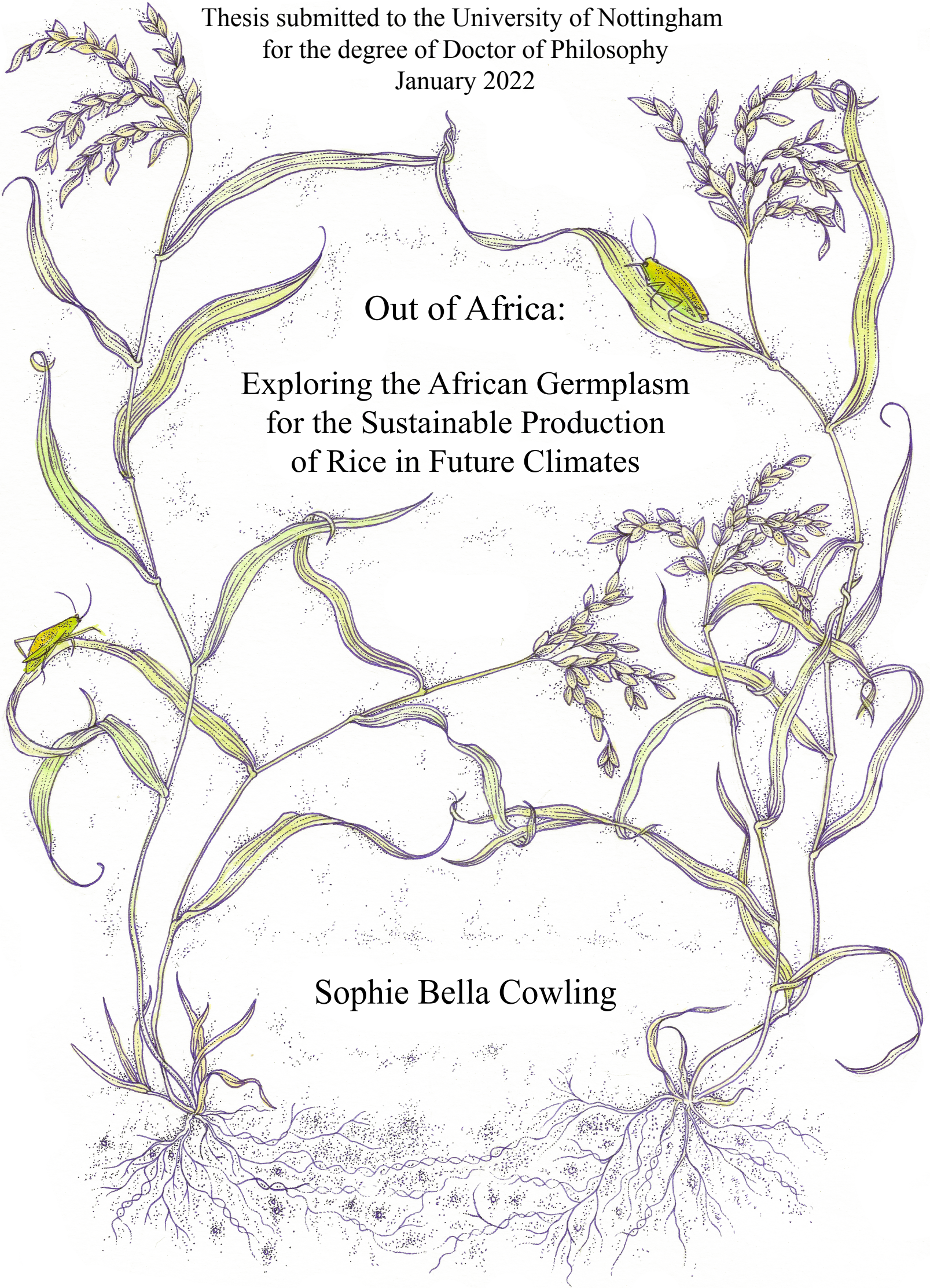
Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy
January 2022



Out of Africa:

Exploring the African Germplasm for the Sustainable Production of Rice in Future Climates

Sophie Bella Cowling

ABSTRACT

Asian rice (*Oryza sativa*) is a staple food for billions of people worldwide. There has been increasing focus on the introgression of diverse material to develop productive rice cultivars that are resilient to a changing climate. African rice (*Oryza glaberrima*) evolved independently to Asian rice and demonstrates exceptional abiotic stress resilience. *O. glaberrima* has been recognised as a potential source for crop improvement for over two decades but the heritable variation in the species remains relatively uncharacterised. The research described here uses a new genomic resource of 155 re-sequenced accessions to explore the diversity *O. glaberrima* has to offer.

