb 7;21(1):30.

21. Ankenbrand MJ, Pfaff S, Terhoeven N, Qureischi M, Gündel M, Weiß CL, et al. chloroExtractor: extraction and assembly of the chloroplast genome from whole genome shotgun data. J Open Source Softw. 2018 Jan 16;3(21):464.

22. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019;37(5):540–6.

23. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol. 2017 Jun 8;13(6):e1005595.

24. Wang D, Wu YW, Shih ACC, Wu CS, Wang YN, Chaw SM. Transfer of Chloroplast Genomic DNA to Mitochondrial Genome Occurred At Least 300 MYA. Mol Biol Evol. 2007 Sep 1;24(9):2040–8.

25. Stern DB, Lonsdale DM. Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. Nature. 1982 Oct;299(5885):698–702.

26. Zhang X, Bauman N, Brown R, Richardson TH, Akella S, Hann E, et al. The mitochondrial and chloroplast genomes of the green alga Haematococcus are made up of nearly identical repetitive sequences. Curr Biol. 2019 Aug 5;29(15):R736–7.

27. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. Genome Biol. 2004;9.

28. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009 Dec 15;10(1):421.

29. Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, et al. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 2020 Sep 10;21(1):241.

30. Coster WD, Rijk PD, Roeck AD, Pooter TD, D'Hert S, Strazisar M, et al. Structural variants identified by Oxford Nanopore PromethION sequencing of the human genome. Genome Res. 2019 Jan 7;29(7):1178–87.

31. Wang L, Qu L, Yang L, Wang Y, Zhu H. NanoReviser: An Error-Correction Tool for Nanopore Sequencing Based on a Deep Learning Algorithm. Front Genet [Internet]. 2020 [cited 2022 Jul 27];11. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2020.00900

32. Sahlin K, Medvedev P. Error correction enables use of Oxford Nanopore technology for reference-free transcriptome analysis. Nat Commun. 2021 Jan 4;12(1):2.

33. Stoler N, Nekrutenko A. Sequencing error profiles of Illumina sequencing instruments. NAR Genomics Bioinforma. 2021 Mar 1;3(1):lqab019.

34. Li W, Liu Y, Yang Y, Xie X, Lu Y, Yang Z, et al. Interspecific chloroplast genome sequence diversity and genomic resources in Diospyros. BMC Plant Biol. 2018 Sep 26;18(1):210.

35. Delahaye C, Nicolas J. Sequencing DNA with nanopores: Troubles and biases. PLOS ONE. 2021 Oct 1;16(10):e0257521.

INTENTIONALLY LEFT BLANK

CHAPTER 7

SUMMARY AND FUTURE DIRECTIONS

Wombo Prompt

“Future rice farming”

Wombo Theme

“Comic”

Thesis Summary

The work presented in this thesis resolves decades-long hypotheses about the extent and mechanisms behind some of the abiotic stress tolerances of *O. australiensis*. I also present the first functionally annotated long read based nuclear genome assembly for *O. australiensis* and a tool for the assembly of chloroplast genomes from long read data.

Short-term photosynthetic heat tolerance in *O. australiensis* has been well established, though there have been no reports of how this species tolerates heat events that extend beyond a few hours. Further, while the mechanisms governing short-term tolerance to extreme temperatures have been worked on, there have been no comprehensive analyses of the physiological underpinnings of heat tolerance in either the short- or long term. In Chapter 3, we reveal for the first time that the short-term heat tolerance observed in *O. australiensis* is also seen after five weeks of exposure to extreme temperatures (45°C). *O. australiensis* was able to maintain growth at 45C because of a robust photosynthetic metabolism while *O. sativa* (domestic rice) suffered severe growth penalties. Importantly, we rule out differences in rates of respiration, CO_2 compensation points and electron transport rates as drivers of the observed phenotypes. We identified what appears to be acclimation of respiration to extreme temperature in *O. australiensis*, though the wild and domestic rice species had similar levels of respiration under sustained high temperature conditions. Differences in growth (both above and below ground) must, we hypothesised have come about due to differences in photosynthetic carbon fixation. While *O. sativa* fixed significantly more CO_2 than *O. australiensis* under control conditions, we found that non-significant changes in *O. sativa* and *O. australiensis* CO_2 assimilation meant that these species were fixing similar amounts of CO_2 at 45°C. These changes occurred even though *O.*

sativa accumulated more Rubisco activase (RCA) protein and *O. australiensis* reduced its RCA pool. We hypothesised that these changes are due to species-dependent differences in the efficiency of RCA at different temperatures, though this is something that we did not test directly.

Researchers have hypothesised for decades that *O. australiensis* should be drought tolerant based on the strong seasonality of rainfall in its native range in northern Australia, even though this perennial grass survives the dry season as a rhizome. Because the traits that govern tolerance to drought in some species are also involved in the acquisition of soil N and because *O. australiensis* occurs in low N soils, it has been hypothesised that *O. australiensis* should also be tolerant to low N soils. In Chapter 4, we present the first robust experiments detailing the response of *O. australiensis* and *O. sativa* to drought, low N soils, and their combined effects on above and below ground phenotypes. We used state-of-the-art non-destructive imaging infrastructure to track the above ground growth of these species over ~30 days of water and N stress. At the conclusion of this experiment, we destructively harvested the root systems to evaluate changes in root architecture. We found that *O. australiensis* demonstrates tolerance to drought and low N early in its development, while *O. sativa* was sensitive to both conditions throughout the imaging period. However, *O. australiensis* was not able to tolerate these stresses beyond ~15 days for drought stress and ~10 days for low N stress. Interestingly, the leaves of *O. australiensis* maintained chlorophyll content under drought throughout the imaging period, and low N soils only reduced leaf greenness after 20 days of growth. *O. sativa*, on the other hand, suffered growth penalties as little as 5 days after the commencement of stress. We found that the roots of *O. australiensis* were far more plastic in their response to the imposed stresses than *O. sativa* roots were, and that the changes observed in *O.*

australiensis roots (increased length, decreased thickness) may have allowed the wild rice to access resources deeper in the soil profile earlier than *O. sativa*. We also found that as NUE increased in *O. australiensis* grown on low N soils, WUE decreased. The reason behind this remains mysterious, especially in the context of the leaf N data.

A high-quality nuclear genome assembly for *O. australiensis* was identified as one of the goals of the IOMAP in 2003. However, since then there have been no contiguous assemblies for this important member of the *Oryza* genus. This is because the relatively high repeat content of the *O. australiensis* genome makes it impossible to collapse sequence fragments into larger contigs. In Chapter 5, we present the first Oxford Nanopore Technologies (ONT) long read based nuclear genome assembly of *O. australiensis*. The use of long reads allowed us to assemble the genome into 850 contigs. We used PacBio Iso-Seq long read data to generate gene models, which were subsequently used to aid in functional genome annotation, which identified 34,587 protein-coding genes and 42,329 transcripts. Reads obtained from individual plants representing four populations, which have different levels of salt tolerance, were used to explore intra- and interspecific genetic variation in the genus *Oryza*. We found that, of all species within the *O. officinalis* genome clade of which it is a member, *O. australiensis* is the most distantly related. We also found substantial variation between the *O. australiensis* read libraries from the four populations. Given the array of abiotic stresses that *O. australiensis* is capable of tolerating and the genetic variation that exists between geographically isolated populations of this species, we expect that the presented reference genome will facilitate the exploration of abiotic stress tolerance in the future and aid in new efforts to introduce domestication traits in this species.

As long-read sequencing technology becomes more accessible, it is being adopted by an increasing number of plant scientists with varying levels of bioinformatics experience to assemble bespoke nuclear genomes and answer questions that were previously intractable given the available technologies. When read libraries derived from photosynthetic tissues are sequenced, researchers inadvertently collect reads from the chloroplast genome. This presents researchers with the opportunity to use these reads to glean additional, perhaps unexpected, information from their sequencing runs. Thus, there is a need for accessible, user-friendly tools that will allow non-expert researchers to take full advantage of their data. In Chapter 6, we present *CLAW* (Chloroplast Long-read Assembly Workflow), an automated Snakemake-based tool for the assembly of chloroplast genomes from reads embedded in larger read libraries. We show that *CLAW* can assemble chloroplast genomes of high quality, comparable to pre-existing chloroplast genomes. We also show that individual genes can be isolated from the assemblies and used for simple phylogenetic analyses. This tool is designed for plant scientists with limited bioinformatics training. It is easy to install, easy to use, and all stages are completed automatically requiring only minimal user input. We expect that this tool will increase the number of publicly available chloroplast genomes as researchers leverage reads from their focal species.

The phenotypic and genomic resources generated in this work provide a platform for the exploration of tolerance traits and their genetic underpinnings in the future. For example, studies into population differences in drought tolerance can make use of the reference genome (and long reads from different populations) along with the chloroplast genome assembly tool to explore differences in genes associated with tolerance (or susceptibility) and/or their promoter regions.

Future Directions

The Limitations of Pot Experiments

We acknowledge that a major limitation of the work presented is that it is all derived from pot experiments. While we could not have completed these experiments as field experiments, we must remain aware that the presented results may have been confounded by several of the issues associated with pot experiments (Passioura, 2006). In Chapter 3, in which we presented the heat tolerance work, it is likely that the soil in the pots of the elevated temperature groups will have been warmer than the soil in the control group. We did not monitor soil temperature, nor did we make any attempt to keep soil temperature constant between experimental groups. Given that soil temperature has dramatic effects on plant growth (Poiré *et al.*, 2010), future research into this area should make use of temperature-controlled troughs or pot holders, such as those found in the rare frost chambers available in Australia, to keep soil temperature constant while changing air temperature. In Chapter 4, where we explore the responses of wild and domestic rice species to individual and combined drought and low soil N, a major part of our investigations was characterising the changes in root architecture. The architecture of the root systems we analysed may have been affected by the relatively small pot size. We report that the length of the average axial root for *O. australiensis* and *O. sativa* was, at most, 100 cm and 50 cm, respectively. Given that the pots used were 20 cm in height, these long axial roots were limited in their ability to grown downwards. It may be that the loss of tolerance to drought and low soil N in *O. australiensis* later in development occurred because its roots were unable to mine for water and N lower in the soil profile. Future experiments should either make use of taller pots with larger volumes, such as those employed by Geetika et al. (2019), or be performed in the field such that roots are not obstructed. However, field experiments would make the phenotyping of belowground architecture difficult.

Recently, The Plant Accelerator located at the University of Adelaide acquired an X-Ray CT scanner that allows researchers to non-destructively phenotype belowground traits. Given that tolerance was detected in *O. australiensis* early in its growth, follow up experiments in the same sized pots could be performed on young seedlings with daily X-ray imaging of the root system.

Understanding Growth Differences Between O. australiensis *and* O. sativa *Under Extreme Heat Stress*

In Chapter 3, we show that neither the above- nor belowground growth of *O. australiensis* was affected by prolonged exposure to extreme temperature, while *O. sativa* suffered above- and belowground growth penalties. However, we also show that at 45°C, *O. australiensis* and *O. sativa* have the same rates of respiration and photosynthesis. Given that both species maintained similar rates of photosynthesis and respiration, one might assume that these plants would have similar biomass. However, this is not the case. Why does *O. sativa* stop growing after three weeks at 45°C, while *O. australiensis* grows unabated? It may be that the species are investing their carbon differently. For example, it is known that *O. australiensis* invests much of its photosynthate into the construction of structural carbohydrates to support a taller canopy, while *O. sativa* invests in starch storage in stems in preparation for grain filling (Mathan *et al.*, 2021). This could explain the observed patterns of growth observed for *O. australiensis* and *O. sativa* at 35°C (i.e., the wild rice seedlings grew to be larger than the domestic rice seedlings). However, the large difference in growth between the species at 45°C suggests that some other mechanism is operating that is limiting the growth of *O. sativa*. It could be that because we neglected to control for soil temperature and because we only measured leaf respiration, the roots of *O. sativa* were respiring at a much higher rate than *O. australiensis*. Given the role of root

exudates in the production of vast quantities of methane in rice paddies, it could also be that carbon was exiting *O. sativa* as root exudates (Aulakh *et al.*, 2004). Future experiments that aim to control soil temperature should help resolve this question. Discussions are underway with potential collaborators at The Australian National University (ANU) to explore the effect of temperature on rates of respiration in the roots of *O. australiensis* using a Q2 oxygen sensor (Astec Global, Maarssen, Netherlands) – a high throughput fluorophore-based sensor originally designed for the measurement of seed respiration but which has since been modified for the measurement of plant organs (Scafaro *et al.*, 2017). To the best of our knowledge, this has never been investigated before. If it is true that the suppression of root respiration in *O. australiensis* allows it to maintain growth at high temperature, this would be another exploitable trait.

Exploring Seedling Level Drought and Low N Tolerance

In Chapter 4, we reported on seedling level tolerance of one accession of *O. australiensis* to drought and low soil N from two independent phenotyping studies. These are exciting new observations that should be investigated further. Seedling tolerance to drought and low N are desirable traits for dry-seeded rice, so future experiments should aim to identify the genes and gene products (e.g., proteins) associated with tolerance traits. The best way to achieve this would be through the identification *O. australiensis* genotypes with contrasting tolerance to drought and/or low soil N. Because variation in salt tolerance has been identified between populations of *O. australiensis* it is expected that there will be population differences in tolerance to drought and/or low soil N too. Initially, a screening experiment should be performed to identify population differences in drought or low N tolerance/sensitivity. We have recently come into possession of 14 *O. australiensis* accessions that could be used for

this work. Screening experiments should also include *O. sativa* for comparison and should focus on the first ~30 DAP to gain a deeper understanding of the response of younger (<15 DAP) seedlings to these conditions. These experiments should pay closer attention to crown root angle, as this is a major determining factor of tolerance to both drought a low soil N. Once an appropriate growth period and tolerant and sensitive genotypes have been identified, proteome and transcriptome profiling of above and below ground tissue could be used to identify significantly differentially expressed genes between populations. The expression of these genes could be augmented in *O. sativa* lines, using plant transformation, to see if they can contribute to improving tolerance in domestic rice. It would be good to investigate the possibility of generating QTL mapping populations to investigate beneficial traits as well.

Constructing a Pan-Genome for O. australiensis

In Chapter 5, we presented the first long-read assembly of the *O. australiensis* genome. While individual genomes are useful, more information can be gleaned from understanding the core genes that define a species as well as the accessory genes that differentiate individuals or populations. A recent *O. sativa* pan-genome has been used to identify over 10,000 novel protein coding genes and allowed for the exploration of genetic variation within and among the diverse accessions of *O. sativa*. The pan-genome for *O. sativa* was able to differentiate various *O. sativa* cultivars from different geographical ranges based on accessory gene content. The same analyses could be used to investigate diverse *O. australiensis* populations. These efforts might aid in the identification of genes that allow different populations of this species to thrive in diverse and adverse niches in the northern Australian savannah. To date, we have collected long and short reads from individual plants representing four *O. australiensis* populations with different levels of salt tolerance (Yichie *et al.*, 2018). The 14 lines

previously mentioned could also contribute to the construction of the pan-genome with further sequencing. Discussions are underway with collaborators from The University of Queensland (UQ) for the construction of a pan-genome for *O. australiensis*. A key limitation of the genome presented in this thesis is the use of the *O. sativa* Nipponbare genome to scaffold the assembled contigs into pseudomolecules. To address this, we considered using Hi-C technology to allow for the construction of a chromosome level assembly for *O. australiensis,* but the technology was financially prohibitive for this work. Hi-C sequencing therefore remains a key objective for future efforts to assemble a more contiguous genome for *O. australiensis*.

Domestication Traits for O. australiensis

Proposals have been made for the domestication of native Australian grass species for the purpose of commercialisation. Diverse ecotypes demonstrating tolerance to local conditions could be employed, which may provide advantages over the use of an introduced grass species (e.g., rice, wheat etc.). In this light, the domestication of *O. meridionalis*, a wild relative of rice whose native range overlaps extensively with *O. australiensis*, has been proposed. Given that *O. australiensis* demonstrates superior thermotolerance compared to *O. meridionalis*, and young *O. australiensis* plants are capable of tolerating salt, drought and low N, and that this species is perennial (which offers benefits for soil health; Zhang *et al.*, 2022), we find it reasonable that this species should also be a candidate for commercialisation. However, there are several factors limiting the adoption of species like *O. australiensis*. Key amongst these is the shattering that occurs both when grain is still developing and when it is fully mature. Even a slight breeze, such as that generated from an opening glasshouse door, is strong enough to dislodge grain from the panicles of *O. australiensis*. Also, the long tillers of this species make lodging a significant issue. These are traits that might be

ameliorated through a dedicated breeding program, however *O. australiensis* is a cleistogamous species – that is, fertilisation primarily occurs prior to the opening of the flower. This offers benefits (e.g., the stable inheritance of traits) in breeding programs once desirable mutations have been introduced, though getting to this point may prove difficult. Instead, future research may aim to transform *O. australiensis* in tissue culture to introduce beneficial traits – a feat that would be aided by current synthetic biology technologies. Now that we have the genome of *O. australiensis*, we will be able to identify and edit genes of interest for, for example, reducing shattering and inducing dwarfism in this species. Recently, it was reported that only three genes/loci (*OsSh1*, *SPR3* and *sh4*) need to be modified to eliminate rice grain shattering (Ishikawa *et al.*, 2022). Thus, there are gene targets for the improvement agronomic traits in wild rice species. However, there remain significant obstacles for this route too. If one were to make a genetically modified *O. australiensis* with improved agronomic traits, there remains the issue of public perception and uptake of the seed by growers. Further, there is currently no reliable method for the transformation of *O. australiensis*. The issue seems to come from limited callus formation in tissue culture, and a resistance to regeneration of plants from the few calli that can be produced. Future work should aim to resolve these issues if significant progress in the *O. australiensis* transformation space is to be realised.

Concluding Remarks

There remains much work to be done in both understanding the tolerances of *O. australiensis* to abiotic conditions and in either transferring these traits to *O. sativa* and/or in domesticating *O. australiensis* itself. Even so, *O. australiensis* is a native Australian grass that offers tremendous potential for the fortification of grain yield – both in rice and other grasses – into the future.

References

Aulakh MS, Wassmann R, Bueno C, Kreuzwieser J, Rennenberg H. 2004. Characterization of Root Exudates at Different Growth Stages of Ten Rice (Oryza sativa L.) Cultivars. Plant Biol (Stuttg) **3**, 139–148.

Geetika G, Oosterom EJ van, George-Jaeggli B, *et al.* 2019. Genotypic variation in whole-plant transpiration efficiency in sorghum only partly aligns with variation in stomatal conductance. Functional Plant Biology **46**, 1072–1089.

Ishikawa R, Castillo CC, Htun TM, *et al.* 2022. A stepwise route to domesticate rice by controlling seed shattering and panicle shape. Proceedings of the National Academy of Sciences **119**, e2121692119.

Mathan J, Singh A, Ranjan A. 2021. Sucrose transport and metabolism control carbon partitioning between stem and grain in rice. Journal of Experimental Botany **72**, 4355–4372.

Passioura JB. 2006. The perils of pot experiments. Functional Plant Biology **33**, 1075–1079.

Poiré R, Schneider H, Thorpe MR, Kuhn AJ, Schurr U, Walter A. 2010. Root cooling strongly affects diel leaf growth dynamics, water and carbohydrate relations in Ricinus communis. Plant, Cell & Environment **33**, 408–417.

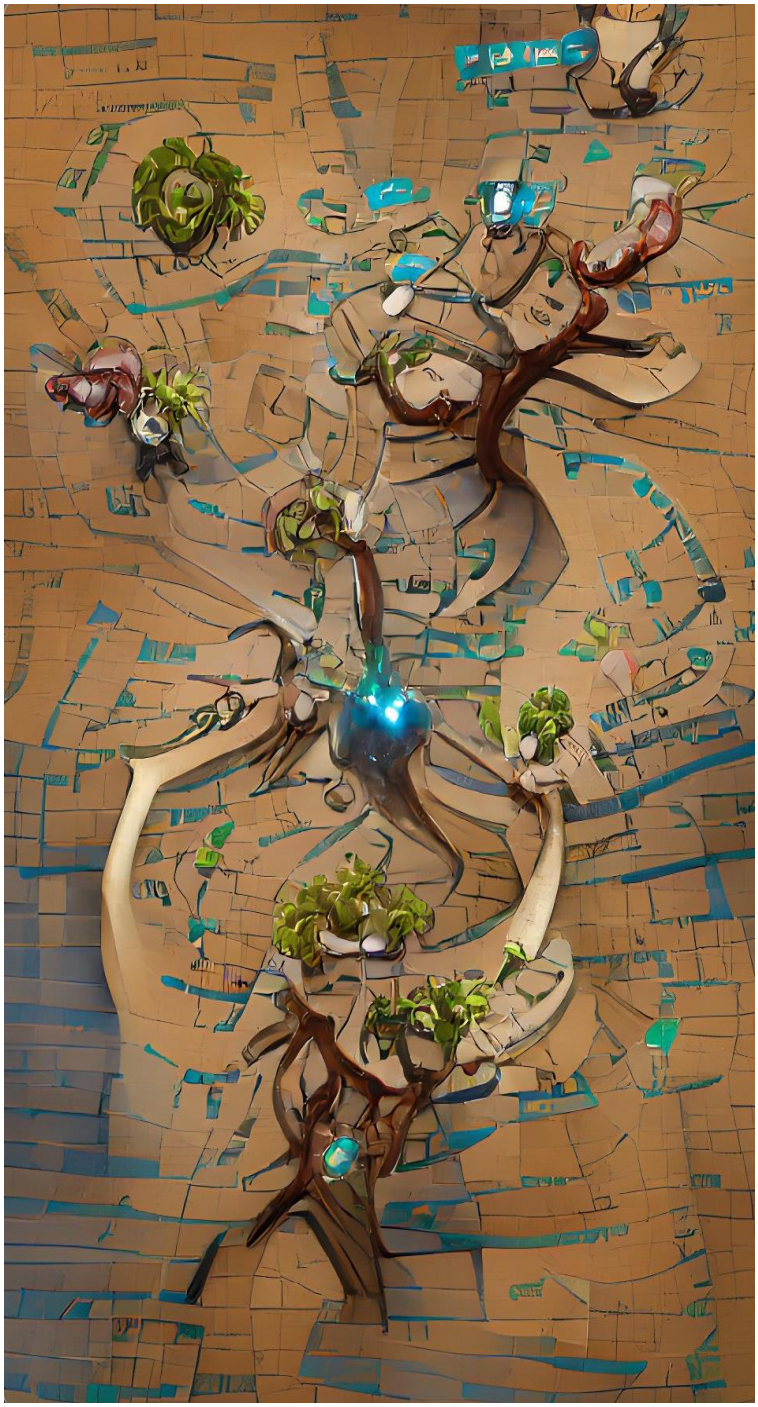
Scafaro AP, Negrini ACA, O'Leary B, *et al.* 2017. The combination of gas-phase fluorophore technology and automation to enable high-throughput analysis of plant respiration. Plant Methods **13**, 16.

Yichie Y, Brien C, Berger B, Roberts TH, Atwell BJ. 2018. Salinity tolerance in Australian wild Oryza species varies widely and matches that observed in O. sativa. Rice **11**, 1–14.

Zhang S, Huang G, Zhang Y, *et al.* 2022. Sustained productivity and agronomic potential of perennial rice. Nature Sustainability, 1–11.

Appendix I

Career and Research Skills Training (CaRST) Accreditation



Wombo Prompt

“Skill-tree progress”

Wombo Theme

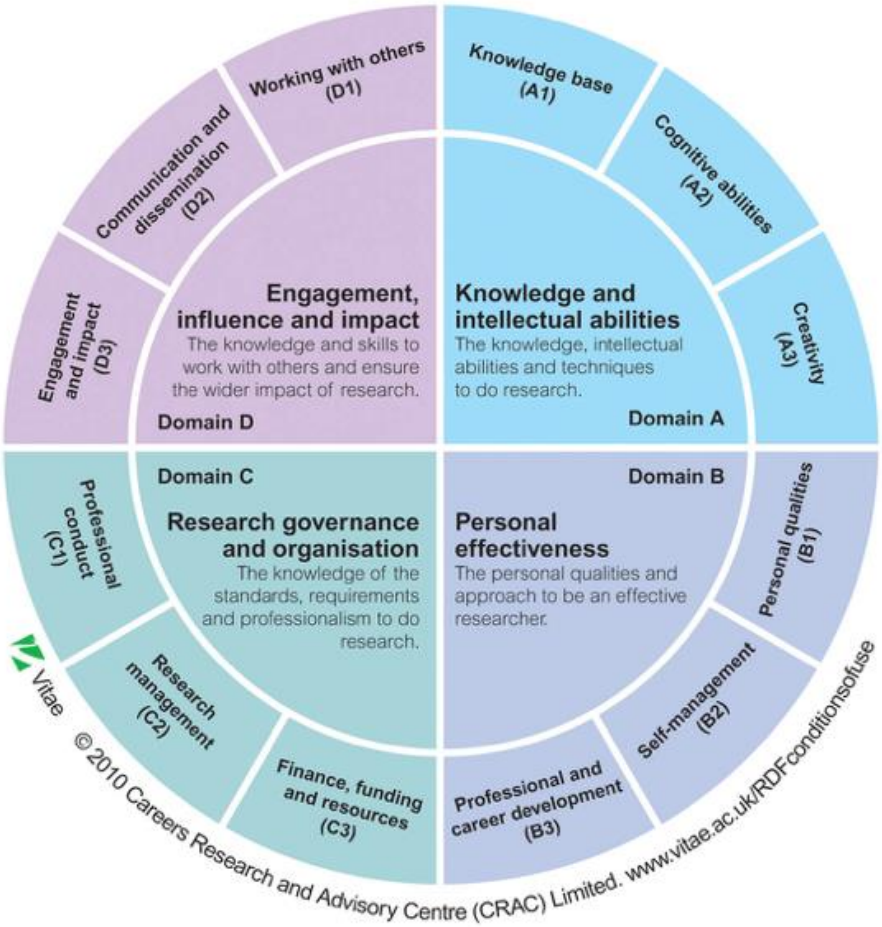
“Provenance”

What is CaRST?

From the CaRST Handbook:

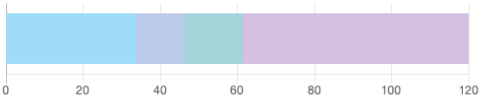
“CaRST is a specialised training and development program for Higher Degree by Research (HDR) students at the University of Adelaide. CaRST was introduced to broaden the research training experience and prepare HDRs for a wide range of careers both within and outside academia. The program aims to facilitate a quality, in-depth research training experience for all HDRs and has been designed to meet the needs of our diverse researcher community.”

HDRs enrolled in a PhD program are expected to complete 120 hours of activities prior to the submission of their thesis. These 120 hours must be spread across four domains, each focusing on a different skill set.



I completed the following activities to meet the requirements of the CaRST program.

CARST CREDITS OVERVIEW



120

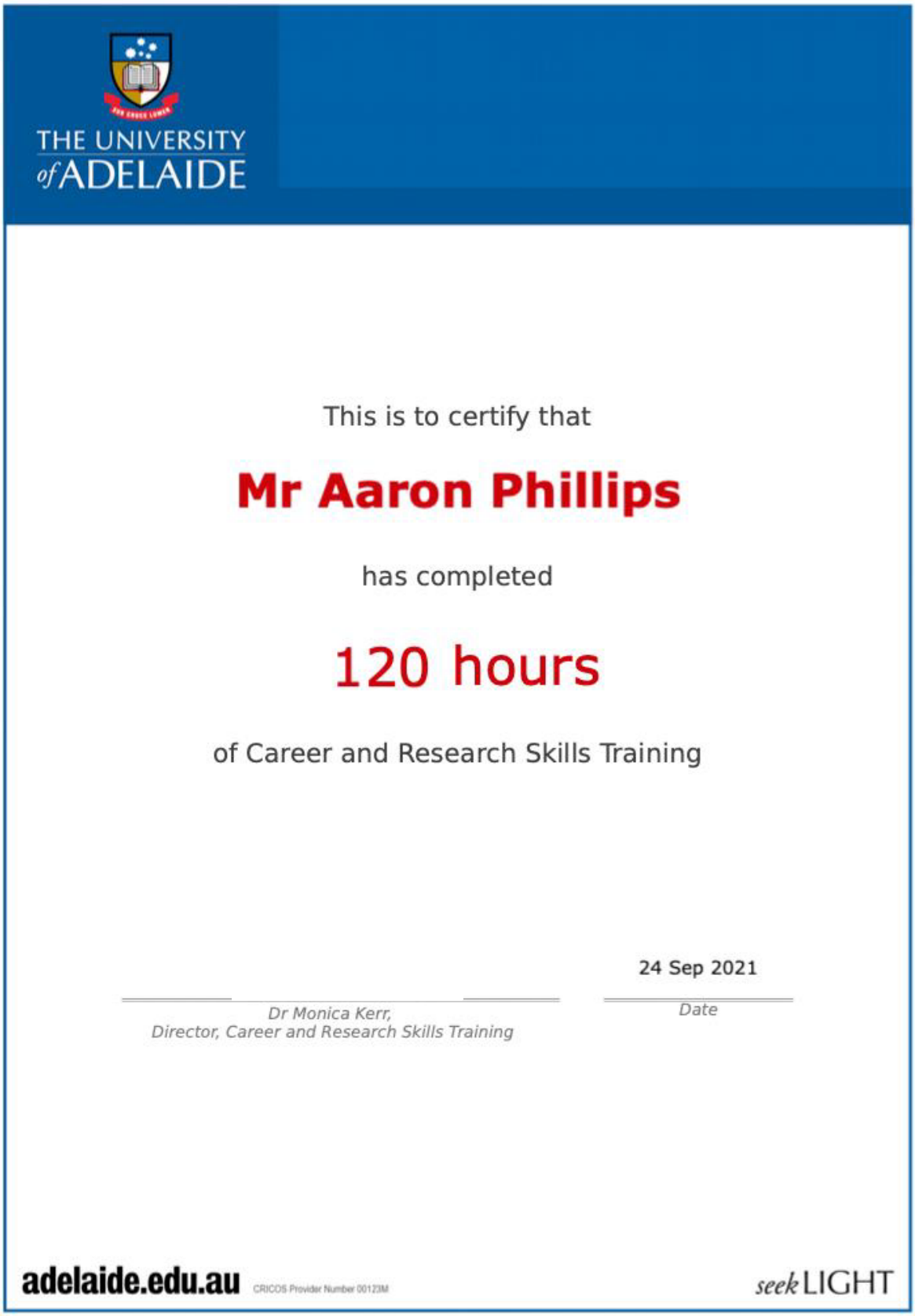
Recognised Credits

Show all

- A: Knowledge and intellectual abilities
- B: Personal effectiveness
- C: Research governance and organisation
- D: Engagement, influence and impact

Date	Title / Summary	Total Credits	Domains
26 Feb 19	PacBio Genome Sequencing Workshop	6.5	Domain A
5 Mar 19	**Foundations in Plant Science (FIPS) II Mentoring - Project Planning**	**6**	**Domain D**
12 Mar 19	**FIPS Mentoring - Mentor and Project Introductory Talk**	**0.5**	**Domain D**
19 Mar 19	**FIPS Mentoring - Running the Project (20 hours)**	**10**	**Domain D**
19 Mar 19	**FIPS II Tutoring (Prac Class)**	**2**	**Domain D**
28 Mar 19	**FreezerPro Biobank Management - Researchers**	**2.5**	**Domain C**
3 Jun 19	**Plant Energy Biology (PEB) Forum - Thredbo**	**12**	**Domain B**
4 Jun 19	**PEB Forum Plenary Presentation**	**6**	**Domain D**
18 Jul 19	Confocal Microscopy Training	3	Domain A
3 Aug 19	**Science Alive! Demonstrating**	**6**	**Domain D**
23 Sep 19	CSIRO's ON Prime Course	30	Domain D
13 Nov 19	**GRDC Research Scholarship Grant Application**	**6**	**Domain C**
29 Nov 19	**APPF Research Scholarship Grant Application**	**6**	**Domain C**
5 Dec 19	**Bioinformatics Workshop - Snakemake Pipelines**	**6**	**Domain A**
22 Jul 20	**The Imposter Syndrome**	**1**	**Domain B**
18 Aug 20	**Research Grants and Fellowship Opportunities (ARC)**	**1**	**Domain C**
23 Jun 21	**ADEPT: Introduction to University Teaching**	**5**	**Domain A**
8 Jul 21	**SAGC 2021 Workshop Series: Single cell RNA-seq analysis**	**7**	**Domain A**
21 Sep 21	**SAGC Epigenomics Workshop**	**7**	**Domain A**

CaRST program certificate



THE UNIVERSITY of ADELAIDE

This is to certify that

Mr Aaron Phillips

has completed

120 hours

of Career and Research Skills Training

24 Sep 2021

Dr Monica Kerr,
Director, Career and Research Skills Training

Date

adelaide.edu.au CRICOS Provider Number 00123M

seekLIGHT

Appendix II

Vacuolar Pyrophosphatase Literature Review

Wombo Prompt

“Plant proton pump”

Wombo Theme

“Backlight”

1.1 Context

Plant growth and yield are multivariate traits, determined by the interaction of various genetic and environmental factors (Kosaki *et al.*, 1989; Bibi *et al.*, 2015). Abiotic stresses (i.e. unfavourable conditions including drought, flooding, saline soils, heat, and chilling) impose a significant yield gap on agricultural systems around the globe and, as a result, realised yield is often far below the potential yield of a crop (Powell *et al.*, 2012). While these systems currently produce enough food and fibre to adequately feed and clothe the global population, crop productivity will have to increase by 100–110% with fewer inputs before 2050 to match increases in demand as the human population exceeds nine billion people (Tilman *et al.*, 2011). This feat is further complicated by the changes in climate observed around the world (Powell *et al.*, 2012). Extremes in temperature (both chilling and heat wave events) and water regimes (i.e. drought and flooding) are increasing in frequency, duration, and intensity, necessitating the generation of crops that are capable of maintaining growth and yield in the face of variable conditions (Frich *et al.*, 2002; Tripathi *et al.*, 2016).

Perhaps the single most important environmental factor contributing to plant performance is the availability of water (Sadras and McDonald, 2012). Too little water (i.e., drought) imposes osmotic stress on plant roots as the concentration of solutes increases in the rhizosphere, and leaf wilting occurs due to loss of turgor pressure ultimately reducing photosynthetic capacity. On the other hand, excess water (i.e. flooding) reduces plant performance by limiting the availability of oxygen to plant tissues, inducing hypoxia or anoxia (Greenway and Armstrong, 2018). While these stresses are different in nature, the net result is reduced availability of ATP and plant cells experience an energy crisis and yield declines (Greenway and Armstrong, 2018). In Australia, yield losses in wheat crops alone grown in water-logged soils costs up to

$300 million per annum (McFarlane and Wheaton, 1990; Dennis *et al.*, 2000). Further, when soils become water logged wheat, barley, and oat crops in Australia can suffer up to 74%, 84%, and 63% yield losses, respectively (Setter and Waters, 2003).