O. glaberrima accessions were grown to late tillering stage in an agronomy glasshouse and measured for traits that contribute to water use efficiency and productivity. Photosynthetic traits were measured by gas exchange and chlorophyll fluorescence, along with root and shoot biomass, stomatal density and leaf area. Photosynthetic steady state and kinetic responses were modelled. Using this phenotypic data, a genome wide association study (GWAS) was completed using four different algorithms to identify significant trait related candidate genes and genomic regions. Multi-variate analyses were used to explore the ecological and environmental factors that contribute to trait and genetic adaption in the *O. glaberrima* collection used here.

This study identified broad heritable variation and candidate genes in a range of useful morphological, steady-state and dynamic photosynthetic traits, including genes known for mitigating drought and heat stress, alongside regulators of key phytohormones. This thesis highlights the importance of *O. glaberrima* as a source of heritable variation. The information compiled here provides a solid base for future elucidation of physiological processes and the functional validation of candidate genes to support future crop improvement efforts.

DECLARATION

Unless otherwise acknowledged, the work presented within this thesis is my own. No part has been submitted for another degree at the University of Nottingham or any other institution of learning.

ACKNOWLEDGEMENTS

Professional

I would like to thank my funding body, the BBSRC, for providing me with the financial support to pursue the research detailed here and develop professional opportunities.

Thank you to my supervisors Prof. Sean Mayes and Prof. Erik Murchie for their support throughout the past four years and during the recent development of manuscripts for publication. Most of all, I would like to thank them for their patience. I would also like to thank Dr. John Ferguson, Dr. Lorna McAusland and Dr. Philippe Cubry for their technical support during my PhD. Finally, I would like to commemorate my colleague and fellow *O. glaberrima* researcher Dr. Pracha Treeintong who sadly died in 2020. While I only knew Pracha as a colleague, he was unerringly kind and a dedicated scientist.

Personal

Bizarrely, I'd first like to acknowledge my two dogs, their constant presence, energy and cuddles kept me calm and anchored. Except for maybe the months spent frequenting super vet clinics and five-figure bill. My partner Alex Jones has been my rock throughout the past year. His unwavering belief in me, even during pain and turmoil, has given me the confidence to strive to be my best and follow my passion. A huge thank you to all my amazing friends, particularly Cameron Grundy, Giorgia Tibaldero, Matthew Hartley, Lauren Baker, Fern Ritchie and Beth Penrose. A special thank you to my cousin Bryony Kemp for designing my beautiful thesis cover page.

Finally, I would like to say thank you to my parents. They always told me I could do anything, and I believed them. Despite my 6-year deviation as a Police Officer, I finally became a biologist! Huzzah! My love of plants originated through my mum, who died when I was 15 years old. I am forever grateful for my memories of her and the love she gave me, I know she would be proud. Dad, you're one of the few people who might read this tome, thank you for being you, teaching me to question everything (pretty much anything, including authority), think for myself and stand up for what I believe in.

ABBREVIATIONS

Abbreviation	Definition	Units
A	Net CO_2 assimilation	μmol m^{-2} s^{-1}
Abaxial SD	SD on the lower leaf side	mm^2
Adaxial SD	SD on the upper leaf side	mm^2
ANOVA	Analysis of variance	
API	Application programming interface	
BLUPs	Best linear unbiased prediction	
DIADE	Diversité Adaptation Developpement des plantes	
ETR	Electron transport rate of PSII.	μmol electrons m^{-2} s^{-1}
FOV	Field of view	
g_s	Stomatal conductance to water vapour	mmol m^{-2} s^{-1}
gs_{max}	Maximum *gs* achieved under 1500 μmol m^{-2} s^{-1} PPFD.	mmol m^{-2} s^{-1}
GWAS	Genome wide association study	
H^2	Broad sense heritability	
H-clustering	Hierarchical clustering	
HCPC	Hierarchical clustering on principal components	
iWUE	Intrinsic water use efficiency.	μmol mol^{-1}
IQR	Inter-quartile range	
IRGA	Infra-red gas analyser	
LED	Light emitting diode	
NPQ	Non-photochemical quenching.	
PAR	Photosynthetically active radiation	
PCA	Principal component analysis	
PC	Principal component	
PGV	Percentage genetic variation	