The rise of flash flooding due to increasingly erratic rainfall calls for an improvement in our understanding of flood (and drought) tolerance in grass species. Rice (*Oryza sativa*) is considered to be highly tolerant to hypoxic conditions (Atwell *et al.*, 1982, 2015; Narsai *et al.*, 2015). Different cultivars have evolved different growth strategies in response to hypoxia that aid in their survival (Bailey-Serres *et al.*, 2012). For example, mature FR13A plants suppress growth during total submergence, entering a quiescent period until flood waters recede (Bailey-Serres *et al.*, 2010; Yang *et al.*, 2017), while mature C9285 plants undergo rapid internode elongation in response to anoxia in a bid to grow above the water line (Kuroha *et al.*, 2018). *Oryza sativa* is therefore a good genetic background to investigate tolerance of grasses to low-oxygen environments. Genes and regulatory pathways have been identified in rice that may be useful for improving tolerance in other species (e.g. the SUB1A locus; Bailey-Serres *et al.*, 2010, 2012).

However, rice crops remain susceptible to damage from excess water if their tissues remain hypoxic for more than four days (Greenway *et al.*, 2012). This is particularly true for rice crops planted via the dry direct-seeding method (i.e. dry seed is sown onto dry aerobic soils; Farooq *et al.*, 2011). These crops often produce more grain per hectare than transplanted rice and are usually established in rain-fed areas and so their success is reliant on consistent rainfall (Farooq *et al.*, 2011). The increasingly erratic rains experienced in the tropics increases the risk that young rice crops will be inundated with flash flood waters preceded and/or followed by periods of reduced

water availability. It is therefore important that seedling vigour is enhanced if these crops are to maintain productivity in the face of unreliable rainfall. For example, it may be desirable to increase investment in below-ground biomass during times of water shortage (Wasaya *et al.*, 2018), and increased internode elongation in young tissues either before or during exposure to flood conditions (Kuroha *et al.*, 2018). Not only would this ensure that young plants are more likely to grow above the flood water line, but it would also enable them to outcompete any weedy species invading agricultural soil (which is a detriment of the dry seeding method).

For increased seedling vigour to be realised under these conditions, young plants (particularly pre-photosynthetic plants) would need to allocate scarce energy to cell division/elongation and the synthesis of biomolecules (e.g. DNA, protein, and cellulose) both above and below ground. When ATP supply is limiting and young cells experience an energy crisis (such as is common during drought or flooding), the high concentration of pyrophosphate (PP_i) generated by over 200 biological reactions (including DNA, protein, and cellulose synthesis) can be used as an alternative to ATP to fuel a PP_i-dependent metabolism (Maeshima, 2000*a*; Atwell *et al.*, 2015; Greenway and Armstrong, 2018). Vacuolar pyrophosphatases (V-PPases) are increasingly recognised as key enzymes that respond to cellular energy crises. This family of proton transporters are upregulated in response to a number of stresses, and improve the growth, yield, and abiotic stress tolerance in a number of transformed plant species (Arif *et al.*, 2013). This is not fully understood mechanistically (Schilling *et al.*, 2017). Regardless, V-PPases present as an exciting gene family to not only improve the growth of commercial species under optimal conditions, but to enhance the tolerance of these species to a number of relevant abiotic stresses. Additionally, the elucidation

of the role that V-PPases play in stress tolerance is likely to reveal novel aspects of cellular metabolism.

1.2 Vacuolar Pyrophosphatases in their Native State

General information

Vacuolar pyrophosphatases (V-PPases) are ancient, conserved, simple membrane-bound proton pumps that utilise the energy liberated from inorganic pyrophosphate (PP_i; a cellular 'waste' product) hydrolysis to acidify vacuoles in plant cells (Rea and Poole, 1993; Maeshima, 2000*b*; Seufferheld *et al.*, 2011). In doing so, V-PPases use a 'cheap' energy source to establish an electrochemical gradient across the tonoplast and reduce the cytosolic PP_i pool. Initially, it was thought that vacuolar acidification (i.e. the establishment of the proton gradient across the tonoplast) was the main function of V-PPases, though a recent paradigm shift highlights the importance of PP_i hydrolysis in maintaining normal cellular metabolism (Asaoka *et al.*, 2016). However, the function of V-PPases in non-transgenic organisms may differ to their transformed counterparts due to changes in tissue- and developmental stage-specific expression and intracellular localisation of the V-PPases. Further, the copy number of VPPase genes changes considerably between genomes, and, in non-transgenic plants, each isoform appears to demonstrate some level of tissue- or stress-specific expression (Liu *et al.*, 2010*a*; Zhang *et al.*, 2011*a*).

Classification

Pyrophosphatases (H-PPases) have been found across all three Domains of life (i.e. Bacteria, Archaea, and Eukarya; Drozdowicz *et al.*, 1999; Seufferheld *et al.*, 2011). In animal and prokaryotic cells, the latter of which are typically considered to not possess organelles, H-PPases, and likely vacuolar H-PPases (V-PPases), play a crucial role in the acidification of acidocalcisomes – vacuole-like organelles that serve as rich

repositories of calcium and phosphate, which function in cytosolic pH balance, and ion transport and sequestration (Motta *et al.*, 2008; Seufferheld *et al.*, 2011; Docampo and Huang, 2016). The ubiquity of the V-PPases across the tree of life suggests that an enzyme of similar function evolved before the time of the last universal common ancestor (LUCA; Drozdowicz *et al.*, 1999; Seufferheld *et al.*, 2011) and hints at the importance of pyrophosphate as an energy source throughout the course of evolution.

H-PPases are classified as soluble or insoluble (or membrane-bound; i.e. V-PPases), as well as those that are K^+-dependent (type I) and K^+-independent (Type II; Drozdowicz *et al.*, 2000). Targeted to the Golgi apparatus, type-II H-PPases do not require K^+ as a co-factor for activation and are expressed in low abundance (about 0.3% of the type-I enzyme; Segami *et al.*, 2010). On the other hand, type-I H-PPases target the vacuolar membrane where they make up approximately 10% of the tonoplast proteome and require K^+ for activation (Maeshima, 2001). The K^+-dependency of the V-PPases has raised the question of whether these enzymes might also function in K^+ translocation into the vacuole (Davies *et al.*, 1992), though no evidence has been found in support of this hypothesis (Ros *et al.*, 1995).

V-PPases differ from soluble PPases in that soluble PPases do not show propensity for H^+ translocation. They have therefore been said to ‘waste’ the energy liberated by PP_i hydrolysis – the energy is lost as heat (Perez-Castineira *et al.*, 2002) – whereas the energy of V-PPase-mediated PP_i hydrolysis is stored as a membrane potential across the tonoplast (Maeshima, 2000*b*). However, soluble PPases from yeast (*Saccharomyces cerevisiae*) have been shown to rescue the phenotypes of PP_i-hydrolysis deficient V-PPase mutants by reducing the pool of PP_i in the cytosol (Ferjani *et al.*, 2011). This is despite V-PPases and soluble PPases not sharing any sequence

similarity nor, presumably, a common evolutionary history and serves to highlight to importance of PP_i hydrolysis in maintaining normal cellular function (Drozdowicz and Rea, 2001; Ferjani *et al.*, 2011; Segami *et al.*, 2018*b*).

Structure and function

V-PPases are single polypeptide chains composed of ~780 amino acid residues, with molecular masses of ~80 kDa (Maeshima, 2000*b*). The simplicity of their structure makes them an ideal candidate for genetic transformation work as opposed to more complex transporters, like V-ATPases, the overexpression of which would require multiple transformation events (Rea and Poole, 1993; Maeshima, 2000*b*; Asaoka *et al.*, 2016). The unique structure and function of the V-PPase pump necessitated the designation of a new type of cellular pump distinct from V-, P-, and F-type ATPases. The V-PPase monomer is composed of 16 transmembrane domains (TMDs) forming an inner ring and an outer ring (Lin *et al.*, 2012; Segami *et al.*, 2018*a*). The inner ring is composed of six alpha helices that span the tonoplast and contains the hydrolytic centre and the proton transport pathway. The outer ring is composed of the remaining 10 transmembrane alpha helices (Lin *et al.*, 2012). All 16 of the TMDs extend either side of the tonoplast, interacting with the cytosol and the vacuolar lumen.

Among those regions interfacing with the cytoplasm, there are 15 hydrophilic loops. These mostly contain charged residues that are responsible for the binding of both mineral cations (e.g. Mg^{2+}) and the anionic substrate ($MgPP_i^{2-}$) (Maeshima, 2000*b*; Asaoka *et al.*, 2014; Segami *et al.*, 2018*a*). The cytoplasmic loops are highly conserved across the tree of life. Specifically, three regions of the V-PPase enzyme (CS1, CS2, and CS3) are present in V-PPase isoforms from organisms as distantly related as bacteria and modern land plants (Maeshima, 2000*b*; Asaoka *et al.*, 2014). The charged

residues present in these parts of V-PPases have been identified as critical for substrate (PP_i) hydrolysis. For example, mutagenesis of residues Lys261 and Glu263 (located in CS1) in mung bean significantly reduces the activity (i.e. both PP_i-hydrolysis and H^+ translocation) of V-PPase (Nakanishi *et al.*, 2001). However, the function of the first cytoplasmic loop, designated 'loop a', remains unclear, though it has been proposed to function as a pH sensor on the cytoplasmic side of the tonoplast and shows low sequence similarity between organisms .

Membrane bound V-PPases form a functional dimer via the interaction of TDMs 10 and 16 of two monomers (Fig. 1 A & B; Sato *et al.*, 1991; Lin *et al.*, 2012; Segami *et al.*, 2018*a*). The individual V-PPase monomers seem to function independently, however. Slight rotational movement of the inner ring upon substrate binding facilitates the translocation of a proton from the cytosol to the vacuolar lumen. For this to happen, a series of oxidation-reduction reactions occur along the proton transport pathway (Lin *et al.*, 2012). Acidic residues of TMDs 6 and 16 within the hydrolytic centre form part of the proton transport pathway. The site of PP_i hydrolysis is activated by the presence of Mg^{2+} and K^+. These ions position the active-site residues, PP_i and a hydrolytic water molecule in such a way that reduces the pKa of H_2O-mediated PP_i hydrolysis. The subsequent protonation and de-protonation of residues in the proton translocation pathway results in conformational changes to V-PPase that Lin *et al.* (2012) described as the R (resting), I (initiated), and T (transient) states of V-PPase (Fig. 1C). In the R state, the PP_i binding pocket is in a flexible cytosol-facing open-conformation. Upon binding of PP_i to the hydrolytic centre, a proton interacts with D287 and D731 resulting in the closing of the cytosolic side of V-PPase to prevent PP_i escape (I state). Subsequent hydrolysis of PP_i changes the orientation of these residues and the proton is moved along the acidic residues of TDM 6, ejecting a second proton from an acidic

residue (E301) on the vacuolar lumen side of V-PPase (T state). When the initial proton reduces E301, V-PPase returns to the R state. In this way, the vacuole is acidified and the concentration of PP_i in the cytosol is reduced.

The hydrolysis of PP_i is coupled to proton translocation (Maeshima, 2000*b*). Segami *et al.* (2018*a*) report that the coupled nature of V-PPase results in a H^+ translocation:PP_i hydrolysis ratio of roughly 1. That is, the rate of proton transport (~28 s^{-1} per dimer) and the rate of PP_i hydrolysis (23 s^{-1} per dimer) are similar enough to assume that one proton is pumped for each PPi that is hydrolysed. Incredibly, Segami *et al.* (2018*a*) calculate that this rate of reaction is capable of reducing the pH of a 10 μm (diameter) vacuole from 7.0 to 6.0 in one second (assuming that the abundance of V-PPase on the tonoplast is ~10^5-10^6 dimers). Recently, a number of residues that play important roles in the coupling of PP_i hydrolysis to H^+ translocation have been identified (Asaoka *et al.*, 2014). Mutagenesis of residues in the proton translocating pathway have resulted in V-PPase isoforms that are capable of hydrolysing PP_i but incapable of translocating a proton across the tonoplast (Asaoka *et al.*, 2014). That is, the PPi hydrolysis and H^+ pumping reactions were de-coupled. Despite the reduced capacity for vacuolar acidification, these plants exhibited normal growth, supporting the idea that PP_i hydrolysis is the more important function of V-PPases than vacuolar acidification (Asaoka *et al.*, 2016). Indeed, introducing the un-coupled V-PPase isoform into a V-PPase deficient line of *Arabidopsis* completely rescued the defective phenotype exhibited therein (Asaoka *et al.*, 2016).

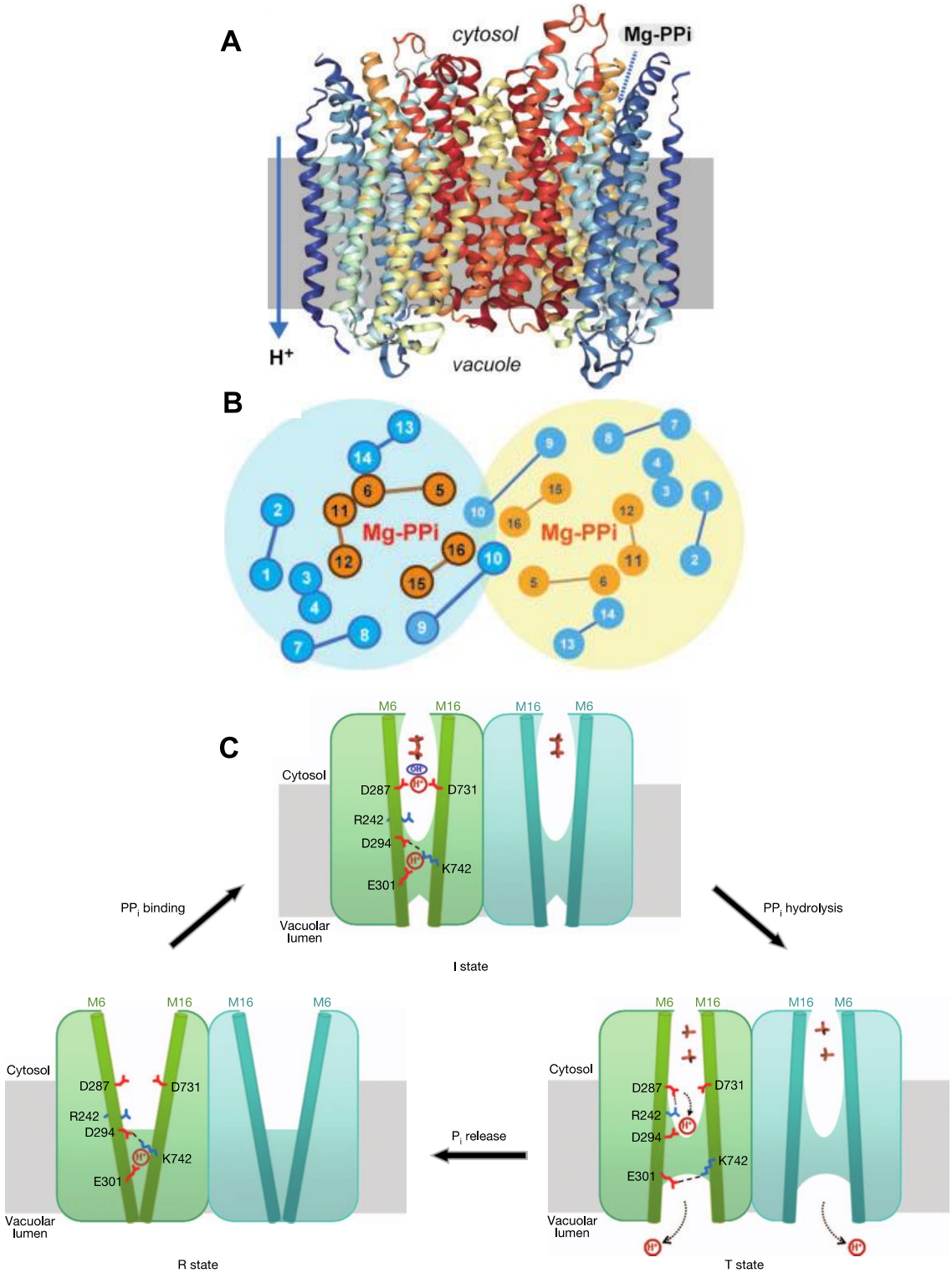


Figure 1. (A) Side view of *Vigna radiata* V-PPase (VVP) quaternary structure (i.e., the V-PPase dimer). (B) Top-down representation of the 16 transmembrane domains of *V. radiata* V-PPase. (C) Working model of *V. radiata* V-PPase PP_i hydrolysis and proton pumping. (A, B) taken from *Segami et al.* (2018*a*); (C) taken from Lin *et al.* (2012).

The importance of PP_i hydrolysis

PP_i is a by-product of a vast number of biochemical reactions in all living cells (Maeshima, 2000*b*; Heinonen, 2001). These reactions include the synthesis of vital biopolymers like DNA, RNA, proteins, sucrose, and cellulose, as well as the synthesis of monomers (e.g. amino acids). It has been estimated that, in spinach leaves, PP_i is produced at a rate of ~15 mM min^{-1} from sucrose synthesis alone (Heinonen, 2001). As a result of its continual production, PP_i must be constantly hydrolysed to avoid the inhibition of the reactions that produce it. In plant cells the hydrolysis of PP_i by V-PPases and partially by soluble PPases maintains cytosolic PP_i at a concentration of ~0.3 mM (Leustek and Saito, 1999; Farquharson, 2018; Segami *et al.*, 2018*b*). Above this concentration, plant growth is reduced, organ development is compromised, and the survival of young tissue is diminished (Ferjani *et al.*, 2011; Fukuda *et al.*, 2016).

Perhaps the plant life stage most vulnerable to excess cytosolic PP_i is post-germinative when photosynthetic machinery is yet to be established and young plants are locked into a period of heterotrophic growth dependent upon the energy reserves of the endosperm (Perata *et al.*, 1997; Shaik *et al.*, 2014). As a result of rapid growth of young tissues, such as coleoptile elongation in rice, and the corresponding production of biomolecules, PP_i production is high (Maeshima, 2000*b*; Asaoka *et al.*, 2016). PP_i catalysis is necessary to prevent the inhibition of gluconeogenesis – sucrose production from endosperm/aleurone lipids and proteins – until the young seedling achieves autotrophic capacity (Ferjani *et al.*, 2011; Takahashi *et al.*, 2017). Ferjani *et al*. (2011) used an AVP1-defective *Arabidopsis* line (*fugu5*) to show that when PP_i concentrations increased during the post-germinative stage, leaf growth and sucrose concentrations decreased but storage lipids remained unchanged, indicating reduced

rates of gluconeogenesis. This phenotype was rescued by the expression of a yeast soluble PPase or by sucrose supplementation in the growth media, which suggests a role of PP_i hydrolysis in sucrose metabolism (Ferjani *et al.*, 2011). Ferjani *et al*. (2018) showed that the check of gluconeogenesis by PP_i comes about by the inhibition of UDP-glucose pyrophosphorylase and thus synthesis of sucrose via UDP-glucose.

Similarly, cell division is inhibited by the accumulation of excess PP_i in the cytosol. Ferjani *et al*. (2011) and Segami *et al*. (2018*b*) showed that in the absence of sucrose supplementation, cell number was reduced in young *fugu5* mutants, but cell size was significantly increased. This is understood as 'compensated cell enlargement', wherein plant growth is driven by the expansion of existing cells rather than by the generation of new cells via division (Hisanaga *et al.*, 2015). There is evidence to suggest that the compensation phenotype is auxin-mediated (Ferjani and Maeshima, 2016; Katano *et al.*, 2016), though Takahashi *et al.* (2017) show that reduced sucrose synthesis from triacylglycerol (TAG; i.e. reduced gluconeogenesis) drives the changes in cell size in young plants. Interestingly, these results were not observed in mature plants with established photosynthetic machinery (and therefore access to an endogenous sucrose supply; Ferjani *et al.*, 2011). Plants overexpressing V-PPases accumulate more biomass but exhibit the opposite cellular phenotype. Mature *Arabidopsis* plants overexpressing AVP1 had lower levels of PP_i in their cytosol, and had smaller but more numerous cells than wild type plants (Asaoka *et al.*, 2016). Taken together, these results highlight the importance of cytosolic PP_i balance for seedling heterotrophic growth via gluconeogenesis and cell division and may have implications for the improvement of seedling vigour (Ferjani *et al.*, 2011; Segami *et al.*, 2018*b*).

Interestingly, it appears that when the *Arabidopsis* V-PPase (AVP1) is targeted to the plasma membrane in phloem companion cells the enzyme switches from exhibiting PP_i hydrolysis behaviour to the synthesis of PP_i (Gaxiola *et al.*, 2015; Pizzio *et al.*, 2015; Khadilkar *et al.*, 2016). This reverse reaction is possible based on the structure of AVP1, which uses the influx of H^+ across the plasma membrane to drive PP_i synthesis (Scholz-Starke *et al.*, 2019). Given the inhibitory effects that excess PP_i exerts on cell growth and metabolism this seems counterintuitive. However, the PP_i synthesis reaction is believed to enhance the glycolytic capacity of a cell by providing substrate to PP_i-dependent variants of glycolytic enzymes (e.g. phosphofructokinase [PFK] and pyruvate orthophosphate dikinase [PPDK]). Key among these is the conversion of sucrose to hexoses via sucrose synthase (SuSy) and uridine diphosphate glucose (UDPG) pyrophosphorylase instead of sucrose invertase, and the subsequent PP_i-driven glycolytic reactions (Huang *et al.*, 2008; Atwell *et al.*, 2015). In this way, PP_i synthesis provides an alternative energy source for glycolysis, facilitates respiration with reduced dependence on ATP and, in fact, increases the ATP yield of glycolysis from 2 to 5 mols (Huang *et al.*, 2008). This ATP can then be used by PM-targeted ATPases (P-ATPases) to acidify the apoplast, which allows for further sucrose/H^+ symporter activity into the cytosol (Schilling *et al.*, 2017).

Additionally, these PP_i-dependent glycolytic pathways may also serve as a ‘back-up’ system for metabolism during times of low ATP supply (Huang *et al.*, 2008; Atwell *et al.*, 2015). For example, when plants experience stress and cytosolic concentrations of ATP are low, cells can shift their metabolism away from the currency of ATP to favour PP_i-driven processes (Huang *et al.*, 2008). This is seen in anoxia-tolerant species, which increase investment in PP_i-dependent glycolytic enzymes in all cell types, not just the phloem companion cells discussed previously, and increase the

expression of V-PPases relative to other tonoplast transporters (Liu *et al.*, 2010*b*; Atwell *et al.*, 2015; Narsai *et al.*, 2015). In this way, the high-energy P-O-P bond of PP_i is utilised in sucrose mobilisation and in vacuolar acidification, allowing energy-deficient cells to maintain cellular processes such as respiration and, as will be discussed below, maintaining membrane potentials. This raises the interesting prospect of altered membrane targeting of V-PPases during times of stress – do cells under stress have PM-localised V-PPases that serve to synthesise PP_i and thus fuel PP_i-dependent glycolytic reactions? If V-PPases can target to the PM, can they target other non-tonoplast membranes and exert influence over other cellular processes (e.g., could chloroplast/mitochondria-targeting V-PPases help establish the proton gradients used to synthesise ATP?)?

The importance of acidifying the vacuole

While it appears that a major function of V-PPase is the hydrolysis of cytosolic PP_i (Asaoka *et al.*, 2016), the proton-pumping nature of the enzyme should not be considered a secondary process. Aside from V-PPases, plants possess another vacuolar proton pump – V-ATPase – that uses the energy from ATP hydrolysis to pump protons across the tonoplast (Maeshima, 2000*b*). The large investment of the energy-rich molecules ATP and PP_i to pumping protons across the tonoplast as well as the considerable costs of producing two different pumps with similar function but distinct evolutionary histories hints at the importance of the process (Taiz, 1992). During times when ATP is in plentiful supply, V-ATPases work with V-PPases to maintain vacuole acidification. ATP hydrolysis by V-ATPases and PP_i hydrolysis by V-PPases establishes a charge difference across the tonoplast and subsequently allows for large ion and solute influx and efflux rates between cellular compartments (Kriegel *et al.*, 2015*a*). As a result, plants can direct and respond to changes in the abundance and

distribution of key metabolites (i.e., sugars and amino acids) and toxic ions or molecules (e.g., Na^+, Cl^- etc.) by remobilising or storing compounds as necessary.

While vacuolar acidification is achieved by the cooperative action of V-ATPases and V-PPases, the relative importance of vacuole acidification by the V-PPases is emphasised during stress (e.g., cold, heat, salt, drought, and flooding stresses). When a cell experiences an energy crisis, defined here as a low ATP:ADP ratio, increased investment in V-PPases allows the cell to maintain charge across the tonoplast (and possibly the plasma membrane; Atwell *et al.*, 2015; Greenway and Armstrong, 2018). The membrane potential generated by V-PPases across the tonoplast is harnessed by secondary transporters to maintain ion and solute homeostasis within the cell by facilitating a reduced, though not insignificant, rate of influx and efflux at the plasma membrane and at the tonoplast (Atwell *et al.*, 2015). Here, the vacuole acts as a battery for the cell when ATP supply is low for prolonged periods. The solutes and ions stored in the vacuole via the action of H^+/ion and H^+/solute antiporters during times of energy abundance can be harnessed by the cell to help keep the charge of the cytoplasm negative. K^+ ions stored in the vacuole leak out of the low energy cell to offset the depolarisation of the PM brought about by H^+ uptake, and the release of stored organic acids neutralises free H^+ while also providing an alternate, though energy poor, substrate for anaerobic respiration (Atwell *et al.*, 2015).

Expression

The number of V-PPase genes in a genome is species-specific, ranging from one type-1 isoform in *Arabidopsis* up to six isoforms in the domestic rice (*Oryza sativa* Japonica) genome. Not only are the V-PPase gene families qualitatively different between species, but the expression profiles of the various isoforms are dynamic between

tissue types within the same species, changing in response to the age of the tissue and to conditions, particularly during episodes of stress (Maeshima, 2000*b*; Liu *et al.*, 2010*b*; Zhang *et al.*, 2011*a*). There is also some evidence to suggest that the role that V-PPases play in cell function depends strongly on the membrane in which it is embedded, which may change depending on the tissue they are expressed in or the stresses a cell is exposed to (Segami *et al.*, 2010; Pizzio *et al.*, 2015).

V-PPases are highly expressed in young tissues undergoing rapid growth such as those in the meristematic regions of mature plants and in developing embryos (Maeshima, 2000*b*; Segami *et al.*, 2014; Asaoka *et al.*, 2016). As discussed above, these high PP_i-producing tissues require enhanced V-PPase activity to prevent the inhibition of biopolymer synthesis and cell division. However, Segami *et al*. (2014) also report the presence of V-PPases in high abundance in cells closely associated with sugar supply. For example, the mesophyll, phloem, nectary, and funicle cells of *Arabidopsis* plants expressing a monomeric GFP-tagged AVP1 had up to 2.5 times the abundance of AVP1 than the highly active, elongating cells of the root (Segami *et al.*, 2014). V-PPases have previously been linked with the activity of phloem companion cells (Gaxiola *et al.*, 2015; Pizzio *et al.*, 2015; Regmi *et al.*, 2016). Interestingly, V-PPase is absent in root columella cells, which contain starchy deposits utilised for gravitropism (Segami *et al.*, 2014). Perhaps the resulting accumulation of PP_i in these cells prevents the catabolism of these starchy granules via the inhibition of α-amylases (Racagni *et al.*, 2008). Taken together, these results indicate an important role for V-PPase activity in young tissues and in sugar metabolism.

V-PPases are capable of targeting non-tonoplast membranes (Gaxiola *et al.*, 2015). Type II V-PPases are known to target the membrane of Golgi apparatus (Segami *et*

al., 2010). However, there is contention over which membranes in the cell type I V-PPase isoforms target (Segami *et al.*, 2018*a*). Recent reports suggest that it is only the vacuolar membrane that houses the V-PPase in Arabidopsis (Segami *et al.*, 2014, 2018*a*), which contradicts the finding of PM-localised V-PPase in phloem companion cells (Paez-Valencia *et al.*, 2011; Pizzio *et al.*, 2015; Regmi *et al.*, 2016). Despite this, the notion of abiotic stresses acting to change the tonoplast-targeting nature of V-PPases has been raised (Schilling *et al.*, 2017).

The expression of V-PPases has been shown to change in response to a variety of abiotic stresses (Maeshima, 2000*b*). Many studies have examined the role that the V-PPase gene family plays during salt stress (Colombo and Cerana, 1993; Fukuda *et al.*, 2004; Guo *et al.*, 2006; Gao *et al.*, 2008). The tonoplast membrane potential generated by V-PPase can be tapped by Na^+/H^+ antiporters for ion sequestration, reducing the concentration of toxic ions in the cytosol. However, V-PPases have been shown to respond to low nutrient availability (Kasai *et al.*, 1998), toxic ion exposure (Yang *et al.*, 2018), cold stress (Carystinos *et al.*, 1995; Zhang *et al.*, 2011*a*), drought stress (Venancio *et al.*, 2014; Wang *et al.*, 2016), and the hypoxic/anoxic conditions induced by flooding (Carystinos *et al.*, 1995; Liu *et al.*, 2010*a*). It is thought that stress-induced changes to the availability of ATP causes a shift in metabolism to PP_i-dependent processes (see above). This is advantageous because the pool of PP_i in the cytosol is stable even during stress, while the pool ATP is not. For species with multiple V-PPase isoforms, the changes in expression involve both the upregulation and downregulation of distinct isoforms. For example, Liu *et al.* (2010*a*) show differential expression patterns of the six *O. sativa* V-PPases following exposure to anoxia. In this study, the expression of OVP3 was substantially increased, while other isoforms, for example OVP1 and OVP2, had decreased expression (Liu *et al.*, 2010*a*). Analysis of the

promoter regions of this gene family revealed that the upstream region of OVP3 contained more anaerobic cis-acting elements than the other OVPs (Liu *et al.*, 2010*b*). Differential expression based on unique promoter elements in the upstream regions of V-PPases is supported by Zhang et al. (2011*a*), who found that OVP1 increased in expression during cold stress specifically. The response of V-PPase isoforms across a multitude of species to stress conditions strongly implicates this gene family in multi-stress tolerance.

1.3 Improved Growth, Yield, and Stress Tolerance in Transgenic Plants Expressing Vacuolar Pyrophosphatases

It is clear that V-PPases are intimately linked to both normal cellular development (i.e., via influencing leaf development and regulating biopolymer synthesis and rates of cell division) and responses to stress. The multiple-stress tolerance achieved in V-PPase overexpressing lines points to the fact that this gene family is likely to be a general responder to reduced energy availability in cells (Greenway and Armstrong, 2018). As such, V-PPases have been the focus of a number of transgenic studies aiming to improve plant performance. However, it is unknown whether the role that V-PPases play in transgenic plants is the same as in wild type lines. For example, it has been hypothesised that V-PPase targets different membranes when the enzyme is hyper-accumulated in a cell (Schilling *et al.*, 2017), and V-PPases are present on the Golgi membrane (Mitsuda *et al.*, 2001; Segami *et al.*, 2010). Gaxiola *et al*. (2015) show that the role of V-PPases changes depending on membrane environment, so there is scope for novel influence over cellular physiology if the V-PPases target non-tonoplast membranes. Several studies have reported increased above- and below-ground biomass accumulation and yield of transgenic plants overexpressing V-PPases under both optimal and stress-inducing conditions (reviewed by Schilling *et al.*, 2017; and

Gutiérrez-Luna *et al.*, 2018). It is conceivable that the overexpression of V-PPases enhances growth in transgenic plants due to the fact that vacuolar acidification is capped – that is, the pH of the vacuole can only be driven so low. Increased expression of V-PPases may allow a cell to achieve maximal vacuolar acidification without the need to invest in V-ATPases, thus conserving energy for other processes. Importantly, the overexpression of V-PPases improves the growth and yield of taxonomically diverse and economically important species, highlighting the potential utility of this gene in fortifying yields in a number of dicot and monocot crops (Arif *et al.*, 2013).

Increased tolerance of stress is often measured as improvements in growth and/or yield (Arif *et al.*, 2013). Growth improvements during stress conditions have been hypothesised to come about via changes in both PP_i hydrolysis and by vacuolar acidification (reviewed in Schilling *et al.*, 2017). Increased sucrose mobilisation in metabolically active cells and improved gluconeogenic capacity in young tissues driven by PP_i hydrolysis leads to increases in shoot and root biomass in transgenic organisms overexpressing V-PPases. Relatively low levels of PP_i in the cytosol of overexpression lines facilitates cell division, so these organisms often have more numerous but smaller cells. Likewise, increased vacuolar acidification resulting from V-PPase overexpression improves growth rates of transgenic plants by facilitating solute accumulation in the vacuole (of both beneficial nutrients and toxic ions), and there is some evidence suggesting that increased auxin mobilisation leads to increased shoot and root growth (Schilling *et al.*, 2017). Li *et al.* (2005) proposed that the over-expression of AVP1 leads to increased cell division via pleiotropic effects on the abundance and activity of plasma membrane-targeted ATPase (P-ATPase) and PIN1 (an auxin efflux protein) facilitating increased auxin transport around the plant. Additionally, an Arabidopsis V-PPase knock out line (*avp1-1*) exhibited developmental

flaws, had reduced size, and reduced auxin transport (Li *et al.*, 2005). Together, these results suggest a link between vacuolar acidification and auxin transport. However, the involvement of auxin in the phenotypes of mutant (either over-expressing or knockout) lines has since been called into question because no auxin-related phenotypes were observed in *fugu5* mutants, and whole-genome sequencing of the *avp1-1* mutant revealed that the T-DNA insertion interrupted GNOM, a gene involved in auxin transport (Kriegel *et al.*, 2015*b*). As discussed above, another line of thinking suggests that the acidified vacuole may serve as a battery for an energy deficient cell (Atwell *et al.*, 2015). Regardless of the cause, a larger root system may allow for increased water uptake and retention and increase the soil-mining capability of the plant, while a larger shoot system is able to support increased rates of carbon fixation.

Interestingly, some V-PPase overexpressing organisms only exhibit improved growth during times of stress (Arif *et al.*, 2013). For example, cotton plants overexpressing AVP1 show improved salt and drought tolerance, exhibiting 20% larger fibre harvests in glasshouse experiments (Pasapula *et al.*, 2011). Importantly, in the field these transgenic plants outperformed their WT counterparts (Pasapula *et al.*, 2011), further highlighting the utility of this gene for crop fortification against abiotic stresses. When grown on phosphorous deficient soils, transgenic tomato plants overexpressing AVP1 had 25% more fruit than the WT tomato (Yang *et al.*, 2014). Finally, Schilling *et al*. (2014) showed that the overexpression of AVP1 in barley improved the yield of barley plants grown on saline soils.