ΦPSII	Relative quantum yield of photochemical energy conversion at steady state *A*.	
PPFD	Photosynthetically active photon flux density	μmol m^{-2} s^{-1}
qP	Photochemical quenching	
R-CNN	Regional convolutional neural networks	
RICE	Rice interspecies comparison & evolution team at University of Montpellier.	
sd	Standard deviation	
SD	Stomatal density	mm^2
SE	Standard error	
Trmmol	Transpiration rate.	mmol H_2O m^{-2} s^{-1}
VPD	Vapour pressure deficit between leaf and air.	kPa
X trait $_{max}$	Maximum of X trait achieved under 1500 mmol m^{-2} s^{-1} PPFD. Examples include A_{max} and gs_{max}.	Unit of X trait

Dynamic modelling definitions for chapters three and four:

X trait $_i$		
X trait $_r$		
A_i | gs_i | NPQ_i	Induction response to a change of PPFD from 0 to 1500 mmol m^{-2} s^{-1} (difference between)	A_i = mmol m^{-2} s^{-1} gs_i = mmol m^{-2} s^{-1}
A_r | gs_r | NPQ_r	Relaxation response to a change of PPFD from 1500 to 100 mmol m^{-2} s^{-1}	A_r = mmol m^{-2} s^{-1} gs_r = mmol m^{-2} s^{-1}
$A_{i\ slope}$ | $gs_{i\ slope}$ | $NPQ_{i\ slope}$	Model estimated slope gradient for the induction ($_i$) and relaxation ($_r$) response curves.	

$A_{r\ slope}$ | $gs_{r\ slope}$ |$NPQ_{r\ slope}$		
$A_{i\ min}$ | $gs_{i\ min}$ | $NPQ_{i\ min}$ $A_{r\ min}$ | $gs_{r\ min}$ | $NPQ_{r\ min}$	Model estimated minimum value at the beginning of an induction ($_i$) curve and at the end of a relaxation ($_r$) curve.	A_i | A_r = mmol m^{-2} s^{-1} gs_i | gs_r = mmol m^{-2} s^{-1}
$gs_{i\ max}$ | $A_{i\ max}$ | $NPQ_{i\ max}$ $gs_{r\ max}$ | $A_{r\ max}$ | $NPQ_{r\ max}$	Model estimated maximum value at the top of an induction ($_i$) or relaxation ($_r$) curve.	A_i | A_r = mmol m^{-2} s^{-1} gs_i | gs_r = mmol m^{-2} s^{-1}
$A_{i\ 10}$ | $gs_{i\ 10}$ | $NPQ_{i\ 10}$ $A_{r\ 10}$ | $gs_{r\ 10}$ | $NPQ_{r\ 10}$	Model estimated time taken to reach 10% of the maximum value achieved on the induction ($_i$) or relaxation ($_r$) response curve.	Time (secs)
$A_{i\ 50}$ | $gs_{i\ 50}$ | $NPQ_{i\ 50}$ $A_{r\ 50}$ | $gs_{r\ 50}$ | $NPQ_{r\ 50}$	Model estimated time taken to reach 50% of the maximum value achieved on the induction ($_i$) or relaxation ($_r$) response curve.	Time (secs)
$A_{i\ 90}$ | $gs_{i\ 90}$ | $NPQ_{i\ 90}$ $A_{r\ 90}$ | $gs_{r\ 90}$ | $NPQ_{r\ 90}$	Model estimated time taken to reach 90% of the maximum value achieved on the induction ($_i$) or relaxation ($_r$) response curve.	Time (secs)
$A_{i\ rate}$ | $gs_{i\ rate}$ | $NPQ_{i\ rate}$	Unit per second taken to induce from 0 seconds to the time taken to achieve 90% of the upper limit induction curve.	Trait unit per second
$A_{r\ rate}$ | $gs_{r\ rate}$ | $NPQ_{r\ rate}$	Unit per second taken to relax from 900 seconds to the time taken to achieve 90% of the lower limit relaxation curve.	Trait unit per second

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTARY TABLES

LIST OF SUPPLEMENTARY FIGURES

LIST OF SUPPLEMENTARY DATA FILES

Data File S3.2: Descriptive statistics of all data from *Oryza glaberrima* accessions.

Data File S3.3: Descriptive statistics of all data from *Oryza sativa* accession IR64.