Osmotic stress – salinity and drought

There is a great deal of evidence for increased salt tolerance in a diverse range of plant species overexpressing either endogenous or foreign V-PPase genes (Silva and

Gerós, 2009). For example, salt tolerance is observed in *Arabidopsis*, tomato, barley, bacopa (or Brahmi), tobacco, creeping bentgrass, alfalfa, poplar and cotton overexpressing V-PPases from diverse sources (Gao *et al.*, 2006; Guo *et al.*, 2006; Brini *et al.*, 2007; Duan *et al.*, 2007; Bao *et al.*, 2009; Li *et al.*, 2010; Bhaskaran and Savithramma, 2011; Liu *et al.*, 2011; Pasapula *et al.*, 2011; Zhang *et al.*, 2011*b*; Schilling *et al.*, 2014; Yang *et al.*, 2015; Ahire *et al.*, 2018). Plants and yeasts overexpressing AVP1 have improved tolerance to salt stress because they have increased capacity for Na^+ transport into the vacuole (Gaxiola *et al.*, 1999, 2001; Duan *et al.*, 2007). Additionally, the cooperative action of V-PPases and ion transporters (e.g. the Na^+/H^+ antiporters that exchange protons for Na^+ ions), important for salinity tolerance, on the tonoplast membrane is exacerbated in *Arabidopsis*, rice, cotton, and tomato plants overexpressing these genes in tandem (Zhao *et al.*, 2006; Brini *et al.*, 2007; Bhaskaran and Savithramma, 2011; Shen *et al.*, 2015; Cheng *et al.*, 2018). The increased membrane potential generated by the V-PPases provides a strong gradient that the Na^+/H^+ antiporters can tap for increased Na^+ transport capacity. Thus, the transgenic organisms (either those overexpressing V-PPases alone or those co-expressing V-PPases and ion transporters) avoid the toxic effects of excess cytosolic Na^+. Graus *et al*. (2018) report that *Nicotiana benthamiana* plants transiently overexpressing NbVHP (*N. benthamiana* endogenous V-PPase) exhibited leaf necrosis in the absence of salt stress. The authors report that this was due to impaired cellular pH homeostasis driven by changes in the vacuolar pH (Graus *et al.*, 2018). This result may suggest that increased vacuolar acidification is only useful when the membrane potential is being tapped by other transporters, though this idea is yet to be tested. Do V-PPase overexpressing plants have increased expression of other tonoplast transporters that utilise the vacuole-cytosol pH gradient for ion translocation? Studies into salt tolerance following overexpression of V-PPases often report

concurrent improvements in drought tolerance in the same transgenic lines mentioned above (Liu *et al.*, 2010*c*). This suggests that the overexpression of V-PPases assists plants in tolerating the osmotic stress, common to both drought and salt stress, by improving water retention and root growth.

Low temperature stress

Exposure to low temperatures reduces the integrity of lipid membranes in cells, leading to increased solute leakage and reduced mitochondrial activity (Lyons and Raison, 1970; Zhang *et al.*, 2011*a*). As a result, plant cells become deficient in ATP and the expression of genes involved in fermentation pathways (e.g. alcohol dehydrogenase) increase (Kato-Noguchi, 2008). The low ATP status of the cell leads to the upregulation of PP_i-dependent enzymes including those involved in glycolysis (discussed above) as well as V-PPases (Greenway and Armstrong, 2018). For example, Carystinos *et al*. (1995) showed that the expression and activity of V-PPase increased in *O. sativa* (domestic rice) during cold stress as V-ATPase activity fell. These results are corroborated by Zhang *et al*. (2011*a*), who showed that the expression of a single V-PPase isoform (OVP1) was upregulated in rice following chilling. The authors also report that the overexpression of OVP1 lead to improved membrane stability and increased proline accumulation during exposure to cold stress (Zhang *et al.*, 2011*a*). Improved membrane stability in overexpression lines may have come about due to increases in soluble carbohydrate concentration because of the close link between sucrose synthesis and cytosolic PP_i concentrations (Ferjani *et al.*, 2018).

Low oxygen stress

Hypoxia (reduced access to oxygen) and anoxia (complete absence of oxygen) reduce the synthesis of ATP via respiration by reducing pyruvate flux into mitochondria and

reduce the efficiency of the respiratory electron transport chain as oxygen is the terminal electron acceptor in the chain (Zabalza *et al.*, 2008). As a result, the pH of the cytosol decreases (a phenomenon referred to as cytosolic acidosis) and anoxia tolerance enzymes are synthesised (Greenway and Armstrong, 2018). Interestingly, it is likely that only a very small number of proteins are upregulated in response to anoxia specifically – the majority are involved in response to a multitude of stresses (Greenway and Armstrong, 2018). Key among these are isoforms of the V-PPase family because survival in hypoxic conditions hinges on maintaining membrane potentials (Carystinos *et al.*, 1995; Liu *et al.*, 2010*a*; Atwell *et al.*, 2015; Greenway and Armstrong, 2018). However, there are no published reports thus far (to the best of our knowledge) exploring the effect of V-PPase overexpression on the tolerance of species to hypoxia/anoxia. In the absence of this work, it is important to note the similarities of low-oxygen responses to, for example, the chilling response (see above). The stark reduction in ATP synthesis and reduced reliance on ATP-requiring reactions during exposure to hypoxia/anoxia makes this stress ideal for studying the role of the V-PPases in general low-energy stress. In fact, Greenway and Armstrong (2018) suggest that as our knowledge of energy-limitations in aerated tissues increases the terms hypoxia/anoxia used herein can be understood more generally as an 'energy crisis' rather than an oxygen-specific phenomenon.

1.4 Rice Vacuolar Pyrophosphatases – the OVP Family

Gene family

Rice is unique in that is highly tolerant to hypoxic conditions (Atwell *et al.*, 1982) and it has a large V-PPase gene family relative to other species, comprised of six type I isoforms (Table X; Sakakibara *et al.*, 1996; Choura and Rebaï, 2005; Liu *et al.*, 2010*a*). In comparison, Arabidopsis has one type I isoform, but possesses two type II isoforms

(Sarafian *et al.*, 1992; Drozdowicz *et al.*, 2000; Asaoka *et al.*, 2016). Wheat has five V-PPase isoforms per genome (total of 15 isoforms when considering the A, B, and D genomes; Brini *et al.*, 2005; Wang *et al.*, 2009), while two isoforms have been identified in cucumber (Kabala *et al.*, 2014), and four in grape (Sun *et al.*, 2017). All six rice OVPs are likely to be type I V-PPase based on sequence homology with AVP1 and dissimilarity with AVP2 (Fig. 2). This means that the activity of these isoforms is probably dependent on the presence of cytosolic K^+, and that they likely target the vacuolar membrane (Segami *et al.*, 2010).

Table 1. Copy number of type I V-PPase genes identified in the indicated species' genomes or transcriptomes (sorted highest to lowest copy number). As far as we are aware, rice (*O. sativa*) possesses the highest copy number of V-PPase genes ***per genome*** compared to all other species known to date.

Species	Genome Size (Mbp)	Likely Copy Number	Reference
Triticum aestivum	17000	15*	(Brini *et al.*, 2005; Wang *et al.*, 2009; Menadue *et al.*, 2018)
Oryza sativa	500	6	(Sakakibara *et al.*, 1996; Liu *et al.*, 2010*a*)
Nicotiana tabacum	4500	4	(Lerchl *et al.*, 1995)
Vitis vinifera x V. labrusca	500	4	(Sun *et al.*, 2017)
Halostachys caspica	NA	3	(Wang *et al.*, 2013)
Sophora alopecuroides	NA	3	(Wang *et al.*, 2013)
Beta vulgaris	714-758	2	(Kim *et al.*, 1994)
Cucumis sativus	195	2	(Kabala *et al.*, 2014)
Glycyrrhiza uralensis	379	2	(Wang *et al.*, 2013)
Hordeum vulgare	5300	2	(Tanaka *et al.*, 1993; Fukuda *et al.*, 2004)
Karelinia caspia	70.1	2	(Wang *et al.*, 2013)
Suaeda salsa	NA	2	(Wang *et al.*, 2013)
Arabidopsis thaliana	NA	1	(Sarafian *et al.*, 1992)
Glycyrrhiza inflata	NA	1	(Wang *et al.*, 2013)
Iris lactea	NA	1	(Meng *et al.*, 2017)
Malus domestica	750	1	(Yao *et al.*, 2011)
Suaeda corniculata	NA	1	(Liu *et al.*, 2011)
Suaeda rigida	NA	1	(Wang *et al.*, 2013)

**Triticum aestivum* is comprised of three sub-genomes (A, B, and D), each of which contain five V-PPase isoforms (Menadue *et al.*, 2018).

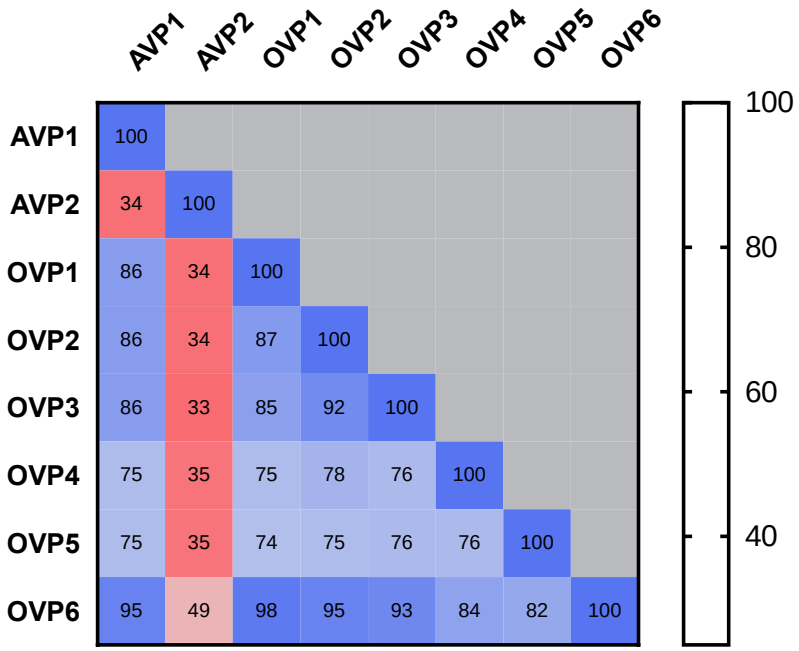


Figure 2. Similarity matrix of amino acid sequences (%) of vacuolar pyrophosphatases found in *Arabidopsis* (AVP1 and AVP2; P31414 and Q56ZN6, respectively) and *O. sativa* (OVP1-6). The AVP sequences were obtained from UniProt, and the OVP amino acid sequences were obtained by BLASTing the OVP1 (D45383) sequence against the *O. sativa* genome in Gramene with default settings. Identified sequences were aligned using the MUSCLE algorithm and the similarity tool was used to obtain the percent similarities of the amino acid sequences of interest in UGENE v1.31.1.

The six OVP isoforms are differentially regulated across tissue types, plant age and in response to various stress types (Carystinos *et al.*, 1995; Liu *et al.*, 2010*b*; Zhang *et al.*, 2011*a*). For example, Liu *et al.* (2010*b*) demonstrate that OVP1 is highly expressed in young leaf blades, but has low expression in embryonic tissue, where the expression of OVP3 is relatively high (Fig. 3) and compared to mature tissue, the expression of the OVPs in young tissue is relatively high. Further, the authors show anoxia-specific upregulation of OVP3, which, despite evidence suggesting a role for OVP1 in response to anoxia, is not observed for the other isoforms (Liu *et al.*, 2010*b*; Zhang *et al.*, 2011*a*). Differences in the promoter regions of these genes likely drives the unique expression patterns in response to stress (Sakakibara *et al.*, 1999; Liu *et al.*, 2010*b*) and tissue-specific expression. For example, Lerchl *et al*. (1995) achieved phloem-specific expression of soluble PPases in Arabidopsis through the use of a phloem-specific

promoter. However, the function of the different OVP isoforms in different tissue types and in response to different stresses is not well understood.

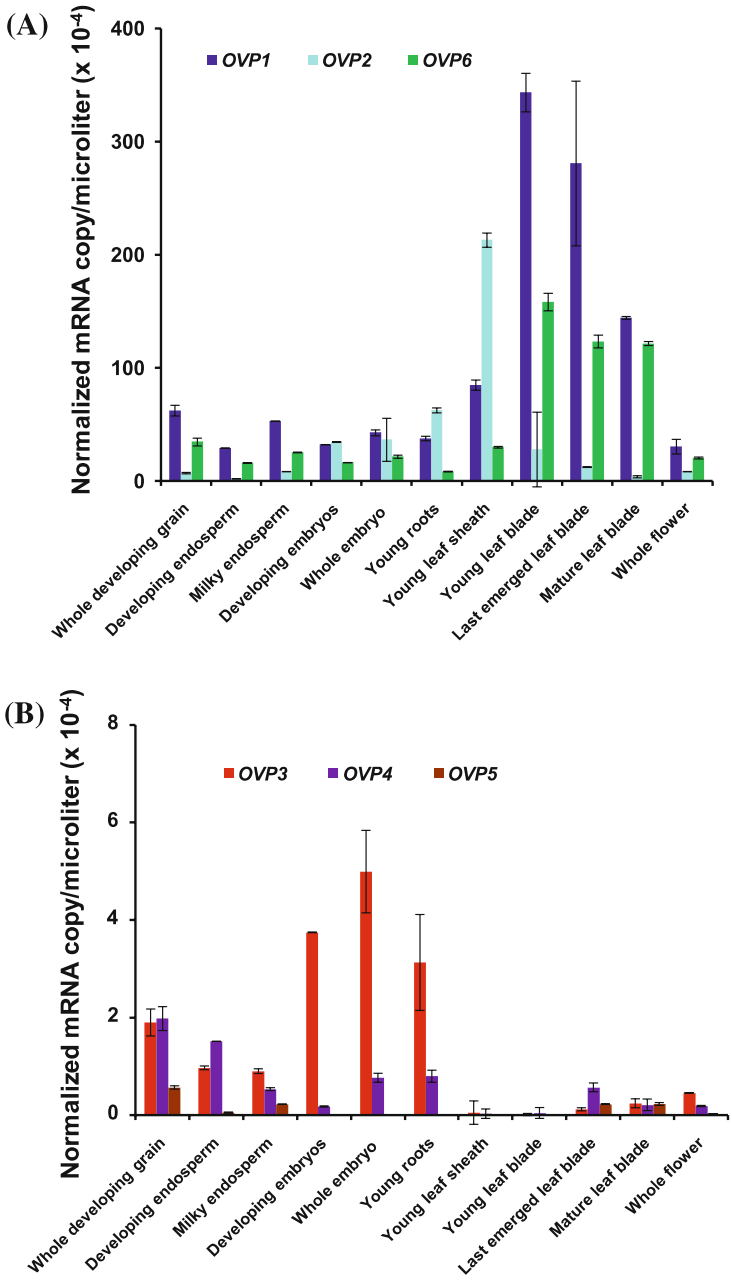


Figure 3. OVP gene family transcript levels in different rice (*O. sativa* cv. Calrose) tissues grown under optimal conditions. (A) Shows isoforms 1, 2, and 6 – the most abundant OVP isoforms – and (B) shows isoforms 3, 4, and 5, which are present in lower abundance. Taken from Liu et al. (2010*b*).

Isoform functions – a case of neofunctionalisation?

There are a limited number of studies looking at the effect of OVP overexpression on the growth and physiology of plants under both optimal conditions and in response to

stress. This paucity in research makes it difficult to draw conclusions on the practicality of maintaining six distinct isoforms of OVP in the *O. sativa* genome. The loss of gene isoforms from genomes is a common occurrence in species that have undergone duplication events, whether that be whole genome duplications or the duplication of individual genes (Ohno, 1970). Following a duplication event, a gene can be maintained in the genome and not undergo any change, be maintained but take on new functions (termed 'neofunctionalisation') or be lost from the genome (Ohno, 1970). While it is impossible to say what the exact function of each isoform is currently, there is evidence to suggest that the different OVPs are functionally divergent (Liu *et al.*, 2010*b*; Zhang *et al.*, 2011*a*). For example, Zhang *et al*. (2011*a*) show that the overexpression of OVP1 increases the tolerance of *O. sativa* to cold stress, though there is no effect of increased OVP1 expression on the tolerance of *O. sativa* to saline conditions. However, the overexpression of OVP1 increased the tolerance of *O. sativa* to both saline conditions and drought, and this tolerance was further exacerbated by the co-overexpression of a Na/H^+ antiporter (Liu *et al.*, 2010*c*). Taken together, these results, along with the increased expression of OVP3 in response to anoxia (Liu *et al.*, 2010*b*), may indicate divergent functions of the OVP isoforms in response to different stress types. If this is true, then the following questions arise: how is tolerance to diverse stress types mediated by distinct OVP isoforms? If it is simply the acidification of the vacuole or the increased hydrolysis of cytosolic PP_i driving normal cellular development and/or stress tolerance, why do OVPs show distinct patterns of expression in different tissue types and in response to different stresses (i.e., OVP1 responds to cold stress, but OVP3 does not)?

OVP3 as a route to improving seedling vigour, growth, and yield of rice

There is considerable evidence in the literature for increased growth rates and yield of plants overexpressing V-PPases under both optimal conditions and conditions that would ordinarily impose an energy crisis (Arif *et al.*, 2013). Increases in growth and yield likely come about due to improved solute (particularly sucrose) mobilisation and metabolism, especially in young tissues in which excess PP_i is removed from the cytosol, while improved tolerance to stresses likely comes about due to increased vacuolar acidification where the membrane potential serves as a 'battery' for a low-energy cell (Atwell *et al.*, 2015). It has been hypothesised that the augmentation of these processes may enhance seedling vigour – the probability that a seed crop is successfully established and undergoes normal post-germinative growth – during times of stress (Ferjani *et al.*, 2011; Schilling *et al.*, 2017). This would be particularly desirable for rice crops established by dry seeding, which can be susceptible to unreliable rainfall. With the increasingly erratic nature of rainfall patterns across the globe due to climate change there is a need to fortify these crops against exposure to drought and flooding conditions associated with changing rainfall patterns.

Given the increased expression of OVP3 (and, to a lesser degree, OVP1) during hypoxia in tolerant tissues (Carystinos *et al.*, 1995; Liu *et al.*, 2010*a*), and the common thread of reduced ATP availability during an energy crisis, it is conceivable that the overexpression of these isoforms would reduce the susceptibility of young rice seedlings to erratic rain via enhanced seedling vigour mediated by improved energy use efficiency (Atwell *et al.*, 2015; Greenway and Armstrong, 2018). This is a major area of research that remains to be investigated. The high tolerance of young rice plants to hypoxic conditions (Atwell *et al.*, 1982; Greenway *et al.*, 2012 report up to 4 days survival in anoxic conditions), the promise of increased multi-stress tolerance

from V-PPase overexpression in a range of species (Arif *et al.*, 2013), and the expression pattern of OVP3 in young, fast growing tissues and in response to anoxia (Liu *et al.*, 2010*b*) makes OVP3-overexpressing rice an excellent model for exploring changes in cellular metabolism during an energy crisis brought on by erratic rain and other abiotic stresses (Atwell *et al.*, 2015).

1.5 Exploiting Wild Germplasm

Another avenue of investigation that has not been explored in the study of V-PPases is the presence of novel isoforms in the germplasm of crop wild relatives. Wild crop relatives represent a vast reservoir of novel genetic material that can be used to improve the tolerance of domestic varieties to abiotic stressors (Dempewolf and Guarino, 2015; Kaur *et al.*, 2018; Mammadov *et al.*, 2018). Many, if not all, crop species have suffered varying degrees of genetic erosion through the intensive breeding processes that have given rise to modern high-yielding cultivars. For example, when compared with wild species the diversity of the soybean (*Glycine max*) genome is reduced by ~50% and the presence of rare alleles reduced by ~60% (Hyten *et al.*, 2006). Such losses in variation render agriculturally important species susceptible to environmental stresses. While different cultivars of the same species can show variation in traits associated with tolerance to extreme conditions, it is much more likely that the functional diversity needed to improve modern cultivars will be found in wild species adapted to marginal environments where local adaptation to extreme conditions has occurred (Atwell *et al.*, 2014; Dempewolf and Guarino, 2015). For example, thermotolerant orthologs of rubisco activase (RCA) have been identified in *O. australiensis* and *O. meridionalis*, wild relatives of rice native to the Northern Territory of Australia (Scafaro *et al.*, 2016). Given the importance of V-PPases in the response of domestic rice to various types of abiotic stress, it stands to reason that

wild rice relatives that are adapted to extreme environments may possess novel variants with unique patterns of expression or activity.

In order to fully take advantage of wild *Oryza* species as sources of novel genes it is important to understand these species in a genomic context (Jacquemin *et al.*, 2013). Recently, Stein *et al*. (2018) compiled reference genomes of 13 *Oryza* species. This study considered ten species/accessions of rice that belong to the AA genome subclade of *Oryza*, one BB genome species, and one FF genome species. In doing so, Stein *et al.* (2018) identified new genetic information that may be of use in the fortification of rice yield. The identification of novel genetic variation within the AA subclade alone makes the prospect of allele mining in other clades desirable. For example, *O. australiensis* is the sole member of the EE genome subclade and has a large genome (~900 Mbp) compared to domestic rice (~500 Mbp; Jacquemin *et al.*, 2013). Given its distribution in extreme environments, populations of this species have evolved tolerances to various abiotic stresses. For example, an accession of *O. australiensis* from the Keep River region has been reported to be thermotolerant (Scafaro *et al.*, 2016), however in comparison to other *O. australiensis* accessions from salt marsh areas, the Keep River rice had low tolerance to salt (Yichie *et al.*, 2018). The size of the *O. australiensis* genome and its occurrence in conditions that would be sub-optimal for domestic rice (Atwell *et al.*, 2014) make it desirable to sequence the genomes of the disparate accessions. This is also true for other *Oryza* species that may possess tolerance to other stress types. Identification of novel genetic variation within key gene families, particularly the V-PPase family, may ultimately serve to improve the ability of taxonomically diverse crop species to withstand a multitude of stresses (Atwell *et al.*, 2014).

1.6 Summary

While rice is generally considered to be one of the most anoxia tolerant crops, erratic rainfall driven by changes in climate will increasingly expose them to periods of excess and/or reduced water supply. As a result, cells enter a ‘low energy’ state and ultimately yield losses may be observed. Improved seedling vigour is one route toward fortifying rice crops, as young plants are more susceptible to environmental perturbations than mature plants. During an energy crisis, cells can shift their metabolisms away from ATP-dependent to PP_i-dependent processes, thus improving the energy status of the cell. Vacuolar pyrophosphatases (V-PPases) play an important role in this transition and their overexpression has been shown to modulate the growth, yield and tolerance of a number of species to abiotic stresses. Improved cell division and carbohydrate metabolism are likely drivers of improved growth, while increased vacuolar acidification may aid in maintaining membrane potentials in a low-energy cell. However, the role of V-PPases has not been explored in flooding-induced anoxia tolerance in rice. The expression of OVP3, one of six V-PPase isoforms in the rice genome, is increased in response to low oxygen conditions. Given the role that V-PPases play in normal cellular development and the response of OVP3 to anoxia (a stress that can be used as a proxy for general tolerance to low energy stress), we hypothesise that the growth, yield, and stress tolerance of *O. sativa* can be improved through the overexpression of OVP3.

1.7 Research Proposal

Materials and Research questions

This project will make use of *O. sativa* (Calrose, Japonica), *O. australiensis* (a wild relative of rice), and seven lines of domestic rice that have been genetically modified to over-express (OE) the OVP3 gene or which have had either all isoforms or OVP3

alone knocked-out (KO). These lines were generated by Dr. Qinxiang Liu in 2009 as part of her work for the degree of Doctor of Philosophy at Macquarie University in Sydney. Lines 1, 3, and 4 are KOs of OVP3, lines 2 and 5 are OVP family KO lines, and lines 6 and 7 are OVP3 OE lines.

The OE lines were constructed by cloning the OVP3 open reading frame (ORF) into a pMDC84 destination vector via a pENTR/D-TOPO entry vector (Invitrogen, Australia). The expression of the OVP3 ORF is under the control of a CaMV 35S promoter and Nos terminator. The constructs were transformed into *O. sativa* (Calrose) via *Agrobacterium tumefaciens* as per (Sallaud *et al.*, 2003).

OVP KO was achieved by gene silencing via double-stranded RNA interference (dsRNAi). The KO lines were generated by cloning a unique OVP3 sequence or a conserved sequence present in all six OVP isoforms into pTOOL1 vectors, via pCR8/GW/TOPO vectors for amplification. *O. sativa* was transformed with these pTOOL1 vectors via *A. tumefaciens*.

O. australiensis is an extremophile wild relative of rice native to northern Australia, particularly the Northern Territory. Its native range exposes *O. australiensis* to extremely high day time temperatures, fluctuating water availability, and inundation with saline water. It is the sole member of the Oryza EE genome subclade, which is approximately double the size of the O. sativa genome (900 Mbp vs. ~500 Mbp).

Overarching Question:

What role does vacuolar pyrophosphatase 3 (OVP3) play in the growth, yield and low-energy stress tolerance of rice?

Sub-questions:

1. Does OVP3 over-expression and knockout affect the vigour of mutant seed?
2. Does OVP3 over-expression and knockout affect the growth rates and yield of *O. sativa* lines gorwn under hypoxia? Does wild type *O. australiensis* tolerate hypoxic growth?
3. What is the mechanism of improved vigour/growth rate in *O. sativ*a mutant lines?
4. Do wild rice genomes possess novel OVP variants or functional genomic variation related to tolerance to abiotic stresses?

Aims and objectives

Aim 1: Characterise the germination traits of OVP3 over-expressing (OE) and knockout (KO) lines and wild type *O. australiensis* during optimal and low-energy conditions

Objective: measure changes in germination rates, radicle and coleoptile elongation rates, cell size and number, soluble (including sucrose, glucose, fructose, and UDP-glucose) and insoluble sugars in post-germinative tissues, the composition of the endosperm during germination (e.g. oil/fat and protein content in relation to gluconeogenesis), and PPi-hydrolysis activity in each rice line during hypoxic stress and with and without sucrose supplied via the growth media.

Aim 2: Characterise the effect of OVP3 OE or KO on the growth and yield of rice plants during optimal and low-energy conditions

Objective: measure changes in the growth rates of roots and shoots of mature plants (those that have achieved photosynthetic capacity) via sequential destructive harvests, cell size and number will be estimated via microscopy, and yield will be measured as

changes in 100-seed weight and number. Similar experiments can also be performed to test whether OVP3 OE improves tolerance to excess salt.

Aim 3: Understand the mechanisms of changed growth or yield in post-germinative and mature OVP3 OE and KO plants

Objective: qPCR and shot-gun or SWATH proteomics will be used to determine, semi-quantitatively, which transcripts and proteins are affected by OVP3 OE and KO in young root and shoot tissue harvested from plants grown for Aim 1 and Aim 2. Microelectrode flux estimation (MIFE) will be used to test differences in membrane potentials across the tonoplast of isolated vacuoles. Atomic absorption spectroscopy will be used to measure ion accumulation in isolated vacuoles and the leakage of solutes from cells. Malondialdehyde (MDA) content will be analysed to assess lipid breakdown, an indicator of membrane damage. PPase and ATPase activities on the isolated tonoplasts and on isolated PM fractions will be measured via a colorimetric method, with PPase measurements validated by a molybdenum blue reaction method. Immunohistochemistry will be used alongside confocal microscopy to identify which membranes are targeted by OVP3 when it is over-expressed.

Aim 4: Identify variations in the genomes of extremophile wild rice species that may contribute to improved multi-stress tolerance.

Objective: Identify genomic elements associated with contrasting tolerances to V-PPase-relevant stresses (e.g. drought and salt) in geographically-isolated accessions of wild rice species (*O. australiensis* and *O. meridionalis*) using third generation MinION sequencing.

Theoretical framework and methods

Aim 1:

There is mounting evidence that the improved growth of young plants overexpressing V-PPases is due to enhanced gluconeogenesis. Improved growth rates in young seedlings would aid in increasing tolerance to inundation of seedlings in the field. As such, germination rates, coleoptile development and reserve carbohydrate and oil content will be studied throughout the germination process under anoxic conditions. It is expected that OVP3 OE plants will have enhanced gluconeogenesis compared to WT and KO lines, leading to faster depletion of endosperm oil content, and correspondingly increased sucrose synthesis, greater flux of sucrose through glycolysis, and faster coleoptile growth rates. KO lines are expected to have reduced gluconeogenic capacity, and as such slower coleoptile elongation rates, though this phenotype should be rescued with sucrose supplementation. Changes in the composition of seed endosperm may affect seedling vigour as well as the quality of the grain, so it is desirable to characterise some of these changes. All work with mutants will be undertaken in a PC2 facility.

Plant Germination and Growth

Seeds (described above) will be surface-sterilised with a bleach:ethanol:triton solution and germinated either in *sealed gel plates*, or in *sealed falcon tubes filled with water*. The plates will be a 0.5% agarose, half-strength Murashige and Skoog Basal media, and supplemented with (0.1 and 1%) sucrose or without. Sucrose supplementation allows the seedling to by-pass gluconeogenic pathways, which will aid in understanding changes to seedling vigour across lines. The gel will be cast in the lid of the plate such that the base can be used to squish the seeds into the gel and create a moisture seal, limiting oxygen availability. In separate experiments, the seeds of each genotype will be placed in 50 mL falcon tubes filled with distilled water and sealed with

screw caps. Tubes and plates will be kept in clear tubs sealed by clamping a rubber-rimmed clear lid to the tub and will be kept in a growth chamber at 30 ^{o}C with no light exposure (no light ensures that all growth is the result of mobilisation of endosperm reserves). Following imbibition, germination rates and coleoptile growth rate will be measured, and seeds/seedlings will be collected every day for one week for coleoptile cell size, sugar, oil and protein content as per Ferjani et al. (2011) and Adachi et al. (2015). The experiment will be repeated with light exposure to see how endogenous sucrose affects germination traits, and with sucrose supplementation to see if exogenous sucrose affects germination traits. An inhibitor of ATPases will be used during germination and early growth to see if the OE of OVP3 is capable of meeting the task of acidifying the vacuole in the absence of V-ATPase activity.

Estimation of Cell Size, Number, and Volume

Light, fluorescent, and confocal microscopy will be used to estimate cell size and volume across a range of tissues in the OE and KO lines. Tissue visualisation on the fluorescent (Axio) microscope and confocal microscope will be achieved by staining both live and fixed tissue with different combinations of cell wall stains (direct red, calcofluor white, SR2200; Ursache *et al.*, 2018), and nuclei stains (syto-9; Thomas and Sekhar, 2014). Tissue fixed in a TEM fixative (formaldehyde, glutaraldehyde, sucrose, PBS solution) will be soaked in ClearSee (a tissue clearing solution that aids in the visualisation of plant tissues) for up to eight weeks (Ursache *et al.*, 2018). Imaris software will be used to estimate cell volume (Pasternak *et al.*, 2015).

Pectin methylesterase (PME)-mediated dispersal of root cells combined with cell-volume based flow cytometry is a novel combination of techniques that might also aid in estimating cell volume.

Soluble and Insoluble Sugars

Soluble sugars will be extracted from the homogenised tissue of etiolated seedlings (endosperm, coleoptile and root tissue separately) following the ethanol extraction methods of Rasse and Tocquin (2006) and Landhausser *et al*. (2018) with modifications. The supernatant of this extraction will be measured on the Dionex, which will give information about the abundance of the soluble sugars generated during gluconeogenesis. Reduced rates of gluconeogenesis have been linked with reduced sucrose synthesis in the presence of excess PP_i.

Total insoluble sugars (namely starch) not extracted in the soluble phase of the AIR method will be treated with α-amylase to hydrolyse the glycosidic bonds linking the glucose monomers (Rasse and Tocquin, 2006; Landhäusser *et al.*, 2018). The abundance of glucose will be estimated on the Dionex, as above, and a starch correction factor (0.9) will be used to estimate starch content (Basu *et al.*, 2011). Single grain starch analysis may also be employed (South and Morrison, 1990).

Oil Content

Gas chromatography mass spectrometry (GC-MS) can be used to measure fatty acid content via conversion of fatty acids to fatty acid methyl esters (FAMEs). If GC-MS is not available, TAG content can be estimated using the Highly Sensitive Triglyceride Fluorometric Assay Kit from Sigma ($686 for 100 rxns = $6.86 per rxn). Alternatively, the Megazyme glycerol kit might be co-opted to estimate TAG content (one mole of glycerol is produced per hydrolysed TAG; $207.41 for 70 rxns = $2.96/rxn). Finally, a microalgae Nile Red-TAG fluorescence method can be optimised for TAG extracted from rice seed (Ruiz-López *et al.*, 2003); changes in fluorescence can be measured

on a fluorescence spectrophotometer (Rumin *et al.*, 2015). I will attempt to generate a standard curve of TAG vs. fluorescence intensity using TAG standards.

Protein Content

Protein from leaf and grain will be extracted following the procedure of (Lin and Wang, 2014), while root proteins will be extracted following (Raorane *et al.*, 2016). Total protein of the tissues will be measured on a Qubit (Invitrogen, US).

Grain Quality (Waxiness and Chalkiness)

The ratio of amylose to amylopectin may change and/or the structure of the starch granules in the endosperm may be altered by OVP OE or KO. Low amylose:amylopectin ratios lead to waxy grains, which is undesirable in that the glycaemic index (GI) of the grain is increased. This is because amylopectin is hydrolysed by α-amylase more rapidly than amylose, and thus has a higher GI. The presence of non-starch polysaccharides (NSPs; e.g. xylan) decreases the rate of α-amylase activity however, and so a high amylopectin-high NSP grain may still have a reasonably low GI. Low GI grain is desirable in that the release of soluble sugars into the blood stream is slowed, which may have benefits for human health (Jovanovski *et al.*, 2015). The Megazyme kit will be used to measure changes in the ratio of amylose to amylopectin in the OE and KO lines to see if the quality of the rice grain is changed in response to OVP3 expression level.

Further, the *Chalk5* locus correlates to V-PPase in rice, and high V-PPase activity has been linked to increased chalkiness (Li *et al.*, 2014). This is an undesirable trait for rice grain. Disturbed pH homeostasis of the endomembrane trafficking system in developing grain affects the formation of protein bodies, which increases the

occurrence of small vesicle-like structures. This, in turn, causes an increase in the formation of air spaces in the grain, leading to the chalky texture. Changes in the size of starch granules in the endosperm will be observed by SEM (Myers *et al.*, 2011), confocal imaging of starch granule extracts (Velde *et al.*, 2002), and by fluorescence imaging of seed sections using fluorescein-5-isothiocyanate (FITC). FITC is a fluorescent stain that can be used to identify the structure of starch granules in seed sections (Gallant and Bouchet, 1986).

Aim 2:

Several studies report increased growth and yield in plants over-expressing V-PPases from diverse sources during growth in optimal and sub-optimal conditions. However, there are no studies published to date that explore the use of V-PPases in mediating tolerance to low-oxygen conditions. Additionally, the effect of OVP3 over-expression on plant growth has never been explored. This study will investigate changes in the growth and yield of *O. sativa* (cv. Calrose) lines in which OVP3 is over-expressed (OE) or knocked out (KO).

Plant Growth

Rice grain will be sterilised as in Aim 1 and germinated on moist filter paper in the absence of light for 48 hours. Seedlings will be grown for one week in a nutrient solution containing 0.5 g L-1 liquid fertiliser (AQUASOL, Yates, Australia) with 10 μM calcium chloride (after five days; Islam *et al.*, 1987) before being installed into a hydroponic growth system at Macquarie University's Plant Growth Facility PC2 glasshouse. The glasshouse will be set to a day/night cycle of 30 ^{o}C/22 ^{o}C, with a 12-hour photoperiod supplemented with purple LED lighting when ambient light levels drop below 700 μmol m^{-2} s^{-1}. Seven-day old seedlings will be transferred into punnets

within 93 L or 53 L black plastic tubs containing nutrient solution. There will be 20 punnets per tub, sitting in a 4x5 rack. The black tubs of the ***aerated*** treatment group will sit on top of 500 L reservoir tubs, which will contain a modified PusaRicH nutrient solution (Sharma *et al.*, 2018), water pumps, and air outlet tubing. The water pumps (which have a pumping capacity of ~2000 L/hr) will shift aerated nutrient solution up into the black tubs, which will subsequently overflow (through piping) back into the reservoir tub (each pump will be pumping to three of the black tubs). All tubs (53-93 L and the reservoir tubs) will be sealed with black plastic sheeting to minimise exposure of the nutrient solution to light, limiting algal growth. Seedlings will be held above their punnets in holes (spaced ~ 10 cm apart) in the sheeting with moist sponges. Each punnet will contain the root system of the respective plant, while allowing fluid movement through and beneath the punnet. The fluid level in the black tubs will allow for an air space between the water line and the top of the plastic sheeting ensuring the root systems never experience hypoxia. On the other hand, the black tubs of the ***hypoxic*** treatment group will sit on 500 L reservoir tubs for consistency, but there will be no media movement between the black tubs and the reservoir tubs. Instead, the hypoxic tubs will be allowed to stagnate, with no flow and nitrogen gas bubbling through the PusaRicH solution to reduce oxygen concentration. There will be no air gap between the black sheeting and the fluid line in these tubs, further reducing the availability of oxygen to the root system. These plants will be grown for six weeks under the conditions described above, with the nutrient solution changed once every two weeks as per Sharma *et al.* (2018)

Sequential Destructive Harvests and Yield

Changes in the growth rates between the OVP3 lines will be measured via sequential destructive harvests. Over the six-week growth period there will be four harvest dates

(one early, one late and two intermediate dates). Root and shoot fresh weight, plant height and tiller number will be recorded. Roots and shoots will be oven dried at 70 ^{o}C for one week and dry weights will be measured. This will allow for the construction of root and shoot growth curves and will allow for an estimate of relative water content. A separate cohort of plants will be grown in soil for grain yield estimates. Harvested seed will be weighed to measure 100-seed weight with and without hulls.

Cell Size, Number, and Volume

As in Aim 1.

Aim 3:

The phenotypes of V-PPase OE plants have been extensively characterised. However, the mechanisms behind improved growth and stress tolerance remain largely unknown. Therefore, this study aims to understand the mechanisms of improved growth, in part by identifying the pathways affected by changes in the expression of OVP3 in young and mature root and shoot tissues as well as the growing coleoptile.

Pleiotropic Effects of OVP3 OE or KO

qPCR and shot-gun or SWATH proteomics will be performed on young root and leaf tissue excised from the plants grown for the sequential harvest described in Aim 2. qPCR will be used to measure changes in the abundance of all OVP transcripts, while shot-gun proteomics will be used to measure which cellular processes are affected by OVP3 OE or KO during stress and non-stress growth conditions. RNA will be extracted from these tissues using the Sigma Spectrum kit. OVP cDNA synthesised from the extracted RNA will be quantified to assess OVP expression as per Liu *et al.* (2010*b*) using a CFX384 Real-Time PCR Detection System (Bio Rad, Australia). Total protein

from the samples will be extracted as per the protocol outlined by the Australian Proteomics Analysis Facility (APAF) located at Macquarie University in Sydney (Wu *et al.*, 2019) or through the use of TRIzol, as per Masood *et al.* (2016), to yield RNA and protein fractions. Protein extracts will be subject to a SWATH or shotgun proteomics pipeline established by staff at APAF.

Photosynthetic Rates

There is some evidence that the increased expression of V-PPases in transgenic organisms enhances photosynthetic capacity (Schilling *et al.*, 2017). Changes in carbon fixation may lead to changes in growth or in carbon metabolism. As such, a Licor 6800 will be used to measure gas exchange parameters on OE and KO lines, as per Caine *et al*. (2019).

Changes in Membrane Dynamics and Integrity

Low-energy conditions, such as those imposed on a cell during hypoxia, causes the plasma membrane to depolarise leading to cell death by the leakage of solutes (Zeng *et al.*, 2014; Atwell *et al.*, 2015). There is potential for an acidified vacuole to act as a battery for the cell during low-energy conditions. As such, changes to the membrane potentials of root cells and their tonoplasts as well as potassium flux will be measured using MIFE (Chakraborty *et al.*, 2016). Membrane leakage will be measured on solutes collected from root tissue incubated in low-oxygen conditions using an osmometer. Damage to membrane integrity, which may be linked to increased solute leakage, can be estimated by measuring changes in the abundance of malondialdehyde (MDA), a breakdown product of lipids. MDA content can be estimated using a spectrophotometer following the reaction of MDA with thiobarbituric acid (TBA) at 532 nm (Zhang and Huang, 2013).

One role of the vacuole during stress responses is the storage of toxic ions or solutes, which are deposited in the vacuole through the action of proton/ion antiporters on the tonoplast. Therefore, ion accumulation in the tonoplast will be estimated using atomic absorption spectroscopy on isolated vacuoles or vacuolar vesicles (AAS; Baetz *et al.*, 2016).

Vacuolar pH and Cytosolic PP_i Content

V-PPase-mediated vacuolar acidification can be measured on crude or vacuolar membrane vesicles by the methods of Maeshima and Yoshida (1989). This is a fluorescence quenching method using acridine orange as a vacuolar pH indicator and so can be performed on a fluorescence spectrophotometer. The more acidic the environment that acridine orange is present in the more red fluorescence is emitted (Schoor *et al.*, 2015).

PP_i concentration in the same tissues as above can be measured by the method of Smyth and Black (1984). The PP_i reagent used therein is rather expensive, and so their extraction method might be combined with a fluorescence-based kit from Sigma or Abcam (this kit or this kit). Further, the V-PPase activity (i.e. PP_i hydrolysis) of the isolated vacuoles can be measured by the methods of Asaoka *et al.* (2016) and Maeshima and Yoshida (1989).

Membrane Localisation of OVP

There is evidence that V-PPase is capable of targeting the Golgi body and the plasma membrane along with the tonoplast, and that its localisation may change in response to stress conditions. In doing so, the function of V-PPases may shift from PPi-

hydrolysis to PPi-synthesis (Gaxiola *et al.*, 2015; Scholz-Starke *et al.*, 2019), which may have implications for normal cellular development and response to stress. To determine the membranes targeted by the OVP gene family, OVP-specific rabbit antibodies (Agrisera, Sweden; or Cosmo Bio, Japan; or PhytoAB, USA) will be used alongside Alexa Fluor 488 anti-rabbit secondary antibodies (Invitrogen, USA) and visualised on Axio and confocal microscopes (Pasternak *et al.*, 2015).

Aim 4:

O. australiensis is a wild rice relative native to the northern regions of Australia. Several geographically-isolated accessions inhabit environments that are considered extreme. An accession from Keep River has demonstrated considerable photosynthetic thermotolerance but low tolerance to saline conditions, while another accession demonstrates considerable tolerance to high salt loads. Because of this, it is desirable to identify variations in the genomes of these accessions as they may contribute to improving multi-stress tolerance in domestic species or may themselves prove to be useful food crops in the future.

Plant Growth and Sequencing

Seeds from four populations of *O. australiensis* with varying levels of salt tolerance, two accessions of *O. meridionalis* (another wild rice that co-occurs with *O. australiensis* but does not demonstrate the level of tolerance to abiotic stress seen in *O. australiensis*), and ISC113 (a Javanica rice demonstrating drought tolerance) will be surface sterilised as above and germinated at 30 ^{o}C on moist filter paper-lined petri dishes in the absence of light. Young seedlings will be sown onto potting mix and grown for ~5 weeks in glasshouse conditions as described above.

Genomic DNA will be extracted from nuclei isolated from the shoots (~500 mg of material per line is required) of the above lines following the methods of Bolger et al. (2014) and Mayjonade et al. (2016) with modifications (Jones and Borevitz, 2018). The genomes of these accessions/species will be sequenced with long-read 3rd generation Oxford Nanopore MinION flow cell technology (Michael *et al.*, 2018). Genomes will be assembled via a collaboration with the Borevitz Lab at ANU and the Bioinformatics Hub at UoA. Isoforms of OVP will be mined from the genomes and checked for SNPs and structural variation. Bioinformatics tools utilised will include a host of Linux tools and packages, Albacore for basecalling (Pomerantz *et al.*, 2018), PoreChop for adaptor trimming (Wick, 2017), Nanolyse for removing control DNA strands (De Coster *et al.*, 2018), FastQC for quality control (Andrews, 2010), MultiQC for visualising read quality (Ewels *et al.*, 2016), NanoFilt for read trimming (based on MultiQC) and filtering out small read lengths (De Coster *et al.*, 2018), NanoPlot for summary plots and statistics (De Coster *et al.*, 2018), Canu for genome assembly (Koren *et al.*, 2017), Bioawk for examining the assembly (https://github.com/lh3/bioawk), Blobtools for the removal of contaminating (i.e. from microbes etc.; Laetsch and Blaxter, 2017), Purge Haplotigs for identifying primary contigs for haploid assembly (Roach *et al.*, 2018), minimap2 to align assembly reads to the assembly (Li, 2018), Racon for polishing the assembly either with MinION sequences or with Illumina short read sequences (https://github.com/isovic/racon), Chromosomer for scaffolding assembled genomes onto previously sequenced genomes for contig joining (Tamazian *et al.*, 2016), Red to mask genomic repeat regions (as required by Chromosomer; Girgis, 2015), BUSCO for genome quality assessment and ortholog completeness (Simão *et al.*, 2015), Sniffles for structural variation calling (Sedlazeck *et al.*, 2018), and ggplot2 and ggbio in the R environment to visualise outputs (Yin *et al.*, 2012; Wickham, 2016).

Significance and contribution to the discipline

Changes in climate and poor management of soils increasingly place plants into a state of low-energy stress. This is affected by increasing global temperatures, changes in the distribution of rain water, and increasing soil salinity. For agricultural species low-energy stress often results in yield losses. For example, changes in rainfall patterns negatively affect the production of rice by the dry-seeding method via both drought and hypoxic stress. Altering the expression of vacuolar pyrophosphatase (VP) genes offers one route to increased growth, yield and multiple stress tolerance in a variety of species. This research aims to establish the role of a VP (OVP3) in the growth, yield and stress tolerance of *Oryza sativa* (domestic rice) and several rice wild relatives. Doing so may not only provide targets for breeding/modification in rice crops but will also further understanding of energy dynamics in plant cells under stress.

Budget

4.1 Will you submit an application to a funding agency for support for your project?

YES ☒ NO ☐

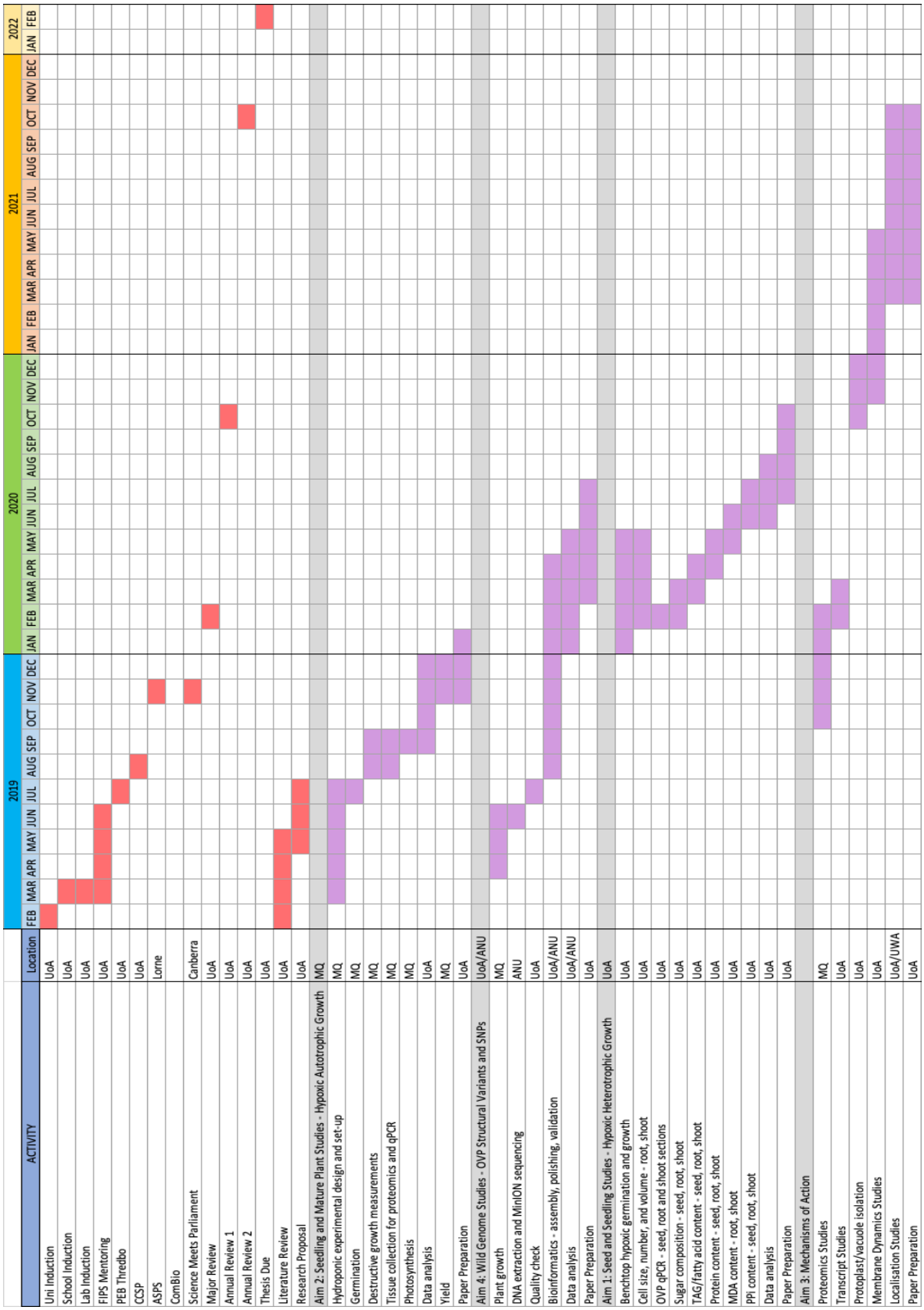
An application to the Deputy Vice Chancellor (Research) at Macquarie University will be made to fund the genomics work at ANU. We have received this funding as of 19/07/19.

4.2 Is this project contingent on funding support from elsewhere?

YES ☐ NO ☒

Research plan timeline

Gaant Chart



Potential risks and limitations

OVP lines

This project will make use of OVP mutants generated by a previous PhD student of Rachel Burton and Brian Atwell. The collection (seven lines in total) is comprised of three OVP3 KO lines, one OVP1 KO lines, one OVP4 KO lines, and two OVP3 OE lines. While the focus of my project will be on OVP3, the use of other OVP KOs may shed light on OVP3 function. However, our collection is limited by replication, especially for OVP1 and OVP4. We have only one KO line for each of these isoforms. Further, we have two OE lines for OVP3. While this is acceptable and does not impede the progression of my project, it would be better to have a third OVP3 OE line to confirm the OE phenotype. Further, we have no null-segregants in our collection, though conversations are underway with researchers at CSIRO in Canberra as they have a collection of transformants that might be utilised. Alternatively, the plant transformation facilities of Waite campus can be used for the generation of a null-segregant.

Growth and maintenance of plants and growth experiments

To grow healthily, rice requires a growth temperature of 30-35 ^{o}C and a constant relative humidity of ~70%. The high use of glasshouses on Waite campus mean that the facilities required to grow my plants (which require PC2 level containment) are often unavailable. To get around this, the PC2 glasshouse facilities at Macquarie University will be used for the growth experiments of my project. This means that all physiological measurements will have to be done offsite and will require travel between universities until I can secure a glasshouse with the right conditions. For experiments at UoA, I can scale down the size of experiments such that they can be performed in the PC2 room of the level 4 WIC laboratory. While this is fine for germination experiments, the lack of adequate light conditions in the PC2 room may affect any

experiments on seedlings. Particularly difficult will be the growth of wild relatives of rice, which have thermal optimum of ~40 ^{o}C for growth and require high light conditions.

The planned hydroponic hypoxic experiment has the potential to kill the plants before the conclusion of the experiment. This in itself would be a result, though would mean the collection of less data and would mean a loss of insight into the function of OVP3 in these plants. Because this experiment is being done off-site, it will be difficult for me to monitor directly. Close monitoring of the plants during the experiment by a research assistant and supervisor in Sydney will alert us to any potential issues arising from the design of the experiment. The hydroponic nature of this experiment also reduces our ability to extrapolate the results to field conditions, though it allows us to answer some fundamental questions about our study system.

Cell size, number and volume

In order to determine cell size, number and volume accurately I will require access to microscopy facilities. The Burton Lab has access to an Axio fluorescence microscope, however heavy use of this microscope renders it inaccessible for considerable chunks of time. To combat issues with access to equipment, I have received training for the use of a confocal microscope at CSIRO at North Terrace.

Measuring TAG content

The measurement of TAG content is typically done by the fatty acid methyl ester (FAME) technique using GC-MS. I don't have access to these facilities here and these techniques require quite a bit of material, which is an issue because we don't have excess seed at UoA. However, laboratories interested in TAG content in algae for the

production of biofuels commonly use a Nile Red staining technique for the quantification of TAG in living cells. I would like to try adapting this method for the rice tissue through the use of a spectrophotometer. This has not been tried before and might not work. However, there are other methods for the quantification of TAGs that can be used in the event of failure. For example, John Carragher and Bob Gibson at North Terrace can be contacted for oil quantification in a sample, though this requires a substantial amount of seed material.

Localisation

We have access to a general VP antibody, though no antibody is available for specific rice isoforms. As such, the immunohistochemistry line of work would involve the visualisation of all OVP isoforms, not just OVP3. While this is not ideal, we will still be able to detect differences in localisation of OVPs in response to stress and in KO and OE lines. The use of the monomeric-GFP-tagged OVP3 isoform is one route in which we might overcome this 'problem', though that itself introduces the possibility of OVP3 mistargeting due to the conditions involved in protoplast generation.

MIFE

The ideal system for studying the effect of OVP3 OE and KO on the membrane dynamics of the vacuole would be to isolate intact vacuoles from rice leaf and root cells. This is notoriously difficult. It may be that I am only successful in isolating vacuolar vesicles, which are not true vacuoles but may be similar enough for the purposes of this study.

Genome sequencing

The inability to extract quality high molecular weight DNA from plants is a hurdle faced by many researchers aiming to sequence the genomes of plants, especially those with large or complex genomes. This is partly because plant genomes contain regions dominated by repeat sequences, rendering short-read sequencing less-desirable than long read sequencing. To combat issues with secondary metabolites while still extracting good quality DNA for long-read sequencing, I will travel to ANU to work with Justin Borevitz and Ashley Jones, who have developed a technique for nuclei isolation and subsequent DNA extraction for eucalyptus species. These methods should prove useful for rice, which contains silica in its leaves, as silica binds DNA.

Grain quality

Assessment of grain quality requires a large amount of material. Given that I currently have relatively sparse seed available at Waite campus this part of the project will be postponed until we are capable of bulking the seed stock either at MQ in Sydney or in Adelaide.

1.8 References

Adachi Y, Sugiyama M, Sakagami J, Fukuda A, Ohe M, Watanabe H. 2015. Seed Germination and Coleoptile Growth of New Rice Lines Adapted to Hypoxic Conditions. Plant Production Science **18**, 471–475.

Ahire ML, Anil Kumar S, Punita DL, Mundada PS, Kavi Kishor PB, Nikam TD. 2018. The vacuolar proton pyrophosphatase gene (SbVPPase) from the Sorghum bicolor confers salt tolerance in transgenic Brahmi [Bacopa monnieri (L.) Pennell]. Physiology and Molecular Biology of Plants **24**, 809–819.

Andrews S. 2010. FastQC: A quality control tool for high throughput sequence data.

Arif A, Zafar Y, Arif M, Blumwald E. 2013. Improved growth, drought tolerance, and ultrastructural evidence of increased turgidity in tobacco plants overexpressing arabidopsis vacuolar pyrophosphatase (AVP1). Molecular Biotechnology **54**, 379–392.

Asaoka M, Segami S, Ferjani A, Maeshima M, Moore I, Pietro G, Sansebastiano D. 2016. Contribution of PPi-Hydrolyzing Function of Vacuolar H + -Pyrophosphatase

in Vegetative Growth of Arabidopsis : Evidenced by Expression of Uncoupling Mutated Enzymes. **7**, 1–12.

Asaoka M, Segami S, Maeshima M. 2014. Identification of the critical residues for the function of vacuolar H+-pyrophosphatase by mutational analysis based on the 3D structure. Journal of Biochemistry **156**, 333–344.

Atwell BJ, Greenway H, Colmer TD. 2015. Tansley review Efficient use of energy in anoxia-tolerant plants with focus on germinating rice seedlings.

Atwell BJ, Wang H, Scafaro AP. 2014. Could abiotic stress tolerance in wild relatives of rice be used to improve Oryza sativa? Plant Science **215–216**, 48–58.

Atwell BJ, Waters I, Greenway H. 1982. The effect of oxygen and turbulence on elongation of coleoptiles of submergence-tolerant and -intolerant rice cultivars. Journal of Experimental Botany **33**, 1030–1044.

Baetz U, Eisenach C, Tohge T, Martinoia E, De Angeli A. 2016. Vacuolar Chloride Fluxes Impact Ion content and Distribution during Early Salinity Stress. Plant Physiology **172**, pp.00183.2016.

Bailey-Serres J, Fukao T, Gibbs DJ, Holdsworth MJ, Lee SC, Licausi F, Perata P, Voesenek LACJ, van Dongen JT. 2012. Making sense of low oxygen sensing. Trends in Plant Science **17**, 129–138.

Bailey-Serres J, Fukao T, Ronald P, Ismail A, Heuer S, Mackill D. 2010. Submergence tolerant rice: SUB1's journey from landrace to modern cultivar. Rice **3**, 138–147.

Bao AK, Wang SM, Wu GQ, Xi JJ, Zhang JL, Wang CM. 2009. Overexpression of the Arabidopsis H+-PPase enhanced resistance to salt and drought stress in transgenic alfalfa (Medicago sativa L.). Plant Science **176**, 232–240.

Basu S, Roychoudhury A, Sanyal S, Sengupta DN. 2011. Carbohydrate content and antioxidative potential of the seed of three edible indica rice (Oryza sativa L.) cultivars. Indian Journal of Biochemistry and Biophysics **49**, 115–123.

Bhaskaran S, Savithramma DL. 2011. Co-expression of Pennisetum glaucum vacuolar Na +/H + antiporter and Arabidopsis H +-pyrophosphatase enhances salt tolerance in transgenic tomato. Journal of Experimental Botany **62**, 5561–5570.

Bibi I, Ali Q, Niazi NK, Kanwal N, Ahsan M, Ali F. 2015. Multivariate Analysis of Grain Yield and Its Attributing Traits in Different Maize Hybrids Grown under Heat and Drought Stress. Scientifica **2015**, 1–6.

Bolger A, Scossa F, Bolger ME, *et al.* 2014. The genome of the stress-tolerant wild tomato species Solanum pennellii. Nature Publishing Group **46**, 1034–1038.

Brini F, Gaxiola RA, Berkowitz GA, Masmoudi K. 2005. Cloning and characterization of a wheat vacuolar cation/proton antiporter and pyrophosphatase proton pump. Plant Physiology and Biochemistry **43**, 347–354.

Brini F, Hanin M, Mezghani I, Berkowitz GA, Masmoudi K. 2007. Overexpression of wheat Na+/H+ antiporter TNHX1 and H+-pyrophosphatase TVP1 improve salt-

and drought-stress tolerance in Arabidopsis thaliana plants. Journal of Experimental Botany **58**, 301–308.

Caine RS, Yin X, Sloan J, *et al.* 2019. Rice with reduced stomatal density conserves water and has improved drought tolerance under future climate conditions. New Phytologist **221**, 371–384.

Carystinos G, MacDonald H, Monroy A, Dhindsa R, Poole R. 1995. Vacuolar H+-translocating pyrophosphatase is induced by anoxia or chilling in seedlings of rice. Plant Physiology **108**, 641.

Chakraborty K, Bose J, Shabala L, Shabala S. 2016. Difference in root K+ retention ability and reduced sensitivity of K+-permeable channels to reactive oxygen species confer differential salt tolerance in three Brassica species. Journal of Experimental Botany **67**, 4611–4625.

Cheng C, Zhang Y, Chen X, Song J, Guo Z, Li K, Zhang K. 2018. Co-expression of AtNHX1 and TsVP improves the salt tolerance of transgenic cotton and increases seed cotton yield in a saline field. Molecular Breeding **38**.

Choura M, Rebaï A. 2005. Identification and characterization of new members of vacuolar H +-pyrophosphatase family from Oryza sativa genome. Russian Journal of Plant Physiology **52**, 821–825.

Colombo R, Cerana R. 1993. Enhanced Activity of Tonoplast Pyrophosphatase in NaCl-Grown Cells of Daucus carota. Journal of Plant Physiology **142**, 226–229.

De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: Visualizing and processing long-read sequencing data. Bioinformatics **34**, 2666–2669.

Davies JM, Poolet RJ, Reaf PA, Sanders D. 1992. Potassium transport into plant vacuoles energized directly by a proton-pumping inorganic pyrophosphatase. **89**, 11701–11705.

Dempewolf H, Guarino L. 2015. Reaching back through the domestication bottleneck: Tapping wild plant biodiversity for crop improvement. Acta Horticulturae **1101**, 165–168.

Dennis ES, Dolferus R, Ellis M, Rahman M, Wu Y, Hoeren FU, Grover A, Ismond KP, Good AG, Peacock WJ. 2000. Molecular strategies for improving waterlogging tolerance in plants. Journal of Experimental Botany **51**, 89–97.

Docampo R, Huang G. 2016. Acidocalcisomes of eukaryotes. Current Opinion in Cell Biology **41**, 66–72.

Drozdowicz YM, Kissinger JC, Rea PA. 2000. AVP2, a Sequence-Divergent, K + -Insensitive H + -Translocating Inorganic Pyrophosphatase from Arabidopsis. Plant Physiology **123**, 353–362.

Drozdowicz YM, Miller JH, Rea PA, Fitz-Gibbon S, Patel V, Lu Y-P. 1999. A thermostable vacuolar-type membrane pyrophosphatase from the archaeon Pyrobaculum aerophilum : implications for the origins of pyrophosphate-energized pumps. FEBS Letters **460**, 505–512.

Drozdowicz YM, Rea PA. 2001. Vacuolar H+ pyrophosphatases: From the evolutionary backwaters into the mainstream. Trends in Plant Science **6**, 206–211.

Duan XG, Yang AF, Gao F, Zhang SL, Zhang JR. 2007. Heterologous expression of vacuolar H+-PPase enhances the electrochemical gradient across the vacuolar membrane and improves tobacco cell salt tolerance. Protoplasma **232**, 87–95.

Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics **32**, 3047–3048.

Farooq M, Siddique KHM, Rehman H, Aziz T, Lee DJ, Wahid A. 2011. Rice direct seeding: Experiences, challenges and opportunities. Soil and Tillage Research **111**, 87–98.

Farquharson KL. 2018. Life of PPi: Soluble PPases and H + -PPase Act Cooperatively to Keep Pyrophosphate Levels in Check . The Plant Cell **30**, 951–951.

Ferjani A, Kawade K, Asaoka M, Oikawa A, Okada T, Mochizuki A, Maeshima M, Hirai MY, Saito K, Tsukaya H. 2018. Pyrophosphate inhibits gluconeogenesis by restricting UDP-glucose formation in vivo. Scientific Reports **8**, 1–11.

Ferjani A, Maeshima M. 2016. Editorial: Multiple Facets of H+-Pyrophosphatase and Related Enzymes. Frontiers in Plant Science **7**, 1–3.

Ferjani A, Muto Y, Horiguchi G, Maeshima M, Tsukaya H, Segami S. 2011. Keep an Eye on PPi: The Vacuolar-Type H + -Pyrophosphatase Regulates Postgerminative Development in Arabidopsis. The Plant Cell **23**, 2895–2908.

Frich P, Alexander L V., Della-Marta P, Gleason B, Haylock M, Tank Klein AMG, Peterson T. 2002. Observed coherent changes in climatic extremes during the second half of the twentieth century. Climate Research **19**, 193–212.

Fukuda A, Chiba K, Maeda M, Nakamura A, Maeshima M. 2004. Effect of salt and osmotic stresses on the expression of genes for the vacuolar H + -pyrophosphatase , H + -ATPase subunit A , and Na + / H + antiporter from barley *. **55**, 585–594.

Fukuda M, Segami S, Tomoyama T, Asaoka M, Nakanishi Y, Gunji S, Ferjani A, Maeshima M. 2016. Lack of H+-pyrophosphatase Prompts Developmental Damage in Arabidopsis Leaves on Ammonia-Free Culture Medium. Frontiers in Plant Science **7**, 1–12.

Gallant D, Bouchet B. 1986. Ultrastructure of Maize Starch Granules. A Review. Food Structure **5**, 16.

Gao F, Gao Q, Duan XG, Yue GD, Yang AF, Zhang JR. 2006. Cloning of an H+-PPase gene from Thellungiella halophila and its heterologous expression to improve tobacco salt tolerance. Journal of Experimental Botany **57**, 3259–3270.

Gao Q, Lv S, Zhang K, Lian L, Song Y, Zhang J. 2008. Overexpression of an H+-PPase gene from Thellungiella halophila in cotton enhances salt tolerance and improves growth and photosynthetic performance. Plant and Cell Physiology **49**, 1150–1164.

Gaxiola RA, Li J, Undurraga S, Dang LM, Allen GJ, Alper SL, Fink GR. 2001. Drought- and salt-tolerant plants result from overexpression of the AVP1 H+-pump. Proceedings of the National Academy of Sciences **98**, 11444–11449.

Gaxiola RA, Rao R, Sherman A, Grisafi P, Fink GR. 1999. The Arabidopsis thaliana proton transporters, AtNhx1 and Avp1, can function in cation detoxification in yeast. Proceedings of the National Academy of Sciences of the United States of America **96**, 1480–1485.

Gaxiola RA, Zhang S, Pizzio G, Regmi K, Paez-Valencia J. 2015. Plant H+-PPases: Reversible Enzymes with Contrasting Functions Dependent on Membrane Environment. Molecular Plant **9**, 317–319.

Girgis HZ. 2015. Red: An intelligent, rapid, accurate tool for detecting repeats de-novo on the genomic scale. BMC Bioinformatics **16**, 1–19.

Graus D, Konrad KR, Bemm F, *et al.* 2018. High V-PPase activity is beneficial under high salt loads, but detrimental without salinity. New Phytologist **219**, 1421–1432.

Greenway H, Armstrong W. 2018. Energy-crises in well-aerated and anoxic tissue: Does tolerance require the same specific proteins and energy-efficient transport? Functional Plant Biology **45**, 877–894.

Greenway H, Kulichikhin KY, Cawthray GR, Colmer TD. 2012. PH regulation in anoxic rice coleoptiles at pH 3.5: Biochemical pHstats and net H+ influx in the absence and presence of NO3-. Journal of Experimental Botany **63**, 1969–1983.

Guo S, Yin H, Zhang X, Zhao F, Li P, Chen S. 2006. Molecular cloning and characterization of a vacuolar H + -pyrophos- phatase gene , SsVP , from the halophyte Suaeda salsa and its overexpression increases salt and drought tolerance of Arabidopsis. , 41–50.

Gutiérrez-Luna FM, Hernández-Domínguez EE, Valencia-Turcotte LG, Rodríguez-Sotres R. 2018. Review: “Pyrophosphate and pyrophosphatases in plants, their involvement in stress responses and their possible relationship to secondary metabolism”. Plant Science **267**, 11–19.

Heinonen JK. 2001. *Biological Role of Inorganic Pyrophosphate*. Kluwer Academic Publishers.

Hisanaga T, Kawade K, Tsukaya H. 2015. Compensation : a key to clarifying the organ-level regulation of lateral organ size in plants. **66**, 1055–1063.

Huang S, Colmer TD, Millar AH. 2008. Does anoxia tolerance involve altering the energy currency towards PPi ? , 221–227.

Hyten DL, Song Q, Zhu Y, Choi I-Y, Nelson RL, Costa JM, Specht JE, Shoemaker RC, Cregan PB. 2006. Impacts of genetic bottlenecks on soybean genome diversity. Proceedings of the National Academy of Sciences **103**, 16666–16671.

Islam AKMS, Asher CJ, Edwards DG. 1987. Response of plants to calcium concentration in flowing solution culture with chloride or sulphate as the counter-ion. **395**, 377–395.

Jacquemin J, Bhatia D, Singh K, Wing RA. 2013. The International Oryza Map Alignment Project: Development of a genus-wide comparative genomics platform to help solve the 9 billion-people question. Current Opinion in Plant Biology **16**, 147–156.

Jones A, Borevitz J. 2018. *Nuclear DNA purification from recalcitrant plant species for long-read sequencing*. Canberra.

Jovanovski E, Zurbau A, Vuksan V. 2015. Carbohydrates and Endothelial Function: Is a Low-Carbohydrate Diet or a Low-Glycemic Index Diet Favourable for Vascular Health? Clinical Nutrition Research **4**, 69.

Kabala K, Janicka-Russak M, Reda M, Migocka M. 2014. Transcriptional regulation of the V-ATPase subunit c and V-PPase isoforms in Cucumis sativus under heavy metal stress. Physiologia Plantarum **150**, 32–45.

Kasai M, Nakamura T, Kudo N, Sato H, Maeshima M, Sawada S. 1998. The activity of the root vacuolar H(+)-pyrophosphatase in rye plants grown under conditions deficient in mineral nutrients. Plant & cell physiology **39**, 890–894.

Katano M, Takahashi K, Hirano T, Kazama Y, Abe T. 2016. Suppressor Screen and Phenotype Analyses Revealed an Emerging Role of the Monofunctional Peroxisomal Enoyl-CoA Hydratase 2 in Compensated Cell Enlargement. **7**, 1–11.

Kato-Noguchi H. 2008. Low temperature acclimation mediated by ethanol production is essential for chilling tolerance in rice roots. Plant Signaling and Behavior **3**, 202–203.

Kaur R, Chakraborty A, Bhunia RK, Sen SK, Ghosh AK. 2018. Tolerance to soil water stress by Oryza sativa cv. IR20 was improved by expression of Wsi18 gene locus from Oryza nivara. Biologia Plantarum **62**, 129–139.