Data File S3.4: Box and distribution plots of linear mixed effects model adjusted means for *Oryza glaberrima* and *Oryza sativa*

Data File S3.5: The Python code for extracting a defined percentage of a curve.

Data File S3.6: This shows the Principal Component Analysis of key traits.

Data File S3.7: This shows the Hierarchical Cluster analysis of key traits.

Data File S3.8: This shows the Principal Component Analysis, climate data.

Data File S3.9: This shows the Hierarchical Cluster analysis, climate data .

Data File S3.10: Correlation table of phenotypic traits form *Oryza sativa* and *Oryza glaberrima*

Data File S3.11: List of accessions in each hierarchical cluster clade.

Data File S3.12: Correlation table of Climate and Phenotypic data, 105 accessions.

CHAPTER ONE

Thesis introduction and research rationale

1.1 Introduction

We are living in an era that is widely accepted as a geological epoch and often referred to as the Anthropocene (Lewis & Maslin, 2015; Ellis et al., 2016). Human induced climate change is progressing at an alarming rate, the average global temperature has risen by 1.4^{o}C since the pre-industrial era and is estimated to continue increasing by 0.2^{o}C per decade (Shukla et al., 2019; Yerlikaya et al., 2020). The world-changing consequences are causing dramatic shifts in all natural systems, with species unable to adapt quickly enough to mitigate deleterious effects (Panetta et al., 2018; Capblancq et al., 2020; Román-Palacios & Wiens, 2020). This extends to human managed crop species, which have been selected by farmers for maximum yield gains but not resilience to an increasingly warm and unpredictable climate (Kang et al., 2009; Gao, 2021). The ability to meet global food demands in response to a burgeoning population and a climate crisis is a substantial challenge. Rising atmospheric CO_2 concentration has been stimulating photosynthesis and productivity through increased Rubisco carboxylation efficiency and the reduction of stomatal conductance, which may improve leaf level water use efficiency. However, the magnitude and longevity of such 'CO_2 fertilisation' effects are unclear (Ainsworth & Long, 2005). A recent study casts doubt on how long this effect may continue and suggests that the major crop producing regions will experience the negative effects of climate change sooner than anticipated (Jägermeyr et al., 2021). It is estimated that we could experience up to 37% yield losses by the end of the century, through temperature increases, drought events, climate change associated desertification and extreme weather events (IPCC, 2014; Zhao). These yield losses are anticipated to occur in the regions which supply most of the worlds key cereal crops; maize, wheat, soy and rice (Black et al., 2008; Challinor et al., 2014). Therefore, diverting plant research efforts to the development of climate resilient crop species is imperative to secure future crop productivity and food security in the face of the climate crisis.

1.2 Abiotic stress limitations on rice growth and productivity

Oryza sativa (Asian rice) is a dietary staple to approximately 50% of the global population (Seck et al., 2012; Muthayya et al., 2014). Billions of people depend upon reliable *O. sativa* productivity as their primary source of nutrients, calories and income, particularly in

developing countries. The increasing frequency of drought and natural disaster events have the propensity to cause widespread famine and poverty due to rice yield losses. For example, the 2015-2016 drought event in Thailand, which caused $1.7 billion in economic losses and affected 9.56 million people. Catastrophic events such as this are occurring more regularly due to the effects of climate change (Kang et al., 2021). In conjunction with this, the global rising temperatures are already negatively affecting rice yields (Iizumi et al., 2018; Ray et al., 2019), which are predicted to continue to decline, even under conservative temperature increases (Fig.1.1) (Zhao et al., 2017). While there is a degree of uncertainty between the modelling estimates, the overall trend is consistent and shows future yield reductions (Parry et al., 2005; Zhao et al., 2017; Wang et al., 2020). These projections are in direct conflict with the substantial yield increases that are required to feed an estimated global population of 9.7 billion people by 2050 (Charles et al., 2010). Therefore, it is essential that research is directed into developing abiotic stress tolerant cultivars to mitigate against significant yield losses in the future.

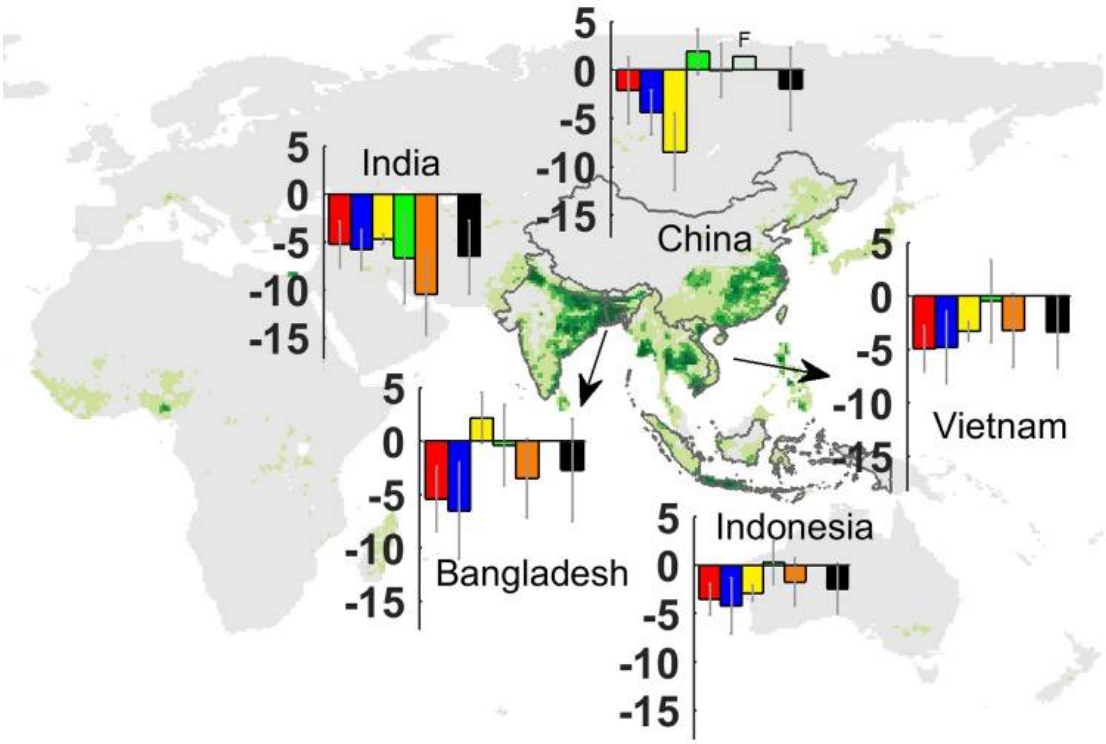


Figure 1.1: Estimates of rice yield changes in response to 1ºC global warming. The five major rice producing countries are shown, accompanied by a plot showing multi-method analyses estimating yield changes. The study uses field warming experiments, statistical regression, global grid-based and local point-based models to generate estimates. Each bar shows the estimate for each model used in the study, where; red = Grid-Sim, blue = Point-Sim, yellow = Point-Obs, green = Regress A, orange =

Regres C, grey = Regres D-K, black = average). The y-axis shows the percentage change for a 1°C temperature increase and error bars show the 95% confidence interval. Image is adapted from Zhao et al. (2017).

Abiotic stress is an umbrella term that is used to describe environmental pressures that may constrain plant growth and productivity, but these individual stressors result in differing negative consequences and physiological responses. The main abiotic stressors are extreme temperature, drought, salinity and flooding; a summary of their effects is discussed below.

- Extreme temperatures

  Increasing levels of atmospheric CO_2 concentrations are responsible for global warming and more frequent heat waves. The rice producing regions of the world are likely to experience extreme temperatures in the form of heat stress, rather than cold (Ramírez & Maiti, 2016). Rice can tolerate relatively high temperatures, with an optimal cultivation temperature range of 25-35°C. Outside of this range, rice growth and yields are substantially reduced, particularly if temperature increases are accompanied by drought or high humidity (Bailey-Serres et al., 2019). The sensitivity to extreme temperature varies throughout the plant growth cycle and development. For example, rice plants are particularly sensitive during the reproductive phase, when heat stress can reduce male fertility, seed quality and lead to yield reduction (Bita & Gerats, 2013). The consequences of exposure to moderate heat stress includes the inhibition of protein synthesis, enzyme inactivation in the mitochondria and chloroplasts, protein denaturation and increased cell membrane fluidity and leakage (Wahid et al., 2007). This results in the production of reactive oxygen species (ROS; $O_2^{\cdot -}$, H_2O_2, $OH^{\cdot}$, $^1O_2$), toxic compounds and the reduction of photosynthesis (Fig. 1.2).