Khadilkar AS, Yadav UP, Salazar C, Shulaev V, Paez-Valencia J, Pizzio GA, Gaxiola RA, Ayre BG. 2016. Constitutive and Companion Cell-Specific Overexpression of AVP1 , Encoding a Proton-Pumping Pyrophosphatase, Enhances Biomass Accumulation, Phloem Loading, and Long-Distance Transport. Plant Physiology **170**, 401–414.

Kim Y, Kim EJ, Rea PA. 1994. Isolation and Characterization of cDNAs Encoding the Vacuolar H + -Pyrophosphatase of Beta vulgaris ’. , 375–382.

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome research **27**, 722–736.

Kosaki T, Wasano K, Juo ASR. 1989. Multivariate statistical analysis of yield-determining factors. Soil Science and Plant Nutrition **35**, 597–607.

Kriegel A, Andrés Z, Medzihradszky A, *et al.* 2015*a*. Job Sharing in the Endomembrane System: Vacuolar Acidification Requires the Combined Activity of V-ATPase and V-PPase. The Plant Cell **27**, 3383–3396.

Kriegel A, Andrés Z, Medzihradszky A, *et al.* 2015*b*. Job Sharing in the Endomembrane System: Vacuolar Acidification Requires the Combined Activity of V-ATPase and V-PPase. The Plant Cell **27**, 3383–3396.

Kuroha T, Nagai K, Gamuyao R, *et al.* 2018. Ethylene-gibberellin signaling underlies adaptation of rice to periodic flooding. **186**, 181–186.

Laetsch DR, Blaxter ML. 2017. BlobTools: Interrogation of genome assemblies. F1000Research **6**, 1287.

Landhäusser SM, Chow PS, Turin Dickman L, *et al.* 2018. Standardized protocols and procedures can precisely and accurately quantify non-structural carbohydrates. Tree Physiology **38**, 1764–1778.

Lerchl J, Konig S, Zrenner R, Sonnewald U. 1995. Molecular cloning, characterization and expression analysis of isoforms encoding tonoplast-bound proton-translocating inorganic pyrophosphatase in tobacco. Plant molecular biology **29**, 833–840.

Leustek, Saito. 1999. Sulfate transport and assimilation in plants. Plant physiology **120**, 637–644.

Li H. 2018. Minimap2: Pairwise alignment for nucleotide sequences. Bioinformatics **34**, 3094–3100.

Li Z, Baldwin CM, Hu Q, Liu H, Luo H. 2010. Heterologous expression of Arabidopsis H+-pyrophosphatase enhances salt tolerance in transgenic creeping bentgrass (Agrostis stolonifera L.). Plant, Cell and Environment **33**, 272–289.

Li Y, Fan C, Xing Y, *et al.* 2014. Chalk5 encodes a vacuolar H+-translocating pyrophosphatase influencing grain chalkiness in rice. Nature Genetics **46**, 398.

Li J, Yang H, Peer WA, Richter G, Titapiwantakun B, Undurraga S, Murphy AS, Gilroy S, Gaxiola R. 2005. Arabidopsis H þ -PPase AVP1 Regulates Auxin-Mediated Organ Development. , 121–126.

Lin S-M, Huang Y-T, Tsai J-Y, Pan R-L, Chiu C-L, Sun Y-J, Hsiao C-D, Liu M-H, Tung J-Y, Liu T-H. 2012. Crystal structure of a membrane-embedded H+-translocating pyrophosphatase. Nature **484**, 399–403.

Lin D-G, Wang C-S. 2014. Extraction of Total Proteins from Rice Plant. Bio-protocol **4**, e1277.

Liu L, Wang Y, Wang N, Dong YY, Fan XD, Liu XM, Yang J, Li HY. 2011. Cloning of a Vacuolar H +-pyrophosphatase Gene from the Halophyte Suaeda corniculata whose Heterologous Overexpression Improves Salt, Saline-alkali and Drought Tolerance in Arabidopsis. Journal of Integrative Plant Biology **53**, 731–742.

Liu Q, Zhang Q, Burton RA, Shirley NJ, Atwell BJ. 2010*a*. Expression of vacuolar H+-pyrophosphatase (OVP3) is under control of an anoxia-inducible promoter in rice. Plant Molecular Biology **72**, 47–60.

Liu Q, Zhang Q, Burton RA, Shirley NJ, Atwell BJ. 2010*b*. Expression of vacuolar H+-pyrophosphatase (OVP3) is under control of an anoxia-inducible promoter in rice. Plant Molecular Biology **72**, 47–60.

Liu S, Zheng L, Xue Y, Zhang Q, Wang L, Shou H. 2010*c*. Overexpression of OsVP1 and OsNHX1 Increases Tolerance to Drought and Salinity in Rice. Journal of Plant Biology **53**, 444–452.

Lyons JM, Raison JK. 1970. Oxidative Activity of Mitochondria Isolated from Plant Tissues Sensitive and Resistant to Chilling Injury. Plant Physiology **45**, 386–389.

Maeshima M. 2000*a*. Vacuolar H+-pyrophosphatase. Biochimica et Biophysica Acta - Biomembranes.

Maeshima M. 2000*b*. Vacuolar H+-pyrophosphatase. Biochimica et Biophysica Acta - Biomembranes.

Maeshima M. 2001. Tonoplast Transporters : Organization and Function. Annual Review of Plant Physiology and Plant Molecular Biology **52**, 469–97.

Maeshima M, Yoshida S. 1989. Purification and properties of vacuolar membrane proton-translocating inorganic pyrophosphatase from mung bean. Journal of Biological Chemistry **264**, 20068–20073.

Mammadov J, Buyyarapu R, Guttikonda SK, Parliament K. 2018. Wild Relatives of Maize , Rice , Cotton , and Soybean : Treasure Troves for Tolerance to Biotic and Abiotic Stresses. **9**.

Masood M, Raftos DA, Nair S V. 2016. Two oyster species that show differential susceptibility to virus infection also show differential proteomic responses to generic dsRNA. Journal of Proteome Research **15**, 1735–1746.

Mayjonade B, Gouzy J, Donnadieu C, Pouilly N, Marande W, Callot C, Langlade N, Muños S. 2016. Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. BioTechniques **61**, 203–205.

McFarlane D, Wheaton B. 1990. The extent and cost of waterlogging. Journal of the Department of Agriculture, Western Australia **31**.

Menadue DJ, Riboni M, Baumann U, Schilling RK, Plett DC, Roy SJ. 2018. Identification And Characterisation Of Proton-Pumping Pyrophosphatase (H+-PPase) Genes In Bread Wheat (Triticum Aestivum). GRDC Update Papers.

Meng L, Li S, Guo J, Guo Q, Mao P, Tian X. 2017. Molecular cloning and functional characterisation of an H+-pyrophosphatase from Iris lacteal. Scientific Reports **7**, 1–10.

Michael TP, Jupe F, Bemm F, Motley ST, Sandoval JP, Lanz C, Loudet O, Weigel D, Ecker JR. 2018. High contiguity Arabidopsis thaliana genome assembly with a single nanopore flow cell. Nature Communications **9**, 1–8.

Mitsuda N, Enami K, Nakata M, Takeyasu K, Sato MH. 2001. Novel Type *Arabidopsis thaliana* H^+-PPase is Localized to the Golgi Apparatus. FEBS Letters **488**, 29–33.

Motta LS, Ramos IB, Machado EA, Santiago MF, Docampo R, Miranda K, Gomes FM, de Souza W, Champagne DE. 2008. Proton-pyrophosphatase and polyphosphate in acidocalcisome-like vesicles from oocytes and eggs of Periplaneta americana. Insect Biochemistry and Molecular Biology **39**, 198–206.

Myers AM, James MG, Yi G. 2011. Maize opaque5 Encodes Monogalactosyldiacylglycerol Synthase and Maize opaque5 Encodes Monogalactosyldiacylglycerol Synthase and Specifically Affects Galactolipids Necessary for Amyloplast and Chloroplast Function.

Nakanishi Y, Saijo T, Wada Y, Maeshima M. 2001. Mutagenic Analysis of Functional Residues in Putative Substrate-binding Site and Acidic Domains of Vacuolar H+-Pyrophosphatase. Journal of Biological Chemistry **276**, 7654–7660.

Narsai R, Edwards JM, Roberts TH, Whelan J, Joss GH, Atwell BJ. 2015. Mechanisms of growth and patterns of gene expression in oxygen-deprived rice coleoptiles. , 25–40.

Ohno S. 1970. *Evolution by Gene Duplication*. Springer, Berlin, Heidelberg.

Paez-Valencia J, Patron-Soberano A, Rodriguez-Leviz A, Sanchez-Lares J, Sanchez-Gomez C, Valencia-Mayoral P, Diaz-Rosas G, Gaxiola R. 2011. Plasma membrane localization of the type I H+-PPase AVP1 in sieve element-companion cell complexes from Arabidopsis thaliana. Plant Science **181**, 23–30.

Pasapula V, Zhu L, Qiu X, *et al.* 2011. Expression of an Arabidopsis vacuolar H+-pyrophosphatase gene (AVP1) in cotton improves drought- and salt tolerance and increases fibre yield in the field conditions. Plant Biotechnology Journal **9**, 88–99.

Pasternak T, Tietz O, Rapp K, Begheldo M, Nitschke R, Ruperti B, Palme K. 2015. Protocol: An improved and universal procedure for whole-mount immunolocalization in plants. Plant Methods **11**, 1–10.

Perata P, Guglielminetti L, Alpi A. 1997. Mobilization of Endosperm Reserves in Cereal Seeds under Anoxia. **79**, 49–56.

Perez-Castineira JR, Losada M, Villalba JM, Lopez-Marques RL, Serrano A. 2002. Functional complementation of yeast cytosolic pyrophosphatase by bacterial and plant H+-translocating pyrophosphatases. Proceedings of the National Academy of Sciences **99**, 15914–15919.

Pizzio GA, Paez-valencia J, Khadilkar AS, *et al.* 2015. Arabidopsis Type I Proton-Pumping Pyrophosphatase Expresses Strongly in Phloem , Where It Is Required for Pyrophosphate Metabolism and. **167**, 1541–1553.

Pomerantz A, Peñafiel N, Arteaga A, Bustamante L, Pichardo F, Coloma LA, Barrio-Amorós CL, Salazar-Valenzuela D, Prost S. 2018. Real-time DNA barcoding in a rainforest using nanopore sequencing: Opportunities for rapid biodiversity assessments and local capacity building. GigaScience **7**, 1–14.

Powell N, Ji X, Ravash R, Edlington J, Dolferus R. 2012. Yield stability for cereals in a changing climate. Functional Plant Biology **39**, 539–552.

Racagni G, Villasuso AL, Pasquaré SJ, Giusto NM, Machado E. 2008. Diacylglycerol pyrophosphate inhibits the α-amylase secretion stimulated by gibberellic acid in barley aleurone. Physiologia Plantarum **134**, 381–393.

Raorane ML, Narciso JO, Kohli A. 2016. Chapter 10 Total Soluble Protein Extraction for Improved Proteomic Analysis of Transgenic Rice Plant Roots. **1385**, 139–147.

Rasse DP, Tocquin P. 2006. Leaf carbohydrate controls over Arabidopsis growth and response to elevated CO2: An experimentally based model. New Phytologist **172**, 500–513.

Rea PA, Poole RJ. 1993. Vacuolar H+ trans locating - pyrophosphatase. Annual Review of Plant Physiology, 157–180.

Regmi KC, Zhang S, Gaxiola RA. 2016. Apoplasmic loading in the rice phloem supported by the presence of sucrose synthase and plasma membrane-localized proton pyrophosphatase. , 257–268.

Roach MJ, Schmidt SA, Borneman AR. 2018. Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. BMC Bioinformatics **19**, 460.

Ros R, Romieu C, Gibrat R, Grignon C. 1995. PPse not transport K+.pdf. Journal of Biological Chemistry **270**, 4368–4374.

Ruiz-López N, Martínez-Force E, Garcés R. 2003. Sequential one-step extraction and analysis of triacylglycerols and fatty acids in plant tissues. Analytical Biochemistry **317**, 247–254.

Rumin J, Bonnefond H, Saint-jean B, Rouxel C, Sciandra A, Bernard O. 2015. The use of fluorescent Nile red and BODIPY for lipid measurement in microalgae. , 1–16.

Sadras VO, McDonald G. 2012. Water use efficiency of grain crops in Australia: principles, benchmarks and management. Grains Research and Development Corporation, 1–28.

Sakakibara Y, Kasamo K, Kobayashi H, Kusakabe I, Kawasaki S. 1999. Identification of the gene structure and promoter region of H+- translocating inorganic pyrophosphatase in rice (Oryza sativa L.). Biochimica et Biophysica Acta - Gene Structure and Expression **1444**, 117–124.

Sakakibara Y, Kobayashi H, Kasamo K. 1996. Isolation and characterization of cDNAs encoding vacuolar H(+)-pyrophosphatase isoforms from rice (Oryza sativa L.). Plant molecular biology **31**, 1029–1038.

Sallaud C, Meynard D, Van Boxtel J, *et al.* 2003. Highly efficient production and characterization of T-DNA plants for rice (Oryza sativa L.) functional genomics. Theoretical and Applied Genetics **106**, 1396–1408.

Sarafian V, Kim Y, Poolet RJ, Rea PA. 1992. Molecular cloning and sequence of cDNA encoding the pyrophosphate-energized vacuolar membrane proton pump of Arabidopsis thaliana (cloning/inorganic pyrophosphatase/nudeotide sequence). Proceedings of the National Academy of Sciences **89**, 1775–1779.

Sato MH, Maeshima M, Ohsumi Y, Yoshida M. 1991. Dimeric structure of H+-translocating pyrophosphatase from pumpkin vacuolar membranes. **290**, 177–180.

Scafaro AP, Gallé A, Van Rie J, Carmo-Silva E, Salvucci ME, Atwell BJ. 2016. Heat tolerance in a wild Oryza species is attributed to maintenance of Rubisco activation by a thermally stable Rubisco activase ortholog. New Phytologist **211**, 899–911.

Schilling RK, Marschner P, Shavrukov Y, Berger B, Tester M, Roy SJ, Plett DC. 2014. Expression of the Arabidopsis vacuolar H + -pyrophosphatase gene (AVP1) improves the shoot biomass of transgenic barley and increases grain yield in a saline field. , 378–386.

Schilling RK, Tester M, Marschner P, Plett DC, Roy SJ. 2017. AVP1: One Protein, Many Roles. Trends in Plant Science.

Scholz-Starke J, Primo C, Yang J, Kandel R, Gaxiola RA, Hirschi KD. 2019. The flip side of the Arabidopsis type I proton-pumping pyrophosphatase (AVP1): Using a transmembrane H gradient to synthesize pyrophosphate. Journal of Biological Chemistry **294**, 1290–1299.

Schoor S, Lung S-C, Sigurdson D, Chuong SDX. 2015. Fluorescent Staining of Living Plant Cells. In: Yeung ECT,, In: Stasolla C,, In: Sumner MJ,, In: Huang BQ, eds. Plant Microtechniques and Protocols. Cham: Springer International Publishing, 153–165.

Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, von Haeseler A, Schatz MC. 2018. Accurate detection of complex structural variations using single-molecule sequencing. Nature methods **15**, 461–468.

Segami S, Makino S, Miyake A, Asaoka M, Maeshima M. 2014. Dynamics of Vacuoles and H+-Pyrophosphatase Visualized by Monomeric Green Fluorescent Protein in Arabidopsis: Artifactual Bulbs and Native Intravacuolar Spherical Structures. The Plant Cell **26**, 3416–3434.

Segami S, Nakanishi Y, Kinoshita S, Asaoka M, Maeshima M, Fukuda M. 2018*a*. Biochemical, Structural and Physiological Characteristics of Vacuolar H+-Pyrophosphatase. Plant and Cell Physiology **59**, 1300–1308.

Segami S, Nakanishi Y, Sato MH, Maeshima M. 2010. Quantification, organ-specific accumulation and intracellular localization of type II H-pyrophosphatase in Arabidopsis thaliana. Plant and Cell Physiology **51**, 1350–1360.

Segami S, Tomoyama T, Sakamoto S, Gunji S, Fukuda M, Kinoshita S, Mitsuda N, Ferjani A, Maeshima M. 2018*b*. Vacuolar H + -Pyrophosphatase and Cytosolic Soluble Pyrophosphatases Cooperatively Regulate Pyrophosphate Levels in Arabidopsis thaliana . The Plant Cell **30**, 1040–1061.

Setter TL, Waters I. 2003. Review of prospects for germplasm improvement for waterlogging tolerance in wheat, barley and oats. Plant and Soil **253**, 1–34.

Seufferheld MJ, Valerio A, Whitfield J, Caetano-Anollés G, Kim K. 2011. Evolution of vacuolar proton pyrophosphatase domains and volutin granules: clues into the early evolutionary origin of the acidocalcisome. Biology Direct **6**, 50.

Shaik SS, Carciofi M, Martens HJ, Hebelstrup KH, Blennow A. 2014. Starch bioengineering affects cereal grain germination and seedling establishment. **65**, 2257–2270.

Sharma S, Borah P, Meena MK, Bindraban P, Pandey R. 2018. Evaluation of genotypic variation for growth of rice seedlings under optimized hydroponics medium. Indian Journal of Genetics and Plant Breeding **78**, 292–301.

Shen G, Wei J, Qiu X, Hu R, Kuppu S, Auld D, Blumwald E, Gaxiola R, Payton P, Zhang H. 2015. Co-overexpression of AVP1 and AtNHX1 in Cotton Further Improves Drought and Salt Tolerance in Transgenic Cotton Plants. Plant Molecular Biology Reporter **33**, 167–177.

Silva P, Gerós H. 2009. Regulation by salt of vacuolar H^+-ATPase and H^+-pyrophosphatase activities and Na^+/H^+ exchange. Plant Signaling and Behavior **4**, 718–726.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva E V., Zdobnov EM. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics **31**, 3210–3212.

Smyth DA, Black CC. 1984. Measurement of the Pyrophosphate Content of Plant Tissues. Plant Physiology **75**, 862–864.

South JB, Morrison WR. 1990. Isolation and analysis of starch from single kernels of wheat and barley. Journal of Cereal Science **12**, 43–51.

Stein JC, Yu Y, Copetti D, *et al.* 2018. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza. Nature Genetics **50**.

Sun T, Xu L, Sun H, Yue Q, Zhai H, Yao Y. 2017. VvVHP1; 2 Is Transcriptionally Activated by VvMYBA1 and Promotes Anthocyanin Accumulation of Grape Berry Skins via Glucose Signal. Frontiers in Plant Science **8**, 1–14.

Taiz BYL. 1992. The plant vacuole. **122**, 113–122.

Takahashi K, Morimoto R, Tabeta H, Asaoka M, Ishida M, Maeshima M, Tsukaya H, Ferjani A. 2017. Compensated Cell Enlargement in fugu5 is Specifically Triggered by Lowered Sucrose Production from Seed Storage Lipids. **58**, 668–678.

Tamazian G, Dobrynin P, Krasheninnikova K, Komissarov A, Koepfli KP, O'Brien SJ. 2016. Chromosomer: A reference-based genome arrangement tool for producing draft chromosome sequences. GigaScience **5**, 1–11.

Tanaka Y, Chiba K, Maeda M, Maeshima M. 1993. Molecular cloning of cDNA for vacuolar membrane proton-translocating inorganic pyrophosphatase in hordeum vulgare. Biochemical and Biophysical Research Communications **190**, 1110–1114.

Thomas P, Sekhar AC. 2014. Live cell imaging reveals extensive intracellular cytoplasmic colonization of banana by normally non-cultivable endophytic bacteria. AoB PLANTS **6**.

Tilman D, Balzer C, Hill J, Befort BL. 2011. Global food demand and the sustainable intensification of agriculture. Proceedings of the National Academy of Sciences **108**, 20260–20264.

Tripathi A, Tripathi DK, Chauhan DK, Kumar N, Singh GS. 2016. Paradigms of climate change impacts on some major food sources of the world: A review on current knowledge and future prospects. Agriculture, Ecosystems and Environment **216**, 356–373.

Ursache R, Andersen TG, Marhavý P, Geldner N. 2018. A protocol for combining fluorescent proteins with histological stains for diverse cell wall components. Plant Journal **93**, 399–412.

Velde F Van De, Riel J Van, Tromp RH. 2002. Visualisation of starch granule morphologies using confocal scanning laser microscopy (CSLM). **1536**, 1528–1536.

Venancio JB, Catunda MG, Ogliari J, Rima JAH, Okorokova-Facanha AL, Okorokov LA, Facanha AR. 2014. A vacuolar H+-pyrophosphatase differential activation and energy coupling integrate the responses of weeds and crops to drought stress. Biochimica et Biophysica Acta - General Subjects **1840**, 1987–1992.

Wang Y, Jin S, Wang M, Zhu L, Zhang X. 2013. Isolation and Characterization of a Conserved Domain in the Eremophyte H+-PPase Family. PLoS ONE **8**.

Wang X, Wang H, Liu S, Ferjani A, Li J, Yan J, Yang X, Qin F. 2016. Genetic variation in ZmVPP1 contributes to drought tolerance in maize seedlings. Nature Genetics **48**, 1233–1241.

Wang Y, Xu H, Zhang G, Zhu H, Zhang L, Zhang Z, Zhang C, Ma Z. 2009. Expression and responses to dehydration and salinity stresses of V-PPase gene members in wheat. Journal of Genetics and Genomics **36**, 711–720.

Wasaya A, Zhang X, Fang Q, Yan Z. 2018. Root Phenotyping for Drought Tolerance: A Review. Agronomy **8**, 241.

Wick R. 2017. Porechop.

Wickham W. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wu Y, Mirzaei M, Pascovici D, Haynes PA, Atwell BJ. 2019. Proteomes of Leaf-Growing Zones in Rice Genotypes with Contrasting Drought Tolerance. Proteomics **19**, 1–11.

Yang Y, Tang RJ, Li B, *et al.* 2015. Research paper Overexpression of a Populus trichocarpa H + -pyrophosphatase gene PtVP1 . 1 confers salt tolerance on transgenic poplar. , 663–677.

Yang Y, Tang RJ, Mu B, Ferjani A, Shi J, Zhang H, Zhao F, Lan WZ, Luan S. 2018. Vacuolar proton pyrophosphatase is required for high magnesium tolerance in Arabidopsis. International Journal of Molecular Sciences **19**, 1–15.

Yang SY, Wu YS, Chen CT, Lai MH, Yen HM, Yang CY. 2017. Physiological and molecular responses of seedlings of an upland rice ('Tung Lu 3') to total submergence compared to those of a submergence-tolerant lowland rice ('FR13A'). Rice **10**.

Yang H, Zhang X, Gaxiola RA, Xu G, Peer WA, Murphy AS. 2014. Over-expression of the Arabidopsis proton-pyrophosphatase AVP1 enhances transplant survival, root mass, and fruit development under limiting phosphorus conditions. Journal of Experimental Botany **65**, 3045–3053.

Yao YX, Dong QL, You CX, Zhai H, Hao YJ. 2011. Expression analysis and functional characterization of apple MdVHP1 gene reveals its involvement in Na +, malate and soluble sugar accumulation. Plant Physiology and Biochemistry **49**, 1201–1208.

Yichie Y, Brien C, Berger B, Roberts TH, Atwell BJ. 2018. Salinity tolerance in Australian wild Oryza species varies widely and matches that observed in O. sativa. Rice **11**, 1–14.

Yin T, Cook D, Lawrence M. 2012. ggbio: an R package for extending the grammar of graphics for genomic data. Genome Biology **13**, R77.

Zabalza A, van Dongen JT, Froehlich A, *et al.* 2008. Regulation of Respiration and Fermentation to Control the Plant Internal Oxygen Concentration. Plant Physiology **149**, 1087–1098.

Zeng F, Konnerup D, Shabala L, Zhou M, Colmer TD, Zhang G, Shabala S. 2014. Linking oxygen availability with membrane potential maintenance and K + retention of barley roots : implications for waterlogging stress tolerance. , 2325–2338.

Zhang Z, Huang R. 2013. Analysis of Malondialdehyde, Chlorophyll Proline, Soluble Sugar, and Glutathione Content in Arabidopsis seedling. Bio-protocol **3**, e817.

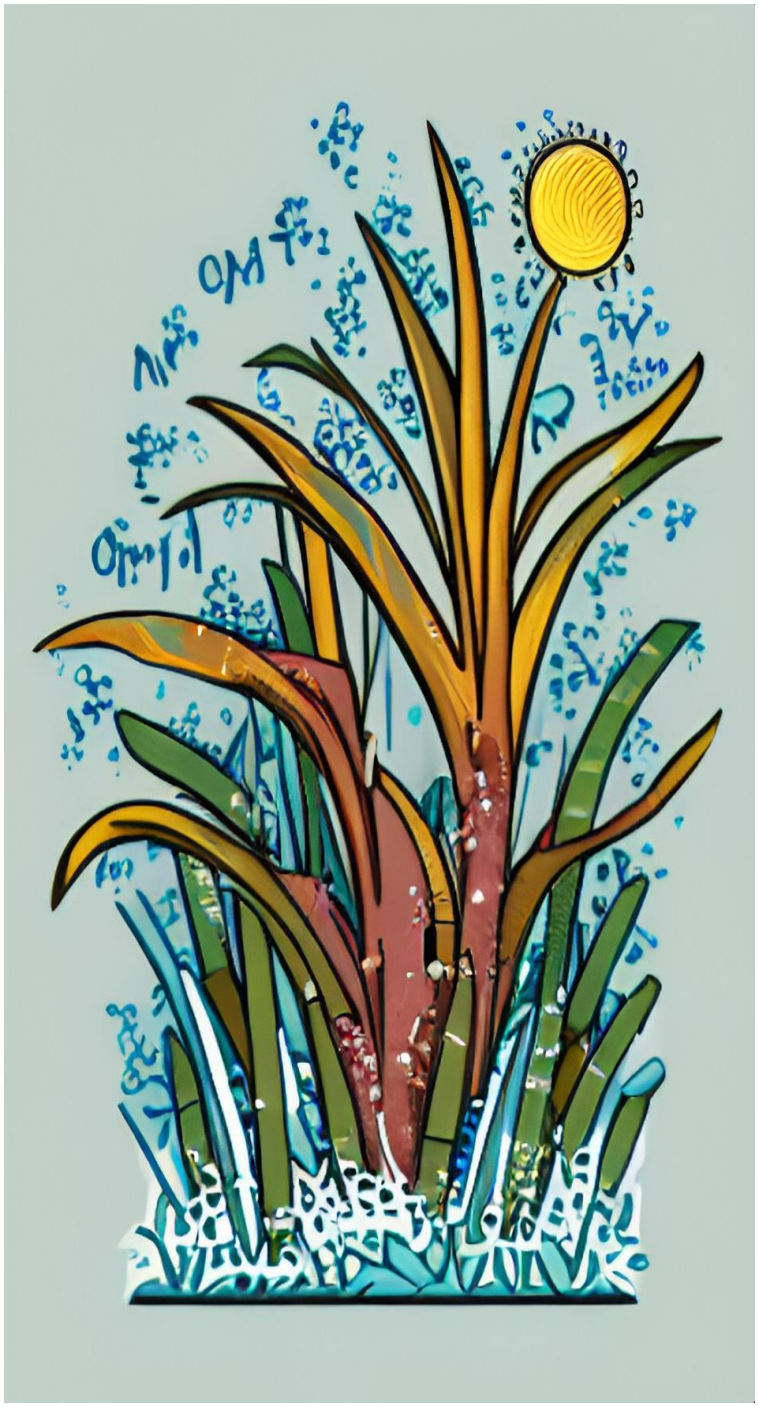
Zhang J, Li J, Wang X, Chen J. 2011*a*. OVP1, a Vacuolar +-translocating inorganic pyrophosphatase (V-PPase), overexpression improved rice cold tolerance. Plant Physiology and Biochemistry **49**, 33–38.

Zhang H, Shen G, Kuppu S, Gaxiola R, Payton P. 2011*b*. Creating drought- and salt-tolerant cotton by overexpressing a vacuolar pyrophosphatase gene. , 861–863.

Zhao FY, Zhang XJ, Li PH, Zhao YX, Zhang H. 2006. Co-expression of the Suaeda salsa SsNHX1 and Arabidopsis AVP1 confer greater salt tolerance to transgenic rice than the single SsNHX1. Molecular Breeding **17**, 341–353.

Appendix III

Literature Review: Improving Australian Rice Cold Tolerance



Wombo Prompt

“Rice paddy frost”

Wombo Theme

“Realistic”

Summary

- Rice is susceptible to cold stress at all growth stages, though the reproductive stage (e.g., the proper functioning of the tapetum in pollen development) is a major bottleneck for improving yield under cold conditions
- A number of Quantitative Trait Loci (QTL) have been identified for cold tolerance at the different stages of rice development
- Tolerance to cold at one growth stage does not always translate to tolerance at later growth stages, which is perhaps why improvements in cold tolerance via breeding have been small and incremental to date
- Many of the genes associated with cold tolerance are involved in cold signalling (e.g., membrane-bound G-proteins) or the activation of cold-responsive genes that lead to cold acclimation (i.e., transcriptions factors), which may suggest that the functional genes involved in cold tolerance exist, but are not switched on, in cold-susceptible varieties
- Key traits associated with cold acclimation are changes in membrane fluidity and the accumulation of osmolytes, both of which affect membrane stability and therefore solute leakage and cell viability
- Improving membrane stability under cold conditions is therefore likely to improve rice cold tolerance, however mechanisms of cold tolerance are highly costly and run counter to mechanisms of heat tolerance and so a constitutively cold-tolerant crop will likely be heat-susceptible
- Genes that increase the ratio of unsaturated:saturated fatty acids in membranes (e.g., OsFAD2) and those that function in osmolyte accumulation (e.g. cell-wall invertases [CWINs] that supply hexoses to developing pollen) are likely to improve membrane stability and cold tolerance, particularly in the tapetum cell layer
- This can be achieved through genetic modification (GM), however, resistance (public and legal) to GM cropping systems is likely to hamper commercialisation (*cv*. Golden Rice), so commercial cold tolerance would have to be approached through non-GM means (e.g. breeding, seed priming, or topical sprays)

- Breeding efforts would have to take some points into consideration:
    - Breeding in multiple QTLs for cold tolerance at multiple growth stages, but particularly during the reproductive stage
    - Cold acclimation should be inducible, not constitutive, to avoid the ongoing cost of cold tolerance and to reduce susceptibility to high temperatures
- Seed priming has not been explored as a route to cold tolerance in rice, though it can ameliorate the impact of other stresses on plants late in their development, likely through epigenetic 'reprogramming'
- There are several commercially available foliar sprays that claim to improve freezing tolerance in several crops, though these could likely improve cold acclimation too, but empirical evidence is lacking for all crop species; foliar sprays could involve hormone-, osmolyte, or miRNA mediated cold acclimation
- To conclude, the utility of GM rice will likely be restricted to the proof-of-concept style studies in a laboratory setting. I think the best hope of improving cold tolerance in Australian rice in a short time frame lies in seed priming or the application of foliar sprays (whether they be cryoprotectant- or miRNA-based) but there are many unanswered questions associated with these two approaches. I would therefore recommend a detailed investigation into the potential of these methods to improve rice cold tolerance.

General Introduction

Temperature is a major determining factor for global species distribution. High latitude plants are more often than not limited in their distribution by extremes of high temperature, while low latitude plants are more limited in their distribution by low temperature extremes (Wen *et al.*, 2018). Due to their evolutionary history, most species in the genus Oryza are adapted to and occur in low latitude environments, which experience warm tropical conditions. Thus, unlike temperate crops such as wheat or barley, rice crops, are susceptible to cold stress. Cold stress, defined as either

chilling (0–15ºC) or freezing (>0ºC) stress, affects every stage of rice development and can reduce yield by 30 to 82% (Rativa *et al.*, 2021) depending on which growth stage is exposed to low temperature conditions (Andaya and Mackill, 2003*a*; Ji *et al.*, 2008; Xu *et al.*, 2008). Temperatures in the range of 12 to 19ºC have negative impacts on germination, seedling establishment, vegetative growth, microspore formation, flowering, and the grain filling period (Satake 1976; (Satake 1976; Ahmed *et al.*, 2008; Ye *et al.*, 2009). Therefore, rice production is limited, in part, by the minimum temperature that occurs during the rice growing season (Farrell *et al.*, 2001).

Rice crops are usually sown in late spring (October in Australia) and the grain is harvested in autumn (March-May in Australia). Thus, vegetative growth, panicle initiation, microsporogenesis, and flowering all occur during the summer months. Grain filling occurs during late summer and grain maturation occurs throughout autumn. Exposure to cold mostly occurs as spring frost events or pre-dawn chilling during early development (germination and seedling stage) and during the reproductive stage (microsporogenesis, flowering, and grain filling). In Australia, a majority of rice production occurs in the New South Wales (NSW) Riverina region (Leeton, Griffith, Deniliquin and Coleambally). Here, according to the Australian Bureau of Meteorology (BOM), lows in the range of 10-17ºC during the rice growing season are common and spring frosts are becoming more frequent and occur later in the season in some parts (BOM *et al.*, 2019). These cold and freezing events have been reported to reduce the yield of Australian rice crops by up to 2.5 t ha^{-1} (Singh *et al.*, 2005). While the effects of cold at each stage of rice development vary, they all manifest as reduced grain yield. It is therefore important to understand how each stage of rice growth is affected by cold if improvements in crop tolerance are to be achieved.

Substantial variation in response to cold stress exists within the Oryza genus. Atwell et al. (2014) identified wild relatives of rice that should, based on their biogeographical context, exhibit tolerance to cold. The authors hypothesised that O. granulata would be cold tolerant (Atwell *et al.*, 2014) though, to the best of our knowledge, this is yet to be tested. Kitazumi et al. (2018) used the species-climate distribution models generated by Atwell et al. (2014) to guide germplasm choice for surveying Oryza cold tolerance. They found that some wild species (i.e., O. australiensis, which is the most high-temperature tolerant Oryza tested to date, and O. barthii) were intolerant to cold stress. However, other wild species, specifically O. longistaminata and O. officinalis, had cold tolerance greater than or similar to domestic rice (O. sativa Nipponbare), respectively.

Given the variety of climates that domestic rice is grown in as a food source, there is also substantial variation in cold tolerance between high-yielding cultivars. For example, O. sativa and O. glaberrima (which were independently domesticated in Asia and Africa, respectively) vary in their tolerance to cold stress, with O. sativa exhibiting greater tolerance to cooler conditions. Within *O. sativa*, Japonica varieties usually demonstrate higher tolerance of cold stress than indica varieties (Cruz and Milach, 2004; Kumar *et al.*, 2017; Dabi and Khanna, 2018). Furthermore, within the Japonica varieties, Nipponbare demonstrates greater cold tolerance than Calrose. Crosses with sensitive (e.g. indica) and tolerant (e.g. Japonica) varieties of domestic rice have provided Quantitative Trait Loci (QTL) associated with increased cold tolerance, though this trait is complicated as different QTL are associated with cold tolerance at different growth stages (reviewed by Najeeb *et al.*, 2021). From these QTL, several candidate genes for improving cold tolerance in rice have been identified. A vast majority of these gene candidates function as transcription factors (TFs) for cold

responsive (COR) genes, or they function in the regulation of cold signal transduction (reviewed by Zhang *et al.*, 2014 and; Lu *et al.*, 2020). For example, the canonical plant cold response involves the activation of COR by C-repeat/dehydration-responsive element binding factor 1 (CBF1). It appears that few 'functional' genes/gene products have been explored for their role in cold tolerance. This is likely because extremes in temperature affect whole-plant physiology – that is, temperature is not a localised stress. Thus, cold tolerance is a complex trait that requires coordinated responses of gene networks rather than individual loci. The complexity of cold tolerance is highlighted by the fact that breeding efforts often fail to make substantial progress towards cold tolerance (Najeeb *et al.*, 2020). Despite this, several phenotypes associated with cold tolerance at different growth stages have emerged that will be explored in this review.