- Drought

  Drought can be defined as a period with below average rainfall, or high evaporation levels due to heat. It affects over a third of the cultivated land area and 4.3×10^7 square hectometres (hm^2) in Asia alone (Singh et al., 2016). Among all the abiotic factors experienced by plants, drought places the greatest limitation on rice yields, which will only worsen under the climate crisis (Pandey & Shukla, 2015; Zhao et al., 2016). Plants experience drought stress through a lack of water supply to the roots or an

excessive transpiration rate. Akin to heat stress, drought stress severely impairs growth and productivity due to the various effects on processes essential to photosynthesis. Stomata close to limit water loss, a lack of stomatal conductance limits carbon assimilation, reduces mesophyll conductance, increases ROS and damages photosynthetic apparatus, resulting in the reduction of photosystem II (PSII) and Rubisco activity (Pandey & Shukla, 2015; Panda et al., 2021). These consequences of drought are associated to a slowing of processes essential to photosynthesis and cell expansion, to help limit water use.

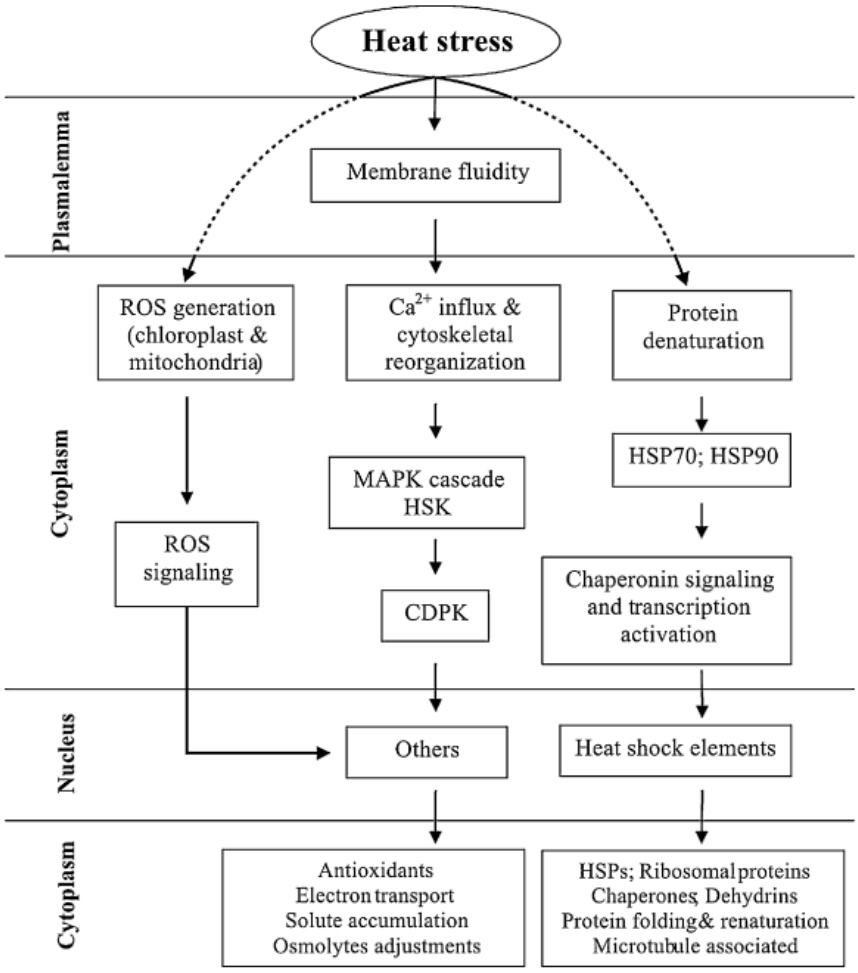


Figure 1.2: Plant heat-shock tolerance mechanisms, showing the key stress-related proteins and their interactions. Reactive oxygen species (ROS), mitogen activated protein kinases (MAPK), histidine kinase (HSK), calcium dependant protein kinase (CDPK) and heat shock protein (HSP). Image taken from Wahid et al. (2007).