Effects Of Cold On Rice Growth, Development And Yield

Germination, Seedling Establishment, and Vegetative Growth

Germination is a crucial step for the establishment of a crop. Cold stress can delay or completely inhibit rice germination (Yoshida, 1981). The sensitivity of germination to cold stress varies widely between O. sativa cultivars and their hybrids and does not necessarily correlate with tolerance to cold at other growth stages. For example, Cruz and Milach (2004) tested germination in 24 rice genotypes (12 indica varieties and 12 Japonica varieties) at 13ºC for 28 days and show a range of germination responses which did not correlate with previously reported cold tolerance/susceptibility. Furthermore, Lone et al. (2018) tested germination in thirty O. sativa indica genotypes at a constant temperature of 13 or 17ºC for 28 days and found that SKUA-402 had 13% and 9% declines in germination, respectively, while IR-52 suffered 85% and 62% declines in germination, respectively (relative to seeds germinated at 28ºC for seven

days). Interestingly, this study shows what may be different temperature tolerance thresholds between the 30 indica genotypes – that is, the impairment of germination at 17ºC was mild for some genotypes that were amongst the most sensitive to 13ºC, while other lines suffered similarly at both 13C and 17C (Lone *et al.*, 2018). Similarly, Ye et al. (Ye *et al.*, 2009) show substantial variation in germination rates between 17 varieties of rice, including 12 Japonicas, two Japonica x indica crosses, two O. glaberrima x O. sativa crosses, and an upland rice variety from the Philippines. Importantly, regardless of how robust germination was to the 15ºC cold treatment in this experiment, the germination of all lines was delayed significantly (Ye *et al.*, 2009). Reduced or delayed germination can impede seedling establishment and can remove uniformity in a crop, reducing harvest potential. Thus, breeding for improved cold tolerance at the germination stage is critical.

Following germination, the next hurdle that a rice crop faces under cold conditions is seedling establishment. Cold stress slows radicle emergence, and reduces coleoptile, root and shoot elongation rates (Cruz and Milach, 2004; Lone *et al.*, 2018; Najeeb *et al.*, 2021). Shoots undergo chlorosis and eventually senesce under cold stress, resulting in seedling death (Cruz *et al.*, 2013; Hussain *et al.*, 2019). Variation in seedling responses to cold stress leads to non-uniform crop development, which can severely affect the final harvest. There is evidence that tolerance to cold is related to the ability of the seedling to quench reactive oxygen species (ROS) (Zhang *et al.*, 2014). The concentration of ROS scavenging compounds increase in rice plants exposed to low temperature (Sato *et al.*, 2011). Teixeira et al. (2021) showed that soaking rice seed in a carrot root extract high in antioxidant phenolic compounds and carotenoids improved shoot and, to a lesser extent, root elongation in rice seedlings grown at 15ºC. Given that seedling-level tolerance to cold stress is highly correlated

with tolerance to cold stress at the reproductive stages of rice development, it is likely that improving seedling tolerance to cold will also lead to improvements in grain yield (Sharma *et al.*, 2021). Thus, there have been a number of QTL identified that are associated with seedling stage cold tolerance (see Najeeb *et al.*, 2021).

Once seedlings have established, the crop enters a period of vegetative growth. During this stage rice plants produce tillers, which will eventually produce the panicles on which flowers develop and grain maturation occurs. Under cold stress conditions, the rate of tillering is reduced, though final tiller numbers are similar under control and cold conditions (Ye *et al.*, 2009). Also, cold stress significantly alters energy metabolism (carbon catabolism and anabolism) in rice plants. Photosynthetic capacity of the rice canopy is reduced by reducing chlorophyll content in leaves (Kumar and Mahajan, 2007; Xiaochuang *et al.*, 2017). Chlorophyll fluorescence and synthesis, and the dark reactions of photosynthesis are inhibited by low temperatures, and thylakoid membranes become leaky (Xiaochuang *et al.*, 2017; van Buer *et al.*, 2019; Adamski *et al.*, 2020). As a result, each rice plant has less leaf area for photosynthesis and reduced photosynthetic rates. Improving photosynthetic rates under cold stress conditions is one route towards improving rice cold stress tolerance during the early seedling and vegetative stages. Mitochondrial structure is altered by cold stress and respiration rates decline (Hussain *et al.*, 2019). Panicle number per plant is not affected by cold, though panicles are smaller and the number of seed per panicle is significantly reduced as investment in reproductive tissues is reduced (Ye *et al.*, 2009).

Microspore Formation and the Tapetum

Panicle initiation, or ‘booting’, signifies the start of the reproductive growth phase in rice crops. The reproductive stage of rice development is the most susceptible to cold

stress (Thakur *et al.*, 2010). Losses in crop yield of up to 3.5 tonnes per hectare have been recorded following chilling events during the reproductive phase (Imin *et al.*, 2004). In particular, the tetrad stage of microsporogenesis, in which mother microsporocytes undergo meiosis to produce daughter haploid microspore cells, is extremely temperature sensitive (Fig. 1; Imin *et al.*, 2004; Sharma and Nayyar, 2016). Microspores undergo mitosis to produce mature pollen. Thus, any stress that interferes with the production of microspores/pollen can render a rice plant sterile.

Cold stress increases the accumulation of abscisic acid (ABA) in the anthers of cold sensitive rice cultivars (Sharma and Nayyar, 2016). ABA is thought to be a signal for cold sensitivity as it acts to repress the expression of cell wall invertases (CWINs) in the tapetum (Fig. 1; Thakur *et al.*, 2010; Ji *et al.*, 2011; Sharma and Nayyar, 2016). The tapetum is a layer of cells in anther tissue that provide nourishment to developing microspores/pollen grains and is pivotal to rice microspore cold tolerance (Parish *et al.*, 2012). CWINs hydrolyse sucrose in the tapetum cell wall to glucose and fructose, which can then be transported to and consumed by the microspores (Parish *et al.*, 2012). By limiting CWIN expression, ABA reduces the accumulation of hexose sugars in the tapetum cell wall and therefore interferes with nutrient supply to the microspores (Ji *et al.*, 2011). Thus, the microspores are starved of an energy source and are aborted (Fig. 1). Cold stress also leads to hypertrophy (abnormal enlargement) of the tapetum and delays tapetum programmed cell death (PCD), both of which can lead to the abortion of microspores and increased plant sterility (Niu *et al.*, 2013). However, Yamamori et al. (2021) show that whole-anther growth abnormalities correlate more strongly with pollen sterility than tapetum morphology alone. Even though whole-locule morphology appears to be a better indicator of pollen fertility than tapetum hypertrophy, we cannot ignore the role of the tapetum in nourishing the developing microspores.

Indeed, the three models of pollen sterility at low temperature proposed by Yamamori et al. (2021) all involve loss of the tapetum.

While loss of the tapetum is required for normal pollen development, this process must occur via a tightly coordinated PCD pathway mediated by ROS (Xu *et al.*, 2020). Excessive ROS leads to premature tapetum PCD, while reduced ROS leads to delayed tapetum PCD – both scenarios lead to reduced pollen viability. Xu et al. (2020) show that a single point mutation in an unknown locus they identify as “low- temperature tolerance 1” (LTT1) leads to increased cold tolerance in japonica rice. Seed set was ~80% in *ltt1* compared to ~8% in the wild type (WT) under cold stress, though no difference was detected under normal conditions. High base-line levels of ROS in *ltt1* lead to increased accumulation of ROS scavenging enzymes. Thus, when a cold period occurs ROS fails to accumulate in the tapetum (as it is catabolised by ROS scavenging enzymes) and normal tapetum PCD can occur, which facilitates normal pollen development. In the WT plant, bursts of ROS under cold-stress lead to premature PCD in the tapetum and then sterile pollen (Xu *et al.*, 2020). Interestingly, Nipponbare had significantly reduced pollen fertility after four days at 12ºC even though it demonstrates robust photosynthetic capacity after three days at 10ºC during the seedling stage (de Freitas *et al.*, 2019; Yamamori *et al.*, 2021). Hoshinoyume, a temperate japonica, and, surprisingly, Silewah, a tropical japonica, demonstrated the highest tolerance (normal locule morphology, unchanged pollen fertility, unchanged anther length, unchanged anther morphology, and pollen number) to cold during the booting stage (Yamamori *et al.*, 2021).

QTL And Genes Associated With Cold Tolerance In Rice

578 QTL associated with cold tolerance have been identified (Lv *et al.*, 2019; Najeeb *et al.*, 2021). They are relatively evenly spread across all 12 of the *O. sativa* chromosomes. Different stages of rice development have different QTL associated with cold tolerance, indicating that cold tolerance is a complex trait that manifests differently depending on developmental stage (Zhang *et al.*, 2014; Lv *et al.*, 2019; Najeeb *et al.*, 2021).

Several QTL associated with increased germination under cold stress have been identified (Lv *et al.*, 2019; Najeeb *et al.*, 2021). Najeeb et al. (2021) reviewed the 578 cold tolerance associated QTL and found that 214 of these are associated with germination stage tolerance to low temperature. These 214 QTL are spread across all 12 of the O. sativa chromosomes. Chromosome seven possesses the highest proportion of cold tolerance QTL (~24) and chromosome eight possesses the least (~12) (Najeeb *et al.*, 2021). Fujino et al. (2004) reveal that single a QTL (that accounted for 30% of phenotypic variation under cold stress) is sufficient for cold-tolerant germination. However, metabolic analyses indicate that response to cold stress at the germination stage is a complex process involving large-scale gene expression changes (Fujino and Matsuda, 2010). The complexity of cold tolerance at the germination stage was confirmed by Najeeb et al. (2020) who identify 16 genes from cold germination- (and seedling vigour-) associated QTLs. The 16 genes are involved in processes including transcription, protein ubiquitination, and protein phosphorylation, all of which play crucial roles in cold perception and signalling (Zhang *et al.*, 2014; Lu *et al.*, 2020). QTL associated with germination under low temperature conditions are often colocalised with QTL associated with seedling survival (Najeeb *et al.*, 2021).

Roughly 249 QTL associated with rice seedling cold tolerance have been identified (Lv *et al.*, 2019; Najeeb *et al.*, 2021). The contribution of each QTL to cold tolerance varies widely (from ~4% up to ~61%). Many of the QTL that co-localise with QTL associated with cold tolerance at other growth stages account for ~40% of the phenotypic variation in cold tolerance at the seedling stage (Andaya and Mackill, 2003*b*; Pan *et al.*, 2015; Yang *et al.*, 2017). Fine mapping of two of these QTL revealed that a calmodulin-binding transcription activator, involved in cold signal transduction, and a Duf6-containing gene, which likely functions in metabolite transport, are associated with cold tolerance at the seedling stage (Kim *et al.*, 2014). Some of the QTL for seedling stage tolerance to cold overlap with QTL associated with tolerance at the germination and reproductive stages as well.

The remaining 115 QTL are associated with cold tolerance at the reproductive stage (Najeeb *et al.*, 2021). This stage is defined by the booting stage, in which the panicles (fruiting bodies) of the rice plant start to grow up through the inside of the stem causing the stem to bulge, the heading stage, in which the panicle is fully emerged from the stem, and flowering, in which flower buds open and begin dispersing pollen grains. Microspore development, anther morphology, panicle emergence, timing of heading, and pollen viability are all negatively affected by exposure to cold stress during the reproductive phase. As discussed above, microspore development and subsequent pollen viability are perhaps the most sensitive to low temperature. Up to 71% of variation in cold tolerance in the booting can be explained by a single QTL (Najeeb *et al.*, 2021), while other QTL account for as little as 2.94% of variation.

Despite the vast number of identified QTL and TFs associated with rice cold tolerance, candidate functional genes for cold tolerance are limited (Xu *et al.*, 2020). This is likely because cold tolerance, which affects all above-ground plant tissues at the same time, is a complex trait. However, even heat stress has a well-defined set of genes that can be augmented for improved heat tolerance. There has therefore been a call for uncovering specific genes/gene families that can be used to improve rice cold stress tolerance at one or more growth stages (Xu *et al.*, 2020). Najeeb et al. (2021) identified 38 candidate genes using the Q-TARO website (http://qtaro.abr.affrc.go.jp/ogro). Interestingly, no genes were identified that were associated with cold tolerance at all three of the growth stages discussed here. Despite the fact that the reproductive phase is the most sensitive to cold stress, only two genes could be identified that contribute to cold tolerance: OsAsr1 (an ABA, stress, and ripening-induced gene) and OsLti6b (small membrane proteins involved in maintaining membrane stability under dehydration stress) (Morsy *et al.*, 2005; Kim *et al.*, 2009). Most of the other genes (25/38) were associated with seedling-stage tolerance. The disparity between functional gene lists for seedling and reproductive stage cold tolerance likely results from a bias in the literature towards seedling experiments. There is therefore a need for '-omics' studies into cold tolerance focussed on the male reproductive tissues (as the ovaries and ova are not sensitive to cold) as it is clear that this stage of the rice life cycle is most prone to cold damage and has the greatest impact on yield components.

Effect Of Cold On Plant Cells

While it may not be known exactly which genes directly affect cold tolerance, there are several phenotypes associated with cold sensitivity that provide insight into the processes which may be augmented to improve cold tolerance. For example, cold stress induces physiological drought and so tolerance to cold, or cold acclimation,

involves maintenance of membrane stability, accumulation of osmolytes, and maintenance of enzymatic activity, like tolerance to drought (Zhang *et al.*, 2019). It is important to note, however, that while drought and cold acclimation and/or tolerance share physiological similarities, not all drought-related traits will enhance cold tolerance, and vice versa.

Membrane Fluidity

Depending on the rapidity of cold exposure, biological membranes enter either a rigid liquid-crystalline (rapid onset of cold) or a thickened gel-like (gradual onset of cold) state (Murata *et al.*, 1982; Zhang *et al.*, 2019). Both states lead to increased solute leakage and ultimately cause cell death, which may translate to tissue death in the absence of recovery. Rapid-onset cold stress causes membrane bursting, and cell contents are released as a result (Shi *et al.*, 2014; Hsu and Hsu, 2019; Zhang *et al.*, 2019). Cold-tolerant plants are capable of recovery from this kind of injury in a short period of time, however cold-sensitive plants usually experience localised or whole-tissue death from rapid-onset cold stress (Zhang *et al.*, 2019). On the other hand, when cold exposure happens gradually and a membrane enters a gel-like state, permeability increases (usually measured as electrolyte leakage or MDA accumulation) (Hongtao *et al.*, 2017; Zhang *et al.*, 2019; Bai *et al.*, 2021). This increases the movement of solutes across membranes and can start to eliminate the chemical disequilibrium necessary for several biological processes, like ATP and NADPH synthesis. In fact, one of the most widely documented responses to cold is the increased expression of ATP synthases, likely to account for reduced ATP synthesis (Zhang *et al.*, 2017; Yang *et al.*, 2020). Ca^{2+} influx is a signal of cold stress for plant cells, leading to the upregulation of cold-responsive genes, because Ca^{2+} accumulates in plant cells due to increased membrane permeability (Guo *et al.*, 2018). Furthermore, the activity of

membrane-bound proteins is reduced in gel-like membranes, which can disrupt metabolic processes, like solute transport, and leads to the accumulation of cellular waste products (Zhang *et al.*, 2019).

In response to cold-induced changes in membrane permeability, cells can adjust the ratios and kinds of biomolecules (fatty acids, sterols, and proteins) in their membranes (Janská *et al.*, 2010; Rogowska and Szakiel, 2020). At low temperature, it is important to buffer the effects of a more viscous membrane by returning the membrane to a more fluid-like state. This process is called ‘cold acclimation’ (Janská *et al.*, 2010; Rogowska and Szakiel, 2020). Cold tolerant plants reduce the ratio of sterols to phospholipids in their membranes, thus increasing membrane fluidity (Palta *et al.*, 1993; Rogowska and Szakiel, 2020). Plants that are cold-sensitive often have higher proportions of saturated fatty acids in their membranes under cold stress and increasing abundance of unsaturated fatty acids is often related to increased cold tolerance (Miquel *et al.*, 1993; Chen and Thelen, 2013; Barrero-Sicilia *et al.*, 2017). This is because unsaturated lipids increase the fluidity of membranes due to the double bonds in their structure reducing the packing density of a membrane (Ballweg *et al.*, 2020). Saturated lipids, on the other hand, are more structurally linear because they lack double bonds, and can therefore be packed at a higher density in a membrane, resulting in reduced membrane fluidity (Ballweg *et al.*, 2020). Thus, it might be possible to increase the tolerance of rice plants to cold by changing the relative abundances of saturated and unsaturated lipids, and sterols, in membranes after the plant experiences cold stress. The timing of lipid content change is important as the current model of cold perception and signalling positions membranes at the beginning of many cold-response pathways that lead to cold acclimation in cold-tolerant species (Janská *et al.*, 2010; Barrero-Sicilia *et al.*, 2017). Thus, any gene products that cause changes in membrane components would

have to be under the control of an inducible promoter that is active only after cold is detected.

Osmolytes and Osmoprotectants

Exposure to cold induces physiological drought (Takahashi *et al.*, 2013; Zhang *et al.*, 2019). That is, cells dehydrate under cold stress, and so many of the responses to cold are analogous to drought responses (Takahashi *et al.*, 2013). Interestingly, the osmotic component of cold stress means that halophytic species are poised for strong cold acclimation responses (Barrero-Sicilia *et al.*, 2017). For example, osmolytes accumulate under drought, salt, and cold conditions. Also, gene expression is similar in plant experiencing drought and cold conditions. This may help cells combat ion and water efflux due to increasingly leaky and rigid membranes. Sugars and sugar alcohols, amino acids, polyamines, and betaines are all osmolytes that have been found to improve plant tolerance to cold stress.

One of the key compounds upregulated in response to cold is trehalose – a nonreducing disaccharide comprised of two glucose monomers joined by a 1-1 glycosidic bond (Iordachescu and Imai, 2008; Hasanuzzaman *et al.*, 2019). Trehalose can mitigate the effects desiccation stress as it is able to enter a ‘glassy’ state in which it can compensate for lost water (Bhandari and Nayyar, 2014; Hasanuzzaman *et al.*, 2019). In doing so, desiccation-sensitive enzymes and membranes are made more stable. Increasing the accumulation of trehalose increases the tolerance of both vegetative and reproductive tissues to cold (Nayyar *et al.*, 2005; Li *et al.*, 2011; Hasanuzzaman *et al.*, 2019). Also, because trehalose is usually present in low abundance in a cell, its increased accumulation under cold conditions serves to

activate signal transduction pathways that eventuate in the synthesis of stress-response enzymes and other osmoprotectants (Iordachescu and Imai, 2008).

Proline, an amino acid, is also upregulated in plants during cold exposure (Bhandari and Nayyar, 2014; Hasanuzzaman *et al.*, 2019). Like trehalose, it functions in the activation of signal transduction pathways that lead to cold acclimation in cold tolerant plants (Hayat *et al.*, 2012; Bhandari and Nayyar, 2014). Proline also plays roles in osmotic adjustment, which could help a cell resist ion leakage and water loss during cold stress (Hasanuzzaman *et al.*, 2019). Further, proline can function as a chaperone for enzymes, stabilising their quaternary structures under stress conditions (Hayat *et al.*, 2012; Bhandari and Nayyar, 2014). In the context of cold tolerance, proline is likely to exert its influence most strongly on membrane bound proteins which are affected both by changes in membrane fluidity (Banu *et al.*, 2009). For example, photosystem II, a regulator of cold tolerance in plants, and photosynthetic pigments are stabilised by proline (Banu *et al.*, 2009). Proline also helps to buffer changes in cytosolic pH, which is challenged by changes in membrane stability under cold and is crucial for continued metabolism (Hayat *et al.*, 2012; Hasanuzzaman *et al.*, 2019).

Polyamines are another class of osmoprotectant that improve tolerance to abiotic stresses (Groppa and Benavides, 2008; Hasanuzzaman *et al.*, 2019). In plants, there are three major polyamines: spermine, spermidine, and putrescine (Groppa and Benavides, 2008). There is evidence that these compounds improve drought and salt tolerance in rice (Farooq *et al.*, 2010; Roychoudhury *et al.*, 2011). However, spermine is re-emerging as an important player in tolerance to range of abiotic stresses where it functions in antioxidant defence (Hasan *et al.*, 2021). Spermine conjugates are highly accumulated in male reproductive tissue of maize (Vogt, 2018) and, in rice, reduced

levels of spermine leads to loss of pollen viability (Chen *et al.*, 2014). These plants are also highly susceptible to cold, salt and drought stress, have reduced water content, increased ion leakage, reduced photosynthetic capacity, and reduced expression of antioxidant enzymes (Chen *et al.*, 2014). Furthermore, the anthers of rice plants classified as cytoplasmic male-sterile (CMS) have reduced levels of the three major polyamines (spermine, spermidine, and putrescine) compared to wild type anthers (Chengye and Xianfeng, 1993). Polyamine downregulation is also observed in the anthers of CMS soybeans (Ding *et al.*, 2019). Spermine may therefore be a good candidate for improving rice tolerance to cold stress.

Glycine betaine (GB), a derivative of methylated glycine, can accumulate in leaves in response to osmotic stress and extremes in temperature (Bhandari and Nayyar, 2014; Hasanuzzaman *et al.*, 2019). Interestingly, GB has been shown to have protective effects that are specific to plant reproductive organs in transgenic plants grown under cold conditions (Chen and Murata, 2008; Bhandari and Nayyar, 2014). While GB does not naturally accumulate to the levels observed in these studies, external application of GB improves the tolerance of rice crops to cold, osmotic stresses and water deficit (Chen and Murata, 2008). In wild-type plants, GB is localised mainly to the chloroplasts during stress events where it functions to stabilise thylakoid membranes and the enzymes involved in CO2 fixation (Wang *et al.*, 2008*b*). Thylakoid membranes are susceptible to cold stress because of the changes that occur in membrane fluidity under cold conditions (van Buer *et al.*, 2019). GB also serves as an osmoprotectant for chloroplasts, which is important because photosynthesis is highly sensitive to the changes induced by cold stress (Chen and Murata, 2008).

Photosynthesis – balancing heat and cold tolerance

CO_2 assimilation is a highly-temperature sensitive set of reactions (Sharkey and Zhang, 2010). Extremes in heat and cold both significantly reduce photosynthetic capacity in domestic rice by affecting membrane integrity (particularly the thylakoid membrane) and reducing the catalytic efficiency of rate-limiting enzymes (e.g., Rubisco) (Hussain *et al.*, 2019). Chlorophyll fluorescence is often used as a measure of photosynthetic cold tolerance as photochemical quenching is reduced (and therefore fluorescence is increased) in cold-sensitive plants (Fracheboud *et al.*, 1999; Mishra *et al.*, 2014). The net effect of reduced photosynthetic rate varies depending on the growth stage affected by cold. In the vegetative stage, growth declines can be seen (Phillips et al., in press), while in the reproductive stage yield and grain filling can be reduced (Kumar and Mahajan, 2007). Despite the impact that cold has on plant cell membranes, the Calvin-Benson cycle (in which CO_2 is fixed) is also susceptible to low temperatures (Ruelland *et al.*, 2009). However, at low temperature ROS are generated at the thylakoid membranes due to the accumulation of the reducing compounds (NADPH) that are ordinarily consumed by the Calvin-Benson cycle, which can reduce photochemical quenching of photosynthetically active radiation and can also lead to chlorosis (Ruelland *et al.*, 2009).

Key to photosynthetic thermotolerance (that is, tolerance to high temperature) is enzyme stability as the excess energy in a heated leaf causes enzymes (e.g. like Rubisco and Rubisco activase (RCA)) to lose catalytic efficiency (Salvucci and Crafts-Brandner, 2004; Scafaro *et al.*, 2016). To increase the stability of an enzyme under heat, they can be made more rigid through changes in enzyme primary structure (i.e., through artificial or natural selection or via genetic engineering) (Shivhare and Mueller-Cajar, 2017; Scafaro *et al.*, 2018, 2019). Thus, it may be possible to modify enzyme

primary structure via site-directed mutagenesis to *reduce* rigidity and allow enzymes to function normally at cooler temperatures (Georlette *et al.*, 2004). If natural selection has already acted on key photosynthetic enzymes for cold tolerance (for example in wild species that are adapted to cool conditions), then it may also be possible to introduce those enzymes into a cold-sensitive genome (Atwell *et al.*, 2014). For example, it may be possible to introduce an isoform of RCA from Antarctic hair grass (Salvucci and Crafts-Brandner, 2004) to improve carbon fixation at low temperature. This may, however, render the plants susceptible to extreme heat due to opposing effects of high and low temperature on cell biology.

Similarly, improving the heat tolerance of membranes can be achieved by making them less fluid (Kan and Lin, 2021), while improving the tolerance of a membrane to cooler conditions involves making it more fluid (Cruz *et al.*, 2013). These outcomes can be achieved by altering the stoichiometry of membrane components (Cruz *et al.*, 2013; Rogowska and Szakiel, 2020; Kan and Lin, 2021). Increased membrane stability at low temperature is likely to improve the efficiency of the ‘light’ reactions of photosynthesis, which would lead to a stable supply of ATP and NADPH for the ‘dark’ reactions of photosynthesis under the same conditions. However, a heat tolerant membrane is likely to be cold sensitive, and vice versa. Improving rice photosynthetic cold tolerance at the cost of photosynthetic heat tolerance would be especially problematic for crops in warmer climates (e.g., Australia) because much of the vegetative stage occurs during summer. Further, a cold tolerant photosynthesis runs counter with the goal of introducing a more C4-like photosynthetic system in rice.

Transgenic Rice Cold Tolerance

Based on the cold-induced phenotypes discussed above, several transgenic rice lines have been made that demonstrate tolerance to cold conditions. A vast majority of these lines have been modified to overexpress genes that facilitate the perception, signalling, and/or activation of cold-responsive pathways (Zhang *et al.*, 2014; Lu *et al.*, 2020). For example, the overexpression of OsDREB (Dehydration-Responsive Element-Binding) transcription factors (TFs) activate gene networks that already exist in sensitive cultivars, but which have lost their responsiveness to cold because of reduced levels of the TFs in question (Wang *et al.*, 2008*a*). These lines of rice often have the same or similar phenotypes under cold conditions – increased osmolyte production and stable membranes (Liu *et al.*, 2007; Yarra and Wei, 2021). As discussed above, while these studies are interesting, it would be more informative to know which functional genes are important for cold tolerance as TFs have pleiotropic effects. This section of the review will therefore focus on non-TF transgenic lines of rice that result in stable or improved survival, growth and/or yield under cold conditions.

As discussed above, several cold-associated QTL have been identified in rice, though relatively few genes have been identified from these QTL (Najeeb *et al.*, 2021). However, several metabolites have been found to improve cold tolerance. For example, brassinosteroids (BRs) and polyols, like myo-inositol, increase in abundance during cold exposure in tolerant plants and can therefore be augmented in sensitive plants to improve tolerance (Hasanuzzaman *et al.*, 2019). Zhao et al. (2017) identified a gene of unknown function, Os09g0410300 in the cold QTL qCTS-9. There is some evidence that the gene product of Os09g0410300 interacts with Brassinosteroid Insensitive 1 (BRI1), a plasma membrane leucine-rich repeat receptor kinase that functions in the BR pathway (Hirabayashi *et al.*, 2004). The exact nature of this

interaction is unknown, though changes in the expression of genes involved in BR synthesis can enhance cold tolerance (Fujii and Saka, 2001; Koh *et al.*, 2007; Divi and Krishna, 2010). In plants, BRs are perceived on the cell surface by BR receptors (like BR1) which activates signal transduction pathways that increase the expression of genes via the TF, Brassinazole Resistant 1 (BZR1) (Ahammed *et al.*, 2020). BZR1-activated genes improve vegetative growth under cold conditions by reducing the inhibition of photosystem II (PSII) by ROS, increasing chlorophyll content, and increasing the accumulation of osmolytes (Singh *et al.*, 2012; Zhao *et al.*, 2019). Oryza officinalis achieves cold tolerance similar to that of O. sativa via a brassinosteroid-mediated response pathway, which has been fragmented in O. sativa (Kitazumi *et al.*, 2018). BRs can also interact with the CBF1 pathway to induce the expression of COR genes that regulate acclimation to cold stress (Kitazumi *et al.*, 2018; Ramirez and Poppenberger, 2020). BRs are also crucial for returning plants to a 'normal' state once cold stress has passed as they interact with and reduce the activity of Inducer of CBF Expression 1 (ICE1), another cold-responsive TF that activates CBF1 genes (Ye *et al.*, 2019). Zhao et al. (2017) show that electrolyte leakage is significantly reduced in Os09g0410300 over-expressing lines exposed to cold stress (relative to the parent line). Also, the susceptible parent line plants had zero survival when exposed to cold during the seedling stage, whereas over-expressor lines had up to 85% survival (Zhao *et al.*, 2017). It would be interesting to see if the over-expression of Os09g0410300, the putative BR-interacting protein, changes post-cold recovery by modulating the activity of ICE1.

Indica rice, which is more cold-sensitive than its japonica relatives, has been the focus of many cold-tolerance transgenic studies. Liao et al. (2017) achieved cold tolerance in Pei'ai 64S (O. sativa ssp. Indica) at the vegetative stage via the overexpression of

OsTMP14 – a component of the thylakoid membrane-bound photosystem 1 (PS1). OsTMP14 is involved in response to several stresses, including drought and salt, and is thought to stabilise light-harvesting complex (LHC) 1 (Subramanyam *et al.*, 2010; Liao *et al.*, 2017). Liao et al. (2017) report increased expression of OsTMP14 in leaves of Pei'ai 64S at the booting stage under cold conditions. Subsequent overexpression of OsTMP14 lead to reduced chlorosis and increased survival under cold stress. Given that cold reduces thylakoid membrane stability and that the activity of membrane proteins is affected by the stability of the membranes they are embedded in, PS1 may be operating inefficiently at low temperatures in indica rice due to membrane destabilisation reducing the efficiency of the LHC1 redox reactions. Liao et al. (2017) did not test chlorophyll fluorescence in the OsTMP14 over-expression lines nor in the parent line, though Adamski et al. (2020) show that cold tolerant indica lines have greater photochemical efficiency at low temperature when compared to their cold-sensitive counterparts. Interestingly, Adamski et al. (2020) report higher water-use efficiency in cold-tolerant indica rice due to a lower rate of transpiration brought about due to stomatal closure. This may implicate the involvement of ABA in indica rice cold tolerance. However, as discussed above, ABA can also play a role in inducing cold sensitivity in the reproductive phase due to its effects on the expression of genes in the tapetum.

An important aspect of cold acclimation is increasing membrane stability. This can be achieved by altering membrane components (Rogowska and Szakiel, 2020). Cold tolerant plants can acclimate to cold conditions by increasing the ratio of unsaturated to saturated fatty acids in their membranes (Barrero-Sicilia *et al.*, 2017). Improving membrane stability is important for maintaining the activity of membrane-bound proteins and preventing loss of ions and water to leakage (Barrero-Sicilia *et al.*, 2017).

The transition from saturated to unsaturated fatty acid composition occurs through the activity of fatty acid desaturases (FADs) (Miquel *et al.*, 1993; Barrero-Sicilia *et al.*, 2017). For example, transgenic rice overexpressing OsFAD2 are more cold-tolerant than the parent line (Shi *et al.*, 2012). These plants convert oleic acid (with one unsaturated double bond) to linoleic acid via a desaturase reaction that introduces an additional double (unsaturated) bond into the fatty acid chain. Germination at 4oC is improved by ~4-6%, photochemical quenching is improved by ~100%, and electrolyte leakage is reduced by ~25% (Shi *et al.*, 2012). Similarly, the over-expression of GmFAD3A (a fatty acid desaturase from Glycine max) in rice increased the abundance of polyunsaturated fatty acids (PUFAs) and lead to increased germination rates and seedling survival at 15oC (Wang *et al.*, 2019). MDA content was reduced, indicating reduced damage to membranes, probably due to the increased activity of antioxidant enzymes and increased accumulation of osmoprotectants (Hasanuzzaman *et al.*, 2019; Wang *et al.*, 2019). PUFAs, like linoleic acid, can also affect grain yield under cold stress. For example, yield is increased ~46% in rice plants overexpressing OsFAD2 when exposed to relatively low temperature (20oC) during flowering, relative to the parent line (Shi *et al.*, 2012). However, the OsFAD2 overexpressing plants still experience a ~60% decline in yield at 20oC relative to control conditions.

Tocopherols (vitamin E) are chloroplast-synthesised lipid-soluble non-enzymatic antioxidants that offer protection against cold stress by reducing damage to PUFAs from ROS, and thus they improve membrane stability (Fukuzawa *et al.*, 1982; Soll *et al.*, 1985; Munné-Bosch and Alegre, 2002). They are not major antioxidants for plants under optimal conditions, instead they function in response to stress and exert influence over rice development (Kanwischer *et al.*, 2005; Zhang *et al.*, 2018). In wheat, tocopherol accumulation in leaves is associated with cold tolerance, and there is

speculation that tocopherol content may bolster the cold tolerance of other cereals (Janeczko *et al.*, 2018). Enzymes involved in tocopherol biosynthesis are crucial for cold tolerance in rice, and, interestingly, exert influence over plant growth and yield under optimal conditions too (Wang *et al.*, 2017; Zhang *et al.*, 2018). SGD-1 (knockout) mutants have a dwarf phenotype, cells in developing grain are smaller, and 1000-grain weight is reduced by ~50% (Wang *et al.*, 2017). There was no survival in sgd-1 mutants following a 2-week recovery period after being exposed to 4oC for three days (Wang *et al.*, 2017). Interestingly, sdg-1 plants had a higher base level of proline than the parent line (Wang *et al.*, 2017). Proline accumulation is usually associated with stress tolerance as it accumulates under stress in tolerant varieties (Hasanuzzaman *et al.*, 2019). The accumulation of proline in these mutants might indicate stress perception even under optimal conditions. Further, the relative conductivity of the growth solution in which the plants were grown increased from 5% to 55% after one day of cold exposure, implicating membrane permeability in the cold-susceptibility of the sgd-1 mutants (Wang *et al.*, 2017). Further, overexpression of tocopherol cyclase, an enzyme that catalyses the synthesis of tocopherols in a number of rice tissues, including anthers, lead to increased stress tolerance (Ouyang *et al.*, 2011). There is also interest in increasing the concentration of tocopherols in cereal grains for human health (Munné-Bosch and Alegre, 2002), and this trait may also increase the cold tolerance of the tapetum membrane during the microspore stage.

Transgenic studies have also uncovered both positive and negative roles for micro-RNAs (miRNAs) in rice cold tolerance (Wang *et al.*, 2014; Sun *et al.*, 2020). Sun et al. (2020) showed that miR535 is a negative regulator of cold tolerance in rice. On the other hand, Wang et al. (2014) showed that miR319b is a positive regulator of cold tolerance in rice. The miRNA family that includes miR535 are believed to target and

silence SPL transcripts (Sun *et al.*, 2020). SPL transcripts are involved in a diverse set of plant phenotypes ranging from plant height to tolerance to drought (Ning *et al.*, 2017; Sun *et al.*, 2019). When miR535 was overexpressed in an elite cold-tolerant rice variety, Kongyu 131 (O. sativa japonica), seedling survival was reduced by ~40% and ion leakage was increased by ~20% (Sun *et al.*, 2020). Importantly, the wild type seedlings and the transgenic seedlings had similar phenotypes under control conditions, implying a specific role of miR535 in cold acclimation. Interestingly, the activities of antioxidant enzymes were increased in the over-expressor seedlings, and soluble sugars accumulated, however proline, a key regulator of tolerance to several abiotic stresses, was reduced in the over-expressor lines (Sun *et al.*, 2020). The abundance of SPL transcripts was reduced in miR535 over-expressor lines and the expression of CBF2 and CBF3, both important for activating the expression of COR genes (Guo *et al.*, 2018), and was also reduced (Sun *et al.*, 2020). On the other hand, the miR319 family is thought to target and silence the transcription factors OsPCF6 and OsTCP21 (Wang *et al.*, 2014). Overexpression of miR319b lead to increased plant survival, increased proline accumulation, and increased antioxidant activities in rice under cold stress (Wang *et al.*, 2014). Thus, miR535 might play an important role in returning plants to a pre-cold physiological state to avoid the cost of maintaining cold acclimation once the stress has passed, while miR319b might function in the activation of cold responsive pathways.

There are therefore targets for improving cold tolerance in rice in a commercial setting. However, given that yield is reduced by 82% in glasshouse conditions and 30% in the field under cold conditions (Rativa *et al.*, 2021), these studies, which mostly occur in glasshouses, should be extended to field trials to see if the reported phenotypes, especially subtle ones, persist. Furthermore, cold acclimation, and therefore cold

tolerance, is a costly process for plants because of the synthesis of osmoprotectants and membrane components (Browse and Lange, 2004). Given the cost of cold acclimation and potential to reduce vigour under normal conditions or high temperature (Kasuga *et al.*, 1999), it seems that any genetic modifications in rice will have to be under the control of an inducible system rather than being constitutively expressed. Even if this were to be achieved, legislature (and public opinion) against GM cropping systems would likely mean that a transgenic cold-tolerant crop would not be feasible in many areas (*cv*. Golden Rice) (Glover *et al.*, 2020), even though laws have relaxed in recent years in Australia. Alternatively, improving a rice crops tolerance to cold may be achieved through non-genetic modification means.

Non-GM Methods To Improve Rice Cold Tolerance

Much of the published research conducted on rice is concerned with tolerance to temperatures below 19oC because rice is, broadly, a tropical species. There is a strong academic interest in cold tolerant rice, and these studies make use of genetically modified rice varieties to address the biology of cold acclimation. However, GM rice varieties will likely not achieve commercial viability (Glover *et al.*, 2020) and so it important to explore non-GM methods of improving cold tolerance. For example, rice growers in Australia may benefit from shifting production to the northern, tropical regions of Australia. Alternatively, wild rice relatives native to the northern regions of Australia (i.e., O. australiensis and O. meridionalis), which thrive under the water and temperature regimes of the regions they are adapted to, may be bred for improved agronomic traits (Henry *et al.*, 2010; Brar and Khush, 2018). Further, incorporating cold-tolerant wild relatives of rice, such as O. granulata (discussed above) is another non-GM avenue that could be explored (Atwell *et al.*, 2014). However, shifting the Australian rice industry northward may not be feasible, and there are many practical

uncertainties around domesticating wild species. Other options may include seed priming, or the application of topical sprays (e.g., cryoprotectants) to enhance cold tolerance.

Seed priming involves exposing seeds to chemical compounds with the aim of enhancing germinability, seedling establishment, and stress tolerance. The effects of seed priming may also exert influence over traits that occur towards the end of plants life cycle (e.g., yield components) (Kant *et al.*, 2003, 2004; Kaur *et al.*, 2005; Farooq *et al.*, 2008). Seed priming techniques include hydropriming (essentially imbibition), chemical priming, osmopriming, hormone priming, halopriming, and nutrient priming (Lal *et al.*, 2018). In rice, seed priming with selenium or salicylic acid can enhance the cold tolerance of germinating seeds and seedlings by, ultimately, reducing damage to membranes by ROS (Pouramir-Dashtmian *et al.*, 2014; Hussain *et al.*, 2016; Wang *et al.*, 2016). These primed plants had higher antioxidant defence capability compared to control plants (Pouramir-Dashtmian *et al.*, 2014), lower rates of membrane damage, and they maintained starch metabolism and respiration rates (Hussain *et al.*, 2016). Furthermore, priming rice seed with spermidine and trehalose (both discussed above) lead to rice cold tolerance (15oC) by increasing the concentration of soluble sugars in seedlings and increasing antioxidant enzyme activity, but via different mechanisms (Fu *et al.*, 2020). That is, priming with spermidine increased the concentration of spermidine in seedlings and increased the expression of spermidine synthesis genes while inhibiting the expression of trehalose synthesis genes, and vice versa. Finally, abiotic stress tolerance was also improved in rice plants primed with brassinosteroids (discussed above) (Basit *et al.*, 2021). Seed priming with these relatively cheap chemicals may allow for increased cold tolerance at the industrial level.

This review has focussed on cold stress and cold acclimation. Companies that produce cryoprotectants claim that their products (there are at least 20 commercially available cryoprotectants) can protect crops from sub-zero temperatures by enhancing *freezing* tolerance, facilitating *freezing* avoidance, or via a combination of both (reviewed recently by Román-Figueroa *et al.*, 2021). However, empirical evidence supporting these claims is sparce and, to the best of our knowledge, none of these products have been used to enhance cold tolerance nor have they been tested in cereals at any growth stage (for example, Francko *et al.*, 2011). Some sprays contain compounds known to be important for cold acclimation in cereal crops (e.g., tocopherol, trehalose; discussed above) so there is a possibility of cold acclimation following treatment (Román-Figueroa *et al.*, 2021). Even though there are some key differences between cold and freezing responses in plants, there are shared mechanisms for reducing the impact of both stresses on crop performance. For example, osmolyte accumulation and improving membrane stability (via ROS quenching and changing lipid profiles) are both important features of acclimation to both freezing and cold stress (Ruelland *et al.*, 2009) and it is claimed that both can be achieved using commercially available topical sprays. Liang et al. (2021) found that exogenous application of trehalose rescued cold-induced spikelet infertility in wheat due to increased accumulation of spermidine, glutathione, and aspartic acid quenching ROS. The use of topical sprays would also allow producers to account for the large temperature fluctuations that characterise springtime growth periods as they can be applied as needed. However, it is unknown how the cryoprotectants would interact with high temperatures if they persist on the leaf surface. Further, given that the microspore stage of the rice reproductive phase is most susceptible to cold, and these products are typically foliage sprays, it is not known if the cryoprotectants could alleviate the damage induced by low temperature on the

internal anther tissue prior to anthesis, though results from wheat are promising (Liang *et al.*, 2021).

Finally, it may be possible to prime a crop against cold stress by foliar application of miRNA. Traditionally, crop species can be genetically modified to encode RNA genes that function in the silencing of a trait of interest. However, Mitter et al. (2017) demonstrated that foliar application of double stranded RNA (dsRNA) loaded onto BioClay can afford a crop protection from viral pathogens for up to 20 days. This is a non-GM technique that results in RNA interference (RNAi) of target RNA molecules. The dsRNA is taken up by plant cells and is processed into small interfering RNA (siRNA) in the plant (Mitter *et al.*, 2017; Betti *et al.*, 2021). In this way, viral RNA is targeted for degradation by siRNA-argonaut complexes, and the plant is protected (Muhammad *et al.*, 2019; Betti *et al.*, 2021). Other uses for this technology include arming plants with insecticidal siRNA to protect against wasp pests and improving crop tolerance to pathogens (Almasi and Amari, 2021; Balassa *et al.*, 2021; Palmer *et al.*, 2021). Research into the application of this technology for enhancing crop abiotic stress tolerance is virtually non-existent. As discussed above, miR319 acts as a positive regulator of cold acclimation in rice by silencing the RNA transcripts of two TFs, OsPCF6 and OsTCP21 (Wang *et al.*, 2014). It might be possible to coat the BioClay developed by Mitter et al. (2017) with miR319 to silence OsPCF6 and OsTCP21 and induce cold acclimation in rice.

Summary And Conclusion

Rice is susceptible to cold stress at all growth stages, though the reproductive stage (e.g., the proper functioning of the tapetum in pollen development) is a major bottleneck for improving yield under cold conditions. A number of Quantitative Trait Loci (QTL)

have been identified for cold tolerance at the different stages of rice development. Tolerance to cold at one growth stage does not always translate to tolerance at later growth stages, which is perhaps why improvements in cold tolerance via breeding have been small and incremental to date. Many of the genes associated with cold tolerance are involved in cold signalling (e.g., membrane-bound G-proteins) or the activation of cold-responsive genes that lead to cold acclimation (i.e., transcriptions factors), which may suggest that the functional genes involved in cold tolerance exist, but are not switched on, in cold-susceptible varieties. Key traits associated with cold acclimation are changes in membrane fluidity and the accumulation of osmolytes, both of which affect membrane stability and therefore solute leakage and cell viability. Improving membrane stability under cold conditions is therefore likely to improve rice cold tolerance, however mechanisms of cold tolerance are highly costly and run counter to mechanisms of heat tolerance and so a constitutively cold-tolerant crop will likely be heat-susceptible. Genes that increase the ratio of unsaturated:saturated fatty acids in membranes (e.g., OsFAD2) and those that function in osmolyte accumulation (e.g. cell-wall invertases [CWINs] that supply hexoses to developing pollen) are likely to improve membrane stability and cold tolerance, particularly in the tapetum cell layer. This can be achieved through genetic modification (GM), however, resistance (public and legal) to GM cropping systems is likely to hamper commercialisation (*cv*. Golden Rice), so commercial cold tolerance would have to be approached through non-GM means (e.g. breeding, seed priming, or topical sprays). Breeding efforts would have to take some points into consideration. First, breeding in multiple QTLs for cold tolerance at multiple growth stages, but particularly during the reproductive stage. Second, cold acclimation should be inducible, not constitutive, to avoid the ongoing cost of cold tolerance and to reduce susceptibility to high temperatures. Further, seed priming has not been explored as a route to cold tolerance in rice, though it can ameliorate the

impact of other stresses on plants late in their development, likely through epigenetic 'reprogramming'. Finally, there are several commercially available foliar sprays that claim to improve freezing tolerance in several crops, though these could likely improve cold acclimation too, but empirical evidence is lacking for all crop species; foliar sprays could involve hormone-, osmolyte-, or miRNA-mediated cold acclimation.

To conclude, the utility of GM rice will likely be restricted to proof-of-concept style studies in a laboratory setting. However, there may be an increasing willingness to develop GM food crops in Australia. Even so, I think the best hope of improving cold tolerance in Australian rice in a short time frame lies in seed priming or the application of foliar sprays (whether they be cryoprotectant- or miRNA-based) but there are many unanswered questions associated with these two approaches. I would therefore recommend a detailed investigation into the potential of these three methods (GM, seed priming, and foliar sprays) to improve rice cold tolerance.

References:

Adamski JM, Rosa LMG, Menezes Peixoto CR de, Pinheiro CL, Fett JP, Sperotto RA. 2020. Photosynthetic activity of indica rice sister lines with contrasting cold tolerance. Physiology and Molecular Biology of Plants **26**, 955–964.

Ahammed GJ, Li X, Liu A, Chen S. 2020. Brassinosteroids in Plant Tolerance to Abiotic Stress. Journal of Plant Growth Regulation **39**, 1451–1464.

Ahmed N, Maekawa M, Tetlow IJ. 2008. Effects of low temperature on grain filling, amylose content, and activity of starch biosynthesis enzymes in endosperm of basmati rice. Australian Journal of Agricultural Research **59**, 599–604.

Almasi R, Amari K. 2021. Harnessing environmental-RNAi to win an enduring battle: Plant versus pathogen. JOURNAL OF PHYTOPATHOLOGY.

Andaya VC, Mackill DJ. 2003*a*. QTLs conferring cold tolerance at the booting stage of rice using recombinant inbred lines from a japonica x indica cross. TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik **106**, 1084–1090.

Andaya VC, Mackill DJ. 2003*b*. Mapping of QTLs associated with cold tolerance during the vegetative stage in rice. Journal of experimental botany **54**, 2579–2585.

Atwell BJ, Wang H, Scafaro AP. 2014. Could abiotic stress tolerance in wild relatives of rice be used to improve Oryza sativa? Plant Science **215–216**, 48–58.

Bai H, Lin P, Li X, *et al.* 2021. DgC3H1, a CCCH zinc finger protein gene, confers cold tolerance in transgenic chrysanthemum. Scientia Horticulturae **281**, 109901.

Balassa K, Balassa G, Gondor OK, Janda T, Almasi A, Rudnoy S. 2021. Changes in physiology, gene expression and ethylene biosynthesis in MDMV-infected sweet corn primed by small RNA pre-treatment. SAUDI JOURNAL OF BIOLOGICAL SCIENCES **28**, 5568–5578.

Ballweg S, Sezgin E, Doktorova M, Covino R, Reinhard J, Wunnicke D, Hänelt I, Levental I, Hummer G, Ernst R. 2020. Regulation of lipid saturation without sensing membrane fluidity. Nature Communications **11**, 756.

Banu MstNA, Hoque MdA, Watanabe-Sugimoto M, Matsuoka K, Nakamura Y, Shimoishi Y, Murata Y. 2009. Proline and glycinebetaine induce antioxidant defense gene expression and suppress cell death in cultured tobacco cells under salt stress. Journal of Plant Physiology **166**, 146–156.

Barrero-Sicilia C, Silvestre S, Haslam RP, Michaelson L V. 2017. Lipid remodelling: Unravelling the response to cold stress in Arabidopsis and its extremophile relative Eutrema salsugineum. Plant Science **263**, 194–200.

Basit F, Chen M, Ahmed T, *et al.* 2021. Seed Priming with Brassinosteroids Alleviates Chromium Stress in Rice Cultivars via Improving ROS Metabolism and Antioxidant Defense Response at Biochemical and Molecular Levels. Antioxidants **10**.

Betti F, Ladera-Carmona MJ, Weits DA, *et al.* 2021. Exogenous miRNAs induce post-transcriptional gene silencing in plants. NATURE PLANTS **7**.

Bhandari K, Nayyar H. 2014. Low Temperature Stress in Plants: An Overview of Roles of Cryoprotectants in Defense BT - Physiological Mechanisms and Adaptation Strategies in Plants Under Changing Environment: Volume 1. In: Ahmad P,, In: Wani MR, eds. New York, NY: Springer New York, 193–265.

BOM, CSIRO, FarmLink. 2019. *Regional weather and climate guide: A guide to weather and climate in the Riverina*.

Brar DS, Khush GS. 2018. Wild relatives of rice: a valuable genetic resource for genomics and breeding research. In: Mondal TK,, In: Henry RJ, eds. The Wild Oryza Genomes. Cham: Springer International Publishing, 1–25.

Browse J, Lange BM. 2004. Counting the cost of a cold-blooded life: Metabolomics of cold acclimation. Proceedings of the National Academy of Sciences of the United States of America **101**, 14996 LP – 14997.

van Buer J, Prescher A, Baier M. 2019. Cold-priming of chloroplast ROS signalling is developmentally regulated and is locally controlled at the thylakoid membrane. Scientific Reports **9**, 3022.

Chen M, Chen J, Fang J, Guo Z, Lu S. 2014. Down-regulation of S-adenosylmethionine decarboxylase genes results in reduced plant length, pollen

viability, and abiotic stress tolerance. Plant Cell, Tissue and Organ Culture (PCTOC) **116**, 311–322.

Chen THH, Murata N. 2008. Glycine betaine: an effective protectant against abiotic stress in plants. Trends in Plant Science **13**, 499–505.

Chen M, Thelen JJ. 2013. ACYL-LIPID DESATURASE2 Is Required for Chilling and Freezing Tolerance in Arabidopsis . The Plant Cell **25**, 1430–1444.

Chengye L, Xianfeng C. 1993. Polyamine contents and antioxidative metabolism in anthers of cytoplasmic male-sterlie rice lines and their corresponding maintainer lines. Journal of Plant Physiology and Molecular Biology **19**, 82–88.

Cruz RP da, Milach SCK. 2004. Cold tolerance at the germination stage of rice: methods of evaluation and characterization of genotypes. Scientia Agricola **61**, 1–8.

Cruz RP da, Sperotto RA, Cargnelutti D, Adamski JM, de FreitasTerra T, Fett JP. 2013. Avoiding damage and achieving cold tolerance in rice plants. Food and Energy Security **2**, 96–119.

Dabi T, Khanna V. 2018. Effect of Climate Change on Rice. Agrotechnology **07**.

Ding X, Wang X, Li Q, Yu L, Song Q, Gai J, Yang S. 2019. Metabolomics Studies on Cytoplasmic Male Sterility during Flower Bud Development in Soybean. International Journal of Molecular Sciences **20**.

Divi UK, Krishna P. 2010. Overexpression of the Brassinosteroid Biosynthetic Gene AtDWF4 in Arabidopsis Seeds Overcomes Abscisic Acid-induced Inhibition of Germination and Increases Cold Tolerance in Transgenic Seedlings. Journal of Plant Growth Regulation **29**, 385–393.

Farooq M, Basra SMA, Rehman H, Saleem BA. 2008. Seed Priming Enhances the Performance of Late Sown Wheat (Triticum aestivum L.) by Improving Chilling Tolerance. Journal of Agronomy and Crop Science **194**, 55–60.

Farooq M, Wahid A, Lee D-J, Cheema SA, Aziz T. 2010. DROUGHT STRESS: Comparative Time Course Action of the Foliar Applied Glycinebetaine, Salicylic Acid, Nitrous Oxide, Brassinosteroids and Spermine in Improving Drought Resistance of Rice. Journal of Agronomy and Crop Science **196**, 336–345.

Farrell TC, Fox KM, Williams RL, Fukai S, Reinke RF, Lewin LG. 2001. Temperature constraints to rice production in Australia and Laos: a shared problem. Increased lowland rice production in the Mekong Region: Proceedings of an International Workshop held in Vientiane, Laos, 30 October-2 November 2000. Canberra: Australian Centre for International Agricultural Research (ACIAR), 129–137.

Fracheboud Y, Haldimann P, Leipner J, Stamp P. 1999. Chlorophyll fluorescence as a selection tool for cold tolerance of photosynthesis in maize (Zea mays L.). Journal of Experimental Botany **50**, 1533–1540.

Francko DA, Wilson KG, Li QQ, Equiza MA. 2011. A Topical Spray to Enhance Plant Resistance to Cold Injury and Mortality. HortTechnology **21**.

de Freitas GM, Thomas J, Liyanage R, Lay JO, Basu S, Ramegowda V, do Amaral MN, Benitez LC, Bolacel Braga EJ, Pereira A. 2019. Cold tolerance response mechanisms revealed through comparative analysis of gene and protein expression in multiple rice genotypes. PloS one **14**, e0218019.

Fu Y, Zhang Z, Liu J, Chen M, Pan R, Hu W, Guan Y, Hu J. 2020. Seed Priming with Spermidine and Trehalose Enhances Chilling Tolerance of Rice via Different Mechanisms. Journal of Plant Growth Regulation **39**, 669–679.

Fujii S, Saka H. 2001. The Promotive Effect of Brassinolide on Lamina Joint-Cell Elongation, Germination and Seedling Growth under Low-Temperature Stress in Rice (Oryza sativa L.). Plant Production Science **4**, 210–214.

Fujino K, Matsuda Y. 2010. Genome-wide analysis of genes targeted by qLTG3-1 controlling low-temperature germinability in rice. Plant Molecular Biology **72**, 137–152.

Fujino K, Sekiguchi H, Sato T, Kiuchi H, Nonoue Y, Takeuchi Y, Ando T, Lin SY, Yano M. 2004. Mapping of quantitative trait loci controlling low-temperature germinability in rice (Oryza sativa L.). Theoretical and Applied Genetics **108**, 794–799.

Fukuzawa K, Tokumura A, Ouchi S, Tsukatani H. 1982. Antioxidant activities of tocopherols on Fe2+-ascorbate-induced lipid peroxidation in lecithin liposomes. Lipids **17**, 511–513.

Georlette D, Blaise V, Collins T, D'Amico S, Gratia E, Hoyoux A, Marx J-C, Sonan G, Feller G, Gerday C. 2004. Some like it cold: biocatalysis at low temperatures. FEMS Microbiology Reviews **28**, 25–42.

Glover D, Kim SK, Stone GD. 2020. Golden Rice and technology adoption theory: A study of seed choice dynamics among rice growers in the Philippines. Technology in Society **60**, 101227.

Groppa MD, Benavides MP. 2008. Polyamines and abiotic stress: Recent advances. AMINO ACIDS **34**, 35–45.

Guo X, Liu D, Chong K. 2018. Cold signaling in plants: Insights into mechanisms and regulation. Journal of Integrative Plant Biology **60**, 745–756.

Hasan MdM, Skalicky M, Jahan MS, Hossain MdN, Anwar Z, Nie Z-F, Alabdallah NM, Brestic M, Hejnak V, Fang X-W. 2021. Spermine: Its Emerging Role in Regulating Drought Stress Responses in Plants. Cells **10**.

Hasanuzzaman M, Anee TI, Bhuiyan TF, Nahar K, Fujita M. 2019. Emerging Role of Osmolytes in Enhancing Abiotic Stress Tolerance in Rice. In: Hasanuzzaman M,, In: Fujita M,, In: Nahar K,, In: Biswas JKBT-A in RR for AST, eds. Woodhead Publishing, 677–708.

Hayat S, Hayat Q, Alyemeni MN, Wani AS, Pichtel J, Ahmad A. 2012. Role of proline under changing environments. Plant Signaling & Behavior **7**, 1456–1466.

Henry RJ, Rice N, Waters DLE, Kasem S, Ishikawa R, Hao Y, Dillon S, Crayn D, Wing R, Vaughan D. 2010. Australian Oryza: utility and conservation. Rice **3**, 235–241.

Hirabayashi S, Matsushita Y, Sato M, Oh-I R, Kasahara M, Abe H, Nyunoya H. 2004. Two Proton Pump Interactors Identified from a Direct Phosphorylation Screening of a Rice cDNA Library by Using a Recombinant BRI1 Receptor Kinase. Plant Biotechnology **21**, 35–45.

Hongtao X, Tongtong W, Dianfeng Z, *et al.* 2017. ABA pretreatment enhances the chilling tolerance of a chilling-sensitive rice cultivar. Brazilian Journal of Botany **40**, 853–860.

Hsu CH, Hsu YT. 2019. Biochemical responses of rice roots to cold stress. Botanical Studies **60**, 14.

Hussain S, Khaliq A, Ali B, Hussain HA, Qadir T, Hussain S. 2019. Temperature Extremes: Impact on Rice Growth and Development. In: Hasanuzzaman M,, In: Hakeem KR,, In: Nahar K,, In: Alharby HF, eds. Plant Abiotic Stress Tolerance: Agronomic, Molecular and Biotechnological Approaches. Cham: Springer International Publishing, 153–171.

Hussain S, Khan F, Hussain HA, Nie L. 2016. Physiological and Biochemical Mechanisms of Seed Priming-Induced Chilling Tolerance in Rice Cultivars . Frontiers in Plant Science **7**, 116.

Imin N, Weinman JJ, Rolfe BG. 2004. *Investigation of the molecular mechanisms causing cold induced sterility in rice*. Canberra, Australia.

Iordachescu M, Imai R. 2008. Trehalose Biosynthesis in Response to Abiotic Stresses. Journal of Integrative Plant Biology **50**, 1223–1229.

Janeczko A, Dziurka M, Pociecha E. 2018. Increased leaf tocopherol and β-carotene content is associated with the tolerance of winter wheat cultivars to frost. Journal of Agronomy and Crop Science **204**, 594–602.

Janská A, Maršík P, Zelenková S, Ovesná J. 2010. Cold stress and acclimation – what is important for metabolic adjustment? Plant Biology **12**, 395–405.

Ji X, Dong B, Shiran B, Talbot MJ, Edlington JE, Hughes T, White RG, Gubler F, Dolferus R. 2011. Control of abscisic acid catabolism and abscisic acid homeostasis is important for reproductive stage stress tolerance in cereals. Plant physiology **156**, 647–662.

Ji S-L, Jiang L, Wang Y-H, Liu S-J, Liu X, Zhai H-Q, Yoshimura A, Wan J-M. 2008. QTL and epistasis for low temperature germinability in rice. Acta Agronomica Sinica **34**, 551–556.

Kan Y, Lin H-X. 2021. Molecular regulation and genetic control of rice thermal response. The Crop Journal **9**, 497–505.

Kant S, Pahuja SS, Pannu RK. 2004. Effect of seed priming on growth and phenology of wheat under late-sown conditions. Tropical Science **44**, 9–15.

Kant S, Verma P, Pahuja SS. 2003. Growth and yield maintenance in bread wheat by seed priming under late-sown conditions. Acta Agronomica Hungarica **51**, 445–453.

Kanwischer M, Porfirova S, Bergmüller E, Dörmann P. 2005. Alterations in Tocopherol Cyclase Activity in Transgenic and Mutant Plants of Arabidopsis Affect Tocopherol Content, Tocopherol Composition, and Oxidative Stress. Plant Physiology **137**, 713–723.

Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K. 1999. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nature biotechnology **17**, 287–291.

Kaur S, Gupta AK, Kaur N. 2005. Seed Priming Increases Crop Yield Possibly by Modulating Enzymes of Sucrose Metabolism in Chickpea. Journal of Agronomy and Crop Science **191**, 81–87.

Kim S-J, Lee S-C, Hong SK, An K, An G, Kim S-R. 2009. Ectopic expression of a cold-responsive OsAsr1 cDNA gives enhanced cold tolerance in transgenic rice plants. Molecules and cells **27**, 449–458.

Kim S-M, Suh J-P, Lee C-K, Lee J-H, Kim Y-G, Jena KK. 2014. QTL mapping and development of candidate gene-derived DNA markers associated with seedling cold tolerance in rice (Oryza sativa L.). Molecular genetics and genomics : MGG **289**, 333–343.

Kitazumi A, Pabuayon ICM, Ohyanagi H, *et al.* 2018. Potential of Oryza officinalis to augment the cold tolerance genetic mechanisms of Oryza sativa by network complementation. Scientific reports **8**, 16346.

Koh S, Lee S-C, Kim M-K, Koh JH, Lee S, An G, Choe S, Kim S-R. 2007. T-DNA tagged knockout mutation of rice OsGSK1, an orthologue of Arabidopsis BIN2, with enhanced tolerance to various abiotic stresses. Plant molecular biology **65**, 453–466.

Kumar M, Gho YS, Jung KH, Kim SR. 2017. Genome-wide identification and analysis of genes, conserved between japonica and indica rice cultivars, that respond to low-temperature stress at the vegetative growth stage. Frontiers in Plant Science **8**, 1–20.

Kumar P, Mahajan V. 2007. Effect of low temperature stress on photosynthesis, total soluble sugars, grain filling rate and yield in rice (Oryza sativa L.). Indian journal of plant physiology **12**, 253–260.

Lal SK, Kumar S, Sheri V, Mehta S, Varakumar P, Ram B, Borphukan B, James D, Fartyal D, Reddy MK. 2018. Seed Priming: An Emerging Technology to Impart Abiotic Stress Tolerance in Crop Plants. In: Rakshit A,, In: Singh HB, eds. Advances in Seed Priming. Singapore: Springer Singapore, 41–50.

Li H-W, Zang B-S, Deng X-W, Wang X-P. 2011. Overexpression of the trehalose-6-phosphate synthase gene OsTPS1 enhances abiotic stress tolerance in rice. Planta **234**, 1007–1018.

Liang Z, Luo J, Wei B, Liao Y, Liu Y. 2021. Trehalose can alleviate decreases in grain number per spike caused by low-temperature stress at the booting stage by

promoting floret fertility in wheat. Journal of Agronomy and Crop Science **207**, 717–732.

Liao Y, Jiang Y, Xu J, *et al.* 2017. Overexpression of a thylakoid membrane protein gene OsTMP14 improves indica rice cold tolerance. Biotechnology & Biotechnological Equipment **31**, 717–724.

Liu K, Wang L, Xu Y, Chen N, Ma Q, Li F, Chong K. 2007. Overexpression of OsCOIN, a putative cold inducible zinc finger protein, increased tolerance to chilling, salt and drought, and enhanced proline level in rice. Planta **226**, 1007–1016.

Lone J, Khan MN, Bhat MA, Shikari A, Wani S, R. Sofi N, Imran Khan M, Lone R. 2018. Cold tolerance at germination and seedling stages of rice: methods of evaluation and characterization of thirty rice genotypes under stress conditions. International Journal of Current Microbiology and Applied Sciences **7**, 1103–1109.

Lu X, Zhou Y, Fan F, Peng J, Zhang J. 2020. Coordination of light, circadian clock with temperature: The potential mechanisms regulating chilling tolerance in rice. Journal of integrative plant biology **62**, 737–760.

Lv Y, Hussain MA, Luo D, Tang N. 2019. Current understanding of genetic and molecular basis of coldtolerance in rice. Molecular Breeding **39**, 159.

Miquel M, James D, Dooner H, Browse J. 1993. Arabidopsis requires polyunsaturated lipids for low-temperature survival. Proceedings of the National Academy of Sciences **90**, 6208 LP – 6212.

Mishra A, Heyer AG, Mishra KB. 2014. Chlorophyll fluorescence emission can screen cold tolerance of cold acclimated Arabidopsis thaliana accessions. Plant Methods **10**, 38.

Mitter N, Worrall EA, Robinson KE, Li P, Jain RG, Taochy C, Fletcher SJ, Carroll BJ, Lu GQ (Max), Xu ZP. 2017. Clay nanosheets for topical delivery of RNAi for sustained protection against plant viruses. Nature Plants **3**, 16207.

Morsy MR, Almutairi AM, Gibbons J, Yun SJ, de Los Reyes BG. 2005. The OsLti6 genes encoding low-molecular-weight membrane proteins are differentially expressed in rice cultivars with contrasting sensitivity to low temperature. Gene **344**, 171–180.

Muhammad T, Zhang F, Zhang Y, Liang Y. 2019. RNA Interference: A Natural Immune System of Plants to Counteract Biotic Stressors. Cells **8**, 38.

Munné-Bosch S, Alegre L. 2002. The Function of Tocopherols and Tocotrienols in Plants. Critical Reviews in Plant Sciences **21**, 31–57.

Murata N, Sato N, Takahashi N, Hamazaki Y. 1982. Compositions and Positional Distributions of Fatty Acids in Phospholipids from Leaves of Chilling-Sensitive and Chilling-Resistant Plants. Plant and Cell Physiology **23**, 1071–1079.

Najeeb S, Ali J, Mahender A, Pang YL, Zilhas J, Murugaiyan V, Vemireddy LR, Li Z. 2020. Identification of main-effect quantitative trait loci (QTLs) for low-temperature stress tolerance germination- and early seedling vigor-related traits in

rice (Oryza sativa L.). Molecular breeding : new strategies in plant improvement **40**, 10.

Najeeb S, Mahender A, Anandan A, Hussain W, Li Z, Ali J. 2021. Genetics and breeding of low-temperature stress tolerance in rice. In: Ali J,, In: Wani SH, eds. Rice Improvement: Physiological, Molecular Breeding and Genetic Perspectives. Cham: Springer International Publishing, 221–280.

Nayyar H, Bains T, Kumar S. 2005. Low temperature induced floral abortion in chickpea: relationship to abscisic acid and cryoprotectants in reproductive organs. Environmental and Experimental Botany **53**, 39–47.

Ning K, Chen S, Huang H, Jiang J, Yuan H, Li H. 2017. Molecular characterization and expression analysis of the SPL gene family with BpSPL9 transgenic lines found to confer tolerance to abiotic stress in Betula platyphylla Suk. Plant Cell, Tissue and Organ Culture (PCTOC) **130**, 469–481.

Niu N, Liang W, Yang X, Jin W, Wilson ZA, Hu J, Zhang D. 2013. EAT1 promotes tapetal cell death by regulating aspartic proteases during male reproductive development in rice. Nature Communications **4**, 1445.

Ouyang S, He S, Liu P, Zhang W, Zhang J, Chen S. 2011. The role of tocopherol cyclase in salt stress tolerance of rice (Oryza sativa). Science China. Life sciences **54**, 181–188.

Palmer S, Dearden PK, Mercier OR, King-Hunt A, Lester PJ. 2021. Gene drive and RNAi technologies: a bio-cultural review of next-generation tools for pest wasp management in New Zealand. Journal of the Royal Society of New Zealand.

Palta JP, Whitaker BD, Weiss LS. 1993. Plasma Membrane Lipids Associated with Genetic Variability in Freezing Tolerance and Cold Acclimation of Solanum Species. Plant Physiology **103**, 793–803.

Pan Y, Zhang H, Zhang D, *et al.* 2015. Genetic analysis of cold tolerance at the germination and booting stages in rice by association mapping. PloS one **10**, e0120590.

Parish RW, Phan HA, Iacuone S, Li SF. 2012. Tapetal development and abiotic stress: a centre of vulnerability. Functional plant biology : FPB **39**, 553–559.

Pouramir-Dashtmian F, Khajeh-Hosseini M, Esfahani M. 2014. Improving chilling tolerance of rice seedling by seed priming with salicylic acid. Archives of Agronomy and Soil Science **60**, 1291–1302.

Ramirez VE, Poppenberger B. 2020. Modes of Brassinosteroid Activity in Cold Stress Tolerance. Frontiers in plant science **11**, 583666.

Rativa AGS, Navarro BB, Gastmann R, *et al.* 2021. Decreased night temperature affects development and grain yield only in cold-susceptible rice (Oryza sativa) plants. Crop and Pasture Science, 782–788.

Rogowska A, Szakiel A. 2020. The role of sterols in plant response to abiotic stress. Phytochemistry Reviews **19**, 1525–1538.

Román-Figueroa C, Bravo L, Paneque M, Navia R, Cea M. 2021. Chemical products for crop protection against freezing stress: A review. Journal of Agronomy and Crop Science **207**, 391–403.

Roychoudhury A, Basu S, Sengupta DN. 2011. Amelioration of salinity stress by exogenously applied spermidine or spermine in three varieties of indica rice differing in their level of salt tolerance. Journal of Plant Physiology **168**, 317–328.