- Salinity

Soils are considered saline when the presence of soluble salts, such as sodium and magnesium chlorides and sulphates, exceed the electrical conductivity concentration of 4 dSm^{-1}. Concentrations beyond this value are considered to impair crop growth and rice is considered a saline sensitive species, with a threshold tolerance of 3 dSm^{-1} (Ghassemi & Nix, 1995). As an abiotic stress factor, salinity has one of the largest impacts on rice yields, second only to drought. Approximately 30% of rice growing land suffers from soil salinity, which is likely to increase in future climates due to rising sea levels and water limitations (Singh et al., 2021). High salinity conditions cause two major stressors, osmotic and ionic stress. Osmotic stress occurs in response to external high salinity, leading to an inhibition in water uptake and cell expansion. While ionic stress occurs in response to an internal toxic accumulation of ions (such as Na^{+}) in plants tissues, leading to reduced photosynthesis and leaf mortality. Both processes lead to a reduction in productivity and tissue necrosis (Horie, Karahara, & Katsuhara, 2012).

Flooding

Cultivated rice is one of the crops at greatest flood risk and experiences substantial annual yield losses as a result, where globally 22 million hectares of rice fields experience annual flooding, affecting over 100 million people (Singh et al., 2016; Panda & Barik, 2021). Flooding has multiple detrimental effects on the plant, which different species mitigate dependent upon their evolutionary adaptation to flooding stress. As a semi-aquatic species, rice demonstrates suites of adaptive strategies to flooding related abiotic stress.

Photosynthesis is impaired during the submergence of plants due to limited opportunity for gas exchange, reduced light levels and disintegration of chloroplasts. Due to submergence and anaerobic soil conditions, leaf CO_2 concentrations decline and produce ROS, which when left unchecked can cause cellular damage and trigger apoptosis (Bailey-Serres et al., 2012; Panda & Barik, 2021). When submerged, some species have developed an ‘escape strategy’ to rise the canopy above the water line and restore gas exchange. Rice is such as species, where the increased ethylene concentration in submerged plant tissues causes the enhancement of shoot growth (Mustroph, 2018; Panda and Barik, 2021). However, this rapid elongation can deplete carbohydrate stores, so rice varieties with slower shoot elongation are preferred in

cultivation areas that experience brief flash flooding (Panda and Barik, 2021).

Root traits are vital in the fight for survival during flooding events. The physical force generated through heavy rainfall and sudden flooding can destabilise a plant and cause physical damage to delicate rice roots and shoots. Some species, such as rice and *Zea mays*, develop adventitious roots that helps to withstand this force and maintain stability during water logging (Mustroph, 2018). Rice stems develop adventitious root primordia at each node, only emerging in response to flooding (Lin and Sauter, 2018). The formation of aerenchyma is essential for the survival of waterlogged plants, which facilitate the diffusion of oxygen from shoot to root. The formation of aerenchyma is ethylene-induced, formed by programmed death of cortex cells. Rice constitutively form aerenchyma even under aerobic conditions, which is then further induced under anoxic conditions, when tissue concentrations of ethylene rise (Mustroph, 2018, Mohammed et al., 2019; Panda and Barik, 2021).

Flooding events can also cause further damage to plants by increasing the solubility of metals in the soil, iron being an example. Iron usually occurs as Fe^{3+} but when flooded for prolonged periods (greater than two to three days), Fe^{3+} is reduced to Fe^{2+}, which is toxic to plants at high concentrations, causing distribution to cellular homeostasis (Oort, 2018). Due to this, paddy grown rice is at particular risk to iron toxicity. Rice utilises four main mechanisms to defend against iron toxicity, each with its own complex molecular pathway. In order of events, these include 1. Suppression of iron uptake by the roots, 2. Iron retention in roots, to prevent damage to the canopy, 3. Iron mitigation at the shoot via compartmentalisation in old leaves and 4. ROS detoxification in the body of the plant in response to iron stress (Aung and Masuda, 2020). While rice demonstrates adaptation to iron stress, this is still an area of concern and ongoing research.

As a semi-aquatic species, rice demonstrates natural adaptation to water logging and flooding, crop development efforts are seeking to enhance these traits and minimise the effects of flooding on future food security (Panda and Barik, 2021).