Ruelland E, Vaultier M-N, Zachowski A, Hurry VBT-A in BR. 2009. Cold Signalling and Cold Acclimation in Plants. Advances in Botnaical Research. Academic Press, 35–150.

Salvucci ME, Crafts-Brandner SJ. 2004. Relationship between the heat tolerance of photosynthesis and the thermal stability of rubisco activase in plants from contrasting thermal environments. Plant Physiology **134**, 1460–1470.

Sato Y, Masuta Y, Saito K, Murayama S, Ozawa K. 2011. Enhanced chilling tolerance at the booting stage in rice by transgenic overexpression of the ascorbate peroxidase gene, OsAPXa. Plant cell reports **30**, 399–406.

Scafaro AP, Atwell BJ, Muylaert S, Reusel B Van, Ruiz GA, Rie J Van, Gallé A. 2018. A thermotolerant variant of Rubisco activase from a wild relative improves growth and seed yield in rice under heat stress. Frontiers in Plant Science **9**, 1–11.

Scafaro AP, Bautsoens N, den Boer B, Van Rie J, Galle A. 2019. A conserved sequence from heat-adapted species improves Rubisco activase thermostability in wheat. Plant Physiology **181**, 43–54.

Scafaro AP, Gallé A, Van Rie J, Carmo-Silva E, Salvucci ME, Atwell BJ. 2016. Heat tolerance in a wild Oryza species is attributed to maintenance of Rubisco activation by a thermally stable Rubisco activase ortholog. New Phytologist **211**, 899–911.

Sharkey TD, Zhang R. 2010. High Temperature Effects on Electron and Proton Circuits of Photosynthesis. Journal of Integrative Plant Biology **52**, 712–722.

Sharma KD, Nayyar H. 2016. Regulatory Networks in Pollen Development under Cold Stress. Frontiers in Plant Science **7**, 402.

Sharma N, Reinke R, Sacks EJ. 2021. Comparison of Methods to Evaluate Rice (Oryza sativa) Germplasm for Tolerance to Low Temperature at the Seedling Stage. Agronomy **11**.

Shi J, Cao Y, Fan X, Li M, Wang Y, Ming F. 2012. A rice microsomal delta-12 fatty acid desaturase can enhance resistance to cold stress in yeast and Oryza sativa. Molecular Breeding **29**, 743–757.

Shi H, Ye T, Zhong B, Liu X, Chan Z. 2014. Comparative proteomic and metabolomic analyses reveal mechanisms of improved cold stress tolerance in bermudagrass (Cynodon dactylon (L.) Pers.) by exogenous calcium. Journal of Integrative Plant Biology **56**, 1064–1079.

Shivhare D, Mueller-Cajar O. 2017. In Vitro Characterization of Thermostable CAM Rubisco Activase Reveals a Rubisco Interacting Surface Loop . Plant Physiology **174**, 1505–1516.

Singh R, Brennan JP, Farrell T, Williams R, Lewin L, Mullen JBT-EA of ICT in R in A. 2005. Economic analysis of improving cold tolerance in rice in Australia.

Singh I, Kumar U, Singh SK, Gupta C, Singh M, Kushwaha SR. 2012. Physiological and biochemical effect of 24-epibrassinoslide on cold tolerance in maize seedlings. Physiology and Molecular Biology of Plants **18**, 229–236.

Soll J, Schultz G, Joyard J, Douce R, Block MA. 1985. Localization and synthesis of prenylquinones in isolated outer and inner envelope membranes from spinach chloroplasts. Archives of Biochemistry and Biophysics **238**, 290–299.

Subramanyam R, Jolley C, Thangaraj B, Nellaepalli S, Webber AN, Fromme P. 2010. Structural and functional changes of PSI-LHCI supercomplexes of Chlamydomonas reinhardtii cells grown under high salt conditions. Planta **231**, 913–922.

Sun M, Shen Y, Li H, Yang J, Cai X, Zheng G, Zhu Y, Jia B, Sun X. 2019. The multiple roles of OsmiR535 in modulating plant height, panicle branching and grain shape. Plant Science **283**, 60–69.

Sun M, Shen Y, Yang J, Cai X, Li H, Zhu Y, Jia B, Sun X. 2020. miR535 negatively regulates cold tolerance in rice. Molecular Breeding **40**, 14.

Takahashi D, Li B, Nakayama T, Kawamura Y, Uemura M. 2013. Plant plasma membrane proteomics for improving cold tolerance. Frontiers in plant science **4**, 90.

Teixeira SB, Pires SN, Ávila GE, Silva BEP, Schmitz VN, Deuner C, da Silva Armesto R, da Silva Moura D, Deuner S. 2021. Application of vigor indexes to evaluate the cold tolerance in rice seeds germination conditioned in plant extract. Scientific Reports **11**, 11038.

Thakur P, Kumar S, Malik JA, Berger JD, Nayyar H. 2010. Cold stress effects on reproductive development in grain crops: An overview. Environmental and Experimental Botany **67**, 429–443.

Vogt T. 2018. Unusual spermine-conjugated hydroxycinnamic acids on pollen: function and evolutionary advantage. Journal of Experimental Botany **69**, 5311–5315.

Wang W, Chen Q, Hussain S, Mei J, Dong H, Peng S, Huang J, Cui K, Nie L. 2016. Pre-sowing Seed Treatments in Direct-seeded Early Rice: Consequences for Emergence, Seedling Growth and Associated Metabolic Events under Chilling Stress. Scientific reports **6**, 19637.

Wang Q, Guan Y, Wu Y, Chen H, Chen F, Chu C. 2008*a*. Overexpression of a rice OsDREB1F gene increases salt, drought, and low temperature tolerance in both Arabidopsis and rice. Plant molecular biology **67**, 589–602.

Wang C, Ma XL, Hui Z, Wang W. 2008*b*. Glycine betaine improves thylakoid membrane function of tobacco leaves under low-temperature stress. Photosynthetica **46**, 400–409.

Wang S, Sun X, Hoshino Y, Yu Y, Jia B, Sun Z, Sun M, Duan X, Zhu Y. 2014. MicroRNA319 positively regulates cold tolerance by targeting OsPCF6 and OsTCP21 in rice (Oryza sativa L.). PloS one **9**, e91357.

Wang D, Wang Y, Long W, *et al.* 2017. SGD1, a key enzyme in tocopherol biosynthesis, is essential for plant development and cold tolerance in rice. Plant science : an international journal of experimental plant biology **260**, 90–100.

Wang X, Yu C, Liu Y, *et al.* 2019. GmFAD3A, A ω-3 Fatty Acid Desaturase Gene, Enhances Cold Tolerance and Seed Germination Rate under Low Temperature in Rice. International journal of molecular sciences **20**.

Wen Y, Qin D, Leng B, Zhu Y, Cao K. 2018. The physiological cold tolerance of warm-climate plants is correlated with their latitudinal range limit. Biology Letters **14**, 20180277.

Xiaochuang C, Chu Z, Lianfeng Z, Junhua Z, Hussain S, Lianghuan W, Qianyu J. 2017. Glycine increases cold tolerance in rice via the regulation of N uptake, physiological characteristics, and photosynthesis. Plant physiology and biochemistry : PPB **112**, 251–260.

Xu Y, Wang R, Wang Y, Zhang L, Yao S. 2020. A point mutation in LTT1 enhances cold tolerance at the booting stage in rice. Plant, cell & environment **43**, 992–1007.

Xu L-M, Zhou L, Zeng Y-W, Wang F-M, Zhang H-L, Shen S-Q, Li Z-C. 2008. Identification and mapping of quantitative trait loci for cold tolerance at the booting stage in a japonica rice near-isogenic line. Plant Science **174**, 340–347.

Yamamori K, Ogasawara K, Ishiguro S, Koide Y, Takamure I, Fujino K, Sato Y, Kishima Y. 2021. Revision of the relationship between anther morphology and pollen sterility by cold stress at the booting stage in rice. Annals of Botany **128**, 559–575.

Yang LM, Liu HL, Lei L, Zhao HW, Wang JG, Li N, Sun J, Zheng HL, Zou DT. 2017. Identification of QTLs controlling low-temperature germinability and cold tolerance at the seedling stage in rice (Oryza Sativa L.). Euphytica **214**, 13.

Yang Y, Saand MA, Abdelaal WB, Zhang J, Wu Y, Li J, Fan H, Wang F. 2020. iTRAQ-based comparative proteomic analysis of two coconut varieties reveals aromatic coconut cold-sensitive in response to low temperature. Journal of Proteomics **220**, 103766.

Yarra R, Wei W. 2021. The NAC-type transcription factor GmNAC20 improves cold, salinity tolerance, and lateral root formation in transgenic rice plants. Functional & Integrative Genomics **21**, 473–487.

Ye C, Fukai S, Godwin I, Reinke R, Snell P, Schiller J, Basnayake J. 2009. Cold tolerance in rice varieties at different growth stages. Crop and Pasture Science **60**, 328–338.

Ye K, Li H, Ding Y, Shi Y, Song C, Gong Z, Yang S. 2019. BRASSINOSTEROID-INSENSITIVE2 Negatively Regulates the Stability of Transcription Factor ICE1 in Response to Cold Stress in Arabidopsis. The Plant Cell **31**, 2682–2696.

Yoshida S. 1981. Climatic environment and its influence. In: Brady N, ed. Fundamentals of rice crop science. Los Banos, Laguna, Phillippines: International RIce Research Institute, 65–110.

Zhang Q, Chen Q, Wang S, Hong Y, Wang Z. 2014. Rice and cold stress: methods for its evaluation and summary of cold tolerance-related quantitative trait loci. Rice (New York, N.Y.) **7**, 24.

Zhang H, Dong J, Zhao X, *et al.* 2019. Research Progress in Membrane Lipid Metabolism and Molecular Mechanism in Peanut Cold Tolerance . Frontiers in Plant Science **10**, 838.

Zhang Z, Li J, Pan Y, *et al.* 2017. Natural variation in CTB4a enhances rice adaptation to cold habitats. Nature communications **8**, 14788.

Zhang Y, Liu K, Zhu X, Wu Y, Zhang S, Chen H, Ling J, Wang Y, Fang X. 2018. Rice tocopherol deficiency 1 encodes a homogentisate phytyltransferase essential for tocopherol biosynthesis and plant development in rice. Plant cell reports **37**, 775–787.

Zhao M, Yuan L, Wang J, Xie S, Zheng Y, Nie L, Zhu S, Hou J, Chen G, Wang C. 2019. Transcriptome analysis reveals a positive effect of brassinosteroids on the photosynthetic capacity of wucai under low temperature. BMC Genomics **20**, 810.

Zhao J, Zhang S, Dong J, Yang T, Mao X, Liu Q, Wang X, Liu B. 2017. A novel functional gene associated with cold tolerance at the seedling stage in rice. Plant biotechnology journal **15**, 1141–1148.

Figures

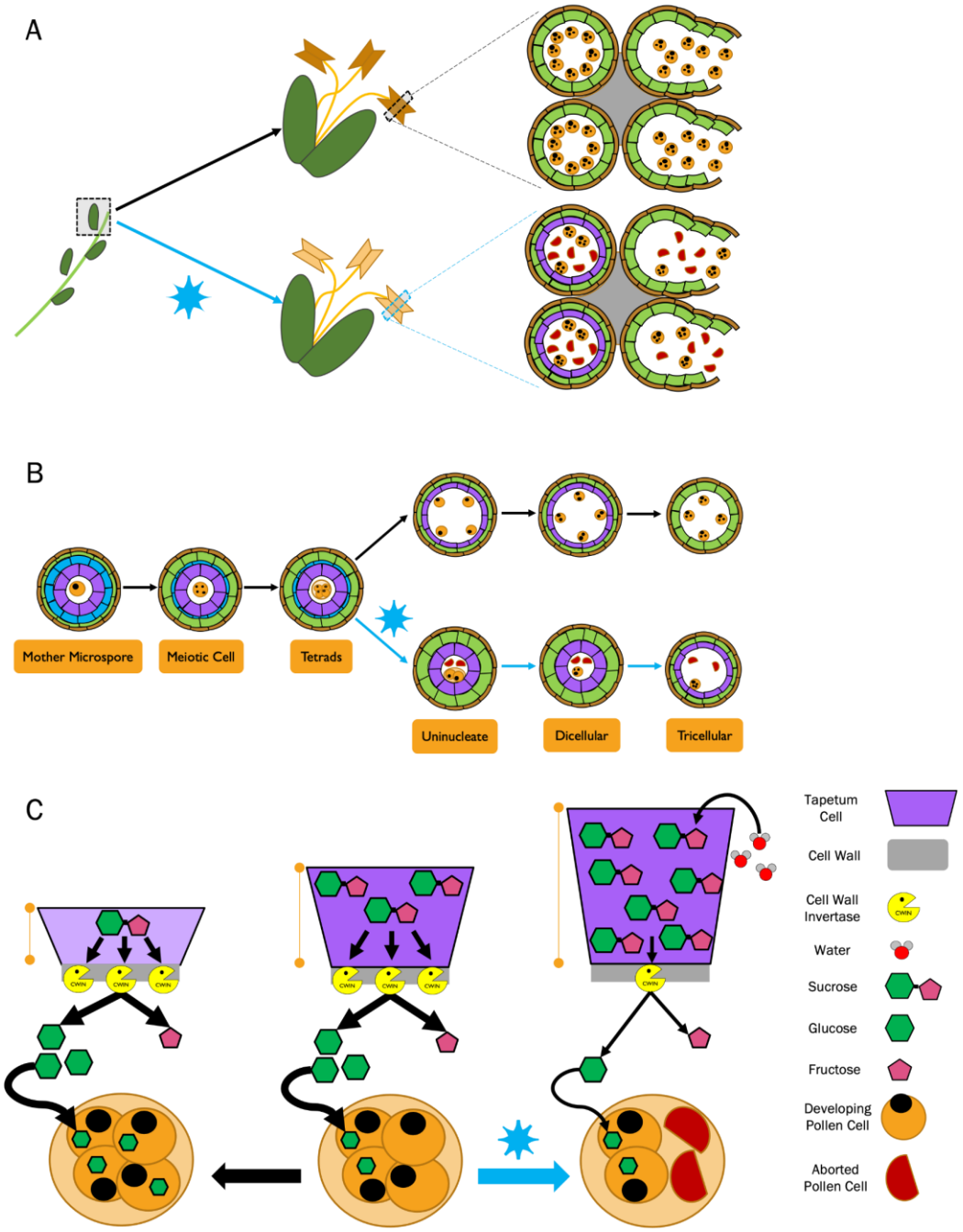
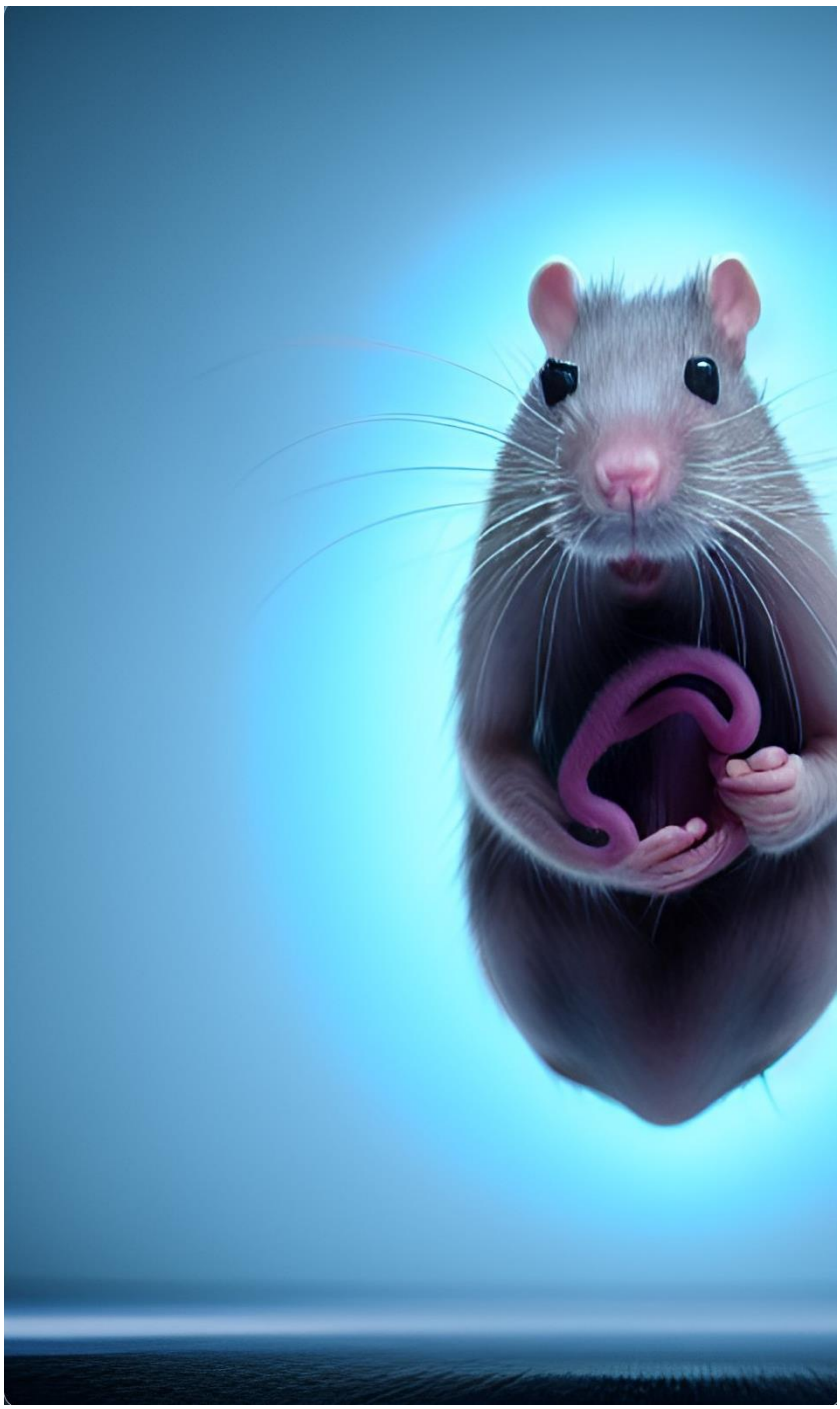


Figure 1. (A) Rice is most sensitive to low temperatures between flag leaf formation and booting because it is when pollen is forming by meiosis inside anthers within the immature flowers; a process called microsporogenesis. (B) Cold stress that coincides with microsporogenesis impairs the development of viable pollen by hindering the function of the tapetum (purple cell layer), causing spikelet sterility and ultimately yield loss. (C) In a complicated signalling pathway, CWIN is switched off by chilling, blocking the splitting of sucrose and starving sugar supply to pollen. The tapetum swells due to osmosis and cannot function. Starved pollen cells develop strange morphologies and are frequently aborted.

Appendix IV

Front Page of: Epigenetic Mechanisms Involved in Intrauterine Growth Restriction and Aberrant Kidney Development and Function



Wombo Prompt:

“Rat epigenetic inheritance uterus”

Wombo Theme:

“HDR”

CAMBRIDGE UNIVERSITY PRESS

Epigenetic mechanisms involved in intrauterine growth restriction and aberrant kidney development and function

Thu N. A. Doan[1,2], Jessica F. Briffa[3], **Aaron L. Phillips**[1], Shalem Y. Leemaqz[2,4,5], Rachel A. Burton[1], Tania Romano[6], Mary E. Wlodek[3] and Tina Bianco-Miotto[1,2,*]

[1]School of Agriculture, Food and Wine & Waite Research Institute, University of Adelaide, Adelaide, South Australia, Australia; [2]Robinson Research Institute, University of Adelaide, South Australia, Australia; [3]Department of Physiology, The University of Melbourne, Parkville, Victoria, Australia; [4]South Australian Health & Medical Research Institute, SAHMRI Women and Kids, Adelaide, Australia; [5]College of Medicine and Public Health, Flinders University, Bedford Park, SA, Australia and [6]Department of Physiology, Anatomy and Microbiology, La Trobe University, Bundoora, Victoria, Australia

Abstract

Intrauterine growth restriction (IUGR) due to uteroplacental insufficiency results in a placenta that is unable to provide adequate nutrients and oxygen to the fetus. These growth-restricted babies have an increased risk of hypertension and chronic kidney disease later in life. In rats, both male and female growth-restricted offspring have nephron deficits but only males develop kidney dysfunction and high blood pressure. In addition, there is transgenerational transmission of nephron deficits and hypertension risk. Therefore, epigenetic mechanisms may explain the sex-specific programming and multigenerational transmission of IUGR-related phenotypes. Expression of DNA methyltransferases (Dnmt1and Dnmt3a) and imprinted genes (Peg3, Snrpn, Kcnq1, and Cdkn1c) were investigated in kidney tissues of sham and IUGR rats in F1 (embryonic day 20 (E20) and postnatal day 1 (PN1)) and F2 (6 and 12 months of age, paternal and maternal lines) generations (n = 6–13/group). In comparison to sham offspring, F1 IUGR rats had a 19% decrease in Dnmt3a expression at E20 ($P < 0.05$), with decreased Cdkn1c (19%, $P < 0.05$) and increased Kcnq1 (1.6-fold, $P < 0.01$) at PN1. There was a sexspecific difference in Cdkn1c and Snrpn expression at E20, with 29% and 34% higher expression in IUGR males compared to females, respectively ($P < 0.05$). Peg3 sex-specific expression was lost in the F2 IUGR offspring, only in the maternal line. These findings suggest that epigenetic mechanisms may be altered in renal embryonic and/or fetal development in growth-restricted offspring, which could alter kidney function, predisposing these offspring to kidney disease later in life.

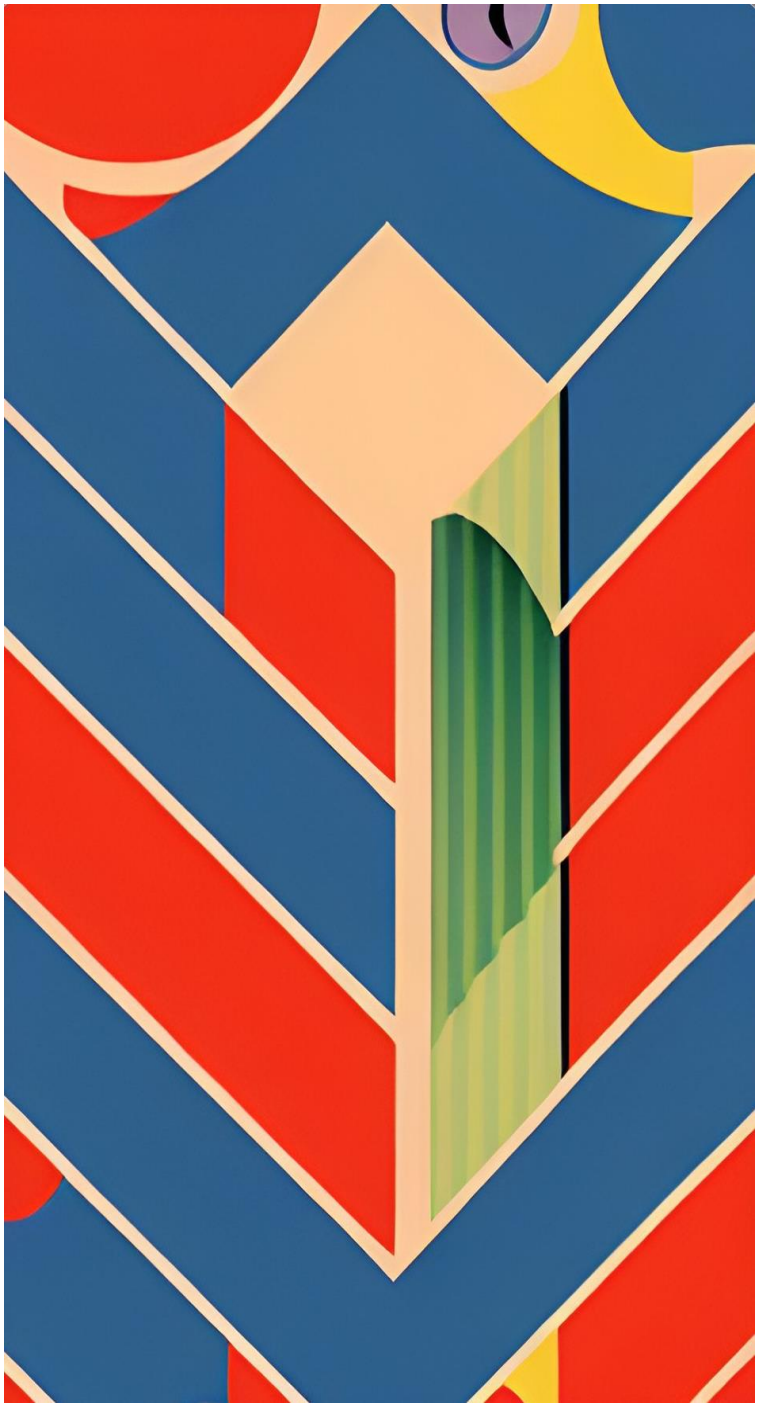
Introduction

Intrauterine growth restriction (IUGR) refers to the restricted growth of a fetus (<10th percentile) compared to its expected capability when examined at a specific gestational age and results in a baby that has low birth weight (LBW).[1] Globally, the prevalence of a baby being born small for gestational age is approximately 20% of all births.[2] In Australia, the incidence of IUGR is approximately 10% of all singleton pregnancies.[3] This high frequency of IUGR is a significant concern, as these individuals are at an increased risk of developing chronic diseases later in life.[4] Indeed, epidemiological studies have shown that LBW babies have a 1.8 times increased risk of developing hypertension as adults.[5,6] Additionally, human IUGR fetuses have reduced nephron number[1,7] and increased expression of pre-apoptotic proteins[7] in their kidney tissues. Furthermore, LBW individuals that have died from cardiovascular disease have a lower number of glomeruli, the key structural unit of the nephron,[8,9] suggesting that an abnormal functioning kidney contributes to the hypertension. In order to reduce the health burden of these diseases, we need to know more about the mechanisms by which IUGR programs chronic disease risk later in life, specifically kidney and heart disease.

One of the most common causes of IUGR is uteroplacental insufficiency (UPI), a pregnancy complication in which the placenta functions poorly causing a limited delivery of nutrients and oxygen to the fetus.[10] There are two methods used to induce UPI in animal models that reflect IUGR metabolic characteristics in humans. The first involves clamping the abdominal aorta and the branches of uterine arteries,[11] while the second involves ligating both the uterine artery and vein of pregnant animals such as rabbits,[12] guinea pigs,[13] or rats.[10,14–18] As expected, these offspring develop

Appendix V

Conference Presentations and Posters



Wombo Prompt

“Presenting research results”

Wombo Theme

“Throwback”

I have presented my research a number of times throughout my candidature. These presentations are listed in chronological order below.

- 2nd June 2019

  Centre of Excellence for Plant Energy Biology Annual Forum plenary talk

  “Defining the role of OVP3 in the growth and yield of rice”

- 23rd August 2019

  PhD Milestone: Core Component of the Structured Program (CCSP)

  “Progress on defining the role of OVP3 in the growth and yield of rice”

- 25th March 2020

  PhD Milestone: Major Review

  “Growth analyses of OVP mutants grown under various stress conditions”

- 30th September 2020

  University of Adelaide Postgraduate Symposium

  “Identifying genomic variation in wild rice species related to salt tolerance”

- 24th August 2021

  Burton Lab Meeting

  “Exploring thermotolerance in an extremophile wild Australian rice”

- 23rd November 2021

  Australian Society of Plant Scientists (ASPS) Adelaide Presentation & Poster

  “The reference genome of *Oryza australiensis*, a wild rice relative with multi-stress tolerance”

- 10th May 2022

  16th Annual Australasian Plant Breeding Conference

  “Wild rice for a changing climate”

The reference genome of *Oryza australiensis*, a wild rice relative with multi-stress tolerance

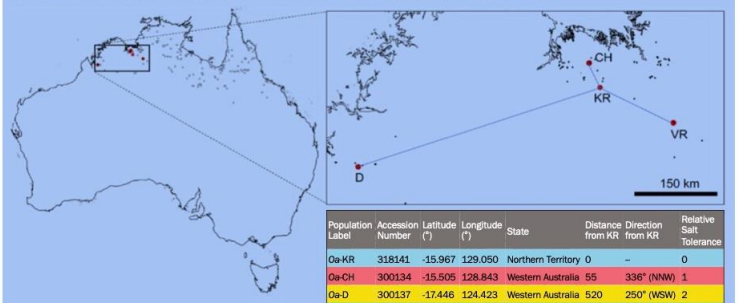
Aaron L. Phillips[1,4,5], Nathan S. Watson-Haigh[2,3], Ashley W. Jones[4], Justin O. Borevitz[4], Rachel A. Burton[1,4], and Brian J. Atwell[5]

aaron.phillips@adelaide.edu.au

University of Adelaide, Department of Food Science, Urrbrae, Australia[1]
South Australian Genomics Centre, Adelaide, Australia[2]
Australian Genome Research Facility, Victoria, Australia[3]
ARC CoE in Plant Energy Biology, University of Adelaide & Australian National University, Canberra, Australia[4]
Macquarie University, Department of Biological Sciences, Sydney, Australia[5]

BACKGROUND & AIM *Oryza australiensis* is a wild relative of rice that is tolerant to several biotic and abiotic stresses. For example, it is tolerant to prolonged extreme heat (up to 45°C) and it demonstrates population-specific salt (drought?) tolerance. The International *Oryza* Map Alignment Project identified the genome of *O. australiensis*, the sole *Oryza* EE genome, as a genome of interest. However, sequencing efforts since 2003 have failed to assemble a highly contiguous genome. Here we present the first high-quality genome for *O. australiensis* from the Northern Territory of Australia.

1. *Oryza australiensis* accessions with differing levels of salt tolerance were sequenced. The four accessions collected from four locations in northern Australia regularly experience multiple environmental stresses including extreme daytime temperature, fluctuating water availability, and saline soils. The accession from Keep River (KR) in the Northern Territory was sequenced to a higher coverage (51×) to generate the reference genome. The remaining three accessions were used to gauge the genomic diversity of *O. australiensis*.

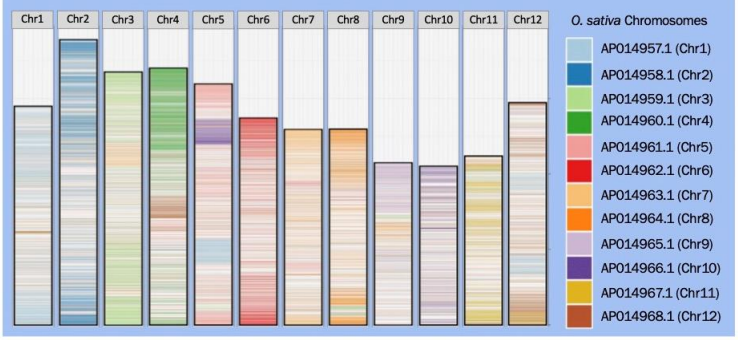


Population Label	Accession Number	Latitude (°)	Longitude (°)	State	Distance from KR	Direction from KR	Relative Salt Tolerance
Oa-KR	318141	-15.967	129.050	Northern Territory	0	–	0
Oa-CH	300134	-15.505	128.843	Western Australia	55	336° (NNW)	1
Oa-D	300137	-17.446	124.423	Western Australia	520	250° (WSW)	2
Oa-VR	300131	-16.625	130.450	Northern Territory	116	116° (ESE)	3

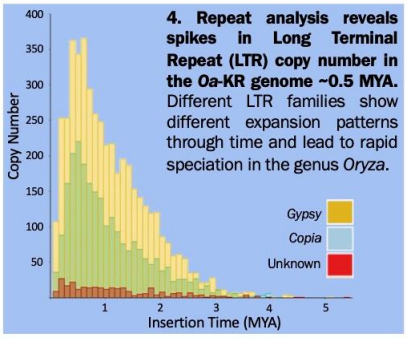
2. A high-quality scaffolded reference genome was assembled for *Oryza australiensis*. High molecular weight DNA was extracted from one plant per accession and sequenced using minION and Illumina flow cells. The genome was assembled using *Canu*, polished using *HyPo*, and ordered into pseudomolecules using *RagTag*. Gene annotation by Omicsbox used Iso-Seq transcripts from five *Oa*-KR tissues and domestic rice proteins as evidence.

Stage	Feature	Value
Long reads	Coverage	51
	Expected Genome Size (Mbp)	965
Genome Scaffolds	Scaffolded Genome Size (Mbp)	998.8
	Placed Contigs	1053
	Unplaced Contigs	903
	Unplaced Contigs Removed (Contamination)	2
	Total Length of Placed Contigs (Mbp)	849.1
	Total Length of Unplaced Contigs (Mbp)	149.7
	Total Length of Contamination (Mbp)	0.552
	Gap (Mbp)	3.86
	BUSCO Score (100% = 4896)	95.6
	Whole-Genome LAI Score	14.9
LTR_Retriever (scaffolds)	Repeat Content (%)	61.6
OmicsBox (scaffolds and contigs)	Repeat Content in Scaffolds (%)	78.2
	Repeat Content in Unplaced Contigs (Mbp)	135
Annotation	No. Genes in Scaffolds	53615
	No. Transcripts in Scaffolds	62101
	No. Genes in Unplaced Contigs	2330
	No. Transcripts in Unplaced Contigs	2330

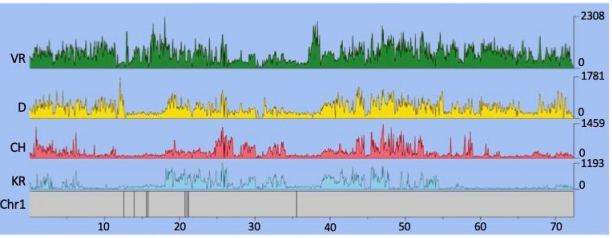
3. Alignment of the *Oa*-KR genome with the reference *O. sativa* Nipponbare genome reveals large-scale chromosomal rearrangements. Colours correspond to homologous *O. sativa* sequences. The *Oa*-KR pseudomolecules are chimeric and have high sequence dissimilarity (white space), which may explain why crossing *O. australiensis* & *O. sativa* is difficult.



4. Repeat analysis reveals spikes in Long Terminal Repeat (LTR) copy number in the *Oa*-KR genome ~0.5 MYA. Different LTR families show different expansion patterns through time and lead to rapid speciation in the genus *Oryza*.



5. Single Nucleotide Polymorphism (SNP) densities are related to degree of salt tolerance, not physical distance between populations. Illumina short reads from the four accessions were mapped to the Oa-KR genome and SNPs were called. SNP densities in 70 Kb windows were drawn for Chr1 as this chromosome contains the important SalTol QTL in *O. sativa* that confers salt tolerance. VR and CH are 215 and 499 km away from the site of collection for *Oa*-KR, respectively, however it is the VR sample, which has higher salt tolerance than CH, that has the higher abundance of SNPs when compared to the *Oa*-KR assembly.



CONCLUSIONS AND FUTURE DIRECTIONS The high-quality reference genome of *O. australiensis* presented here will be used to probe differences in salt tolerance that occur in geographically isolated populations. We using it to uncover heat and drought tolerance mechanisms in this unique wild rice relative.

INTENTIONALLY LEFT BLANK