The effects of abiotic stressors are substantially more complex and interconnected than the simple summary provided here and rarely occur in isolation. For example, drought stress is usually accompanied by high light radiation and heat stress. For this reason, the physiological effects of abiotic stress in plants are complex but often share commonalities, particularly in

the limitation of photosynthesis. Similarly, there are commonalities between plant defence responses to various abiotic factors. As described above, the production of reactive oxygen species (ROS) is a ubiquitous response to abiotic stress in plants. If left unchecked ROS can cause catastrophic cellular damage through oxidative stress, but ROS are also useful transduction signalling molecules that mobilise genetic responses that mitigate the effects of abiotic stress (Choudhury et al., 2017). Alongside ROS, heat shock proteins (HSPs) and proteins in the MAPK and CDPK super family are thought to be key signalling molecules to trigger responses to protect against the effects of abiotic stress on photosynthesis (Fig. 1.2; Cheng et al., 2002; Wahid et al., 2007; Ashraf & Harris, 2013). Phytohormones such as abscisic acid (ABA), gibberellin (GA), cytokinins, salicylic acid (SA), auxin and others also play an essential role in signalling cascades in response to abiotic stress (Fig. 1.3). For example, under heat and drought stress ABA induction is important in the expression of HSPs (Wahid et al., 2007; Wang et al., 2015) and GA signalling, causes a reduction of stomatal conductance (Gaion & Carvalho, 2021). Although there will be variation between plant responses to differing abiotic stress factors, it is worth noting that many mechanisms appear to overlap in the common goal of protecting photosynthesis (Fig. 1.3).

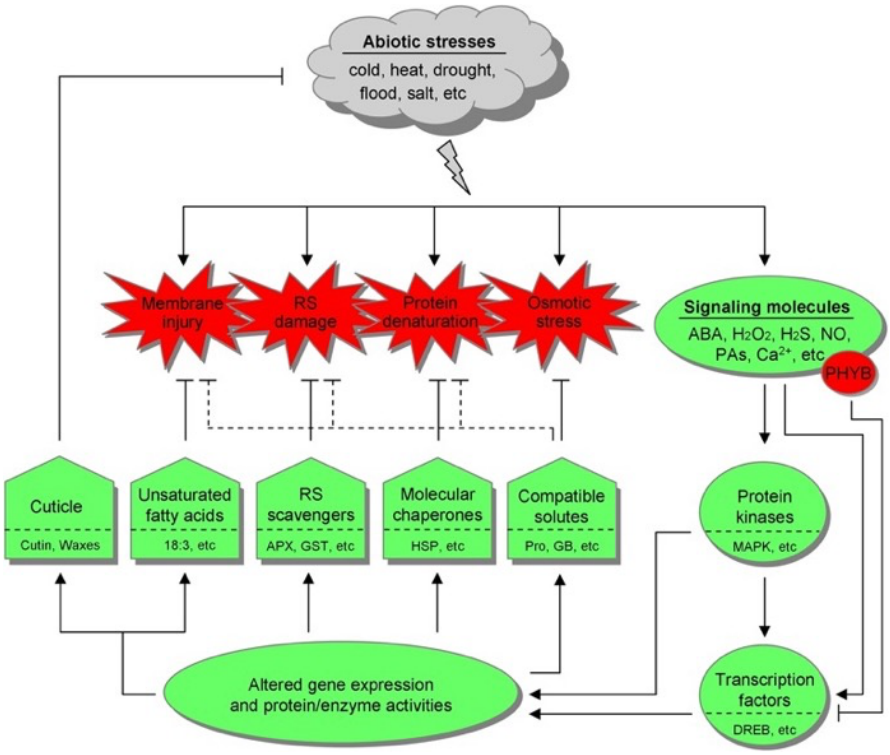


Figure 1.3: Generalised biological responses to abiotic stress factors. Different abiotic stresses can cause similar biological injuries, including protein denaturation, osmotic stress, membrane and reactive species (RS) damage. Upon experiencing abiotic stress, common signalling responses (ROS, calcium ions, phytohormones etc) mobilise protein kinases and transcription factors to alter gene

expression pathways and launch defence systems to protect cellular function. Image taken from (He et al., 2018).

The increasing frequency of extreme weather events and global warming means that crop improvements are required to develop resilience to these abiotic pressures, while simultaneously enhancing yield. A greater understanding of the physiology and genetic components associated with abiotic stress tolerance in rice will be essential for the development of productive climate change resilient rice.

1.3 Desirable traits for drought, heat and high light resilience in rice

Drought and heat stress pose the greatest threat to rice yields, which will decline if global temperatures continue to increase (Zhao et al., 2017). Understanding and harnessing the genetic basis of traits that contribute to drought and heat resilience is essential for securing future food security (Fig. 1.4). The tolerance of high light levels and optimising plant water use efficiency are important components when developing abiotic stress resilient crop varieties. However, the requirement to simultaneously enhance photosynthetic efficiency for yield gains is a substantial challenge (Ashraf & Harris, 2013; Kissoudis et al., 2016; Hubbart et al., 2018).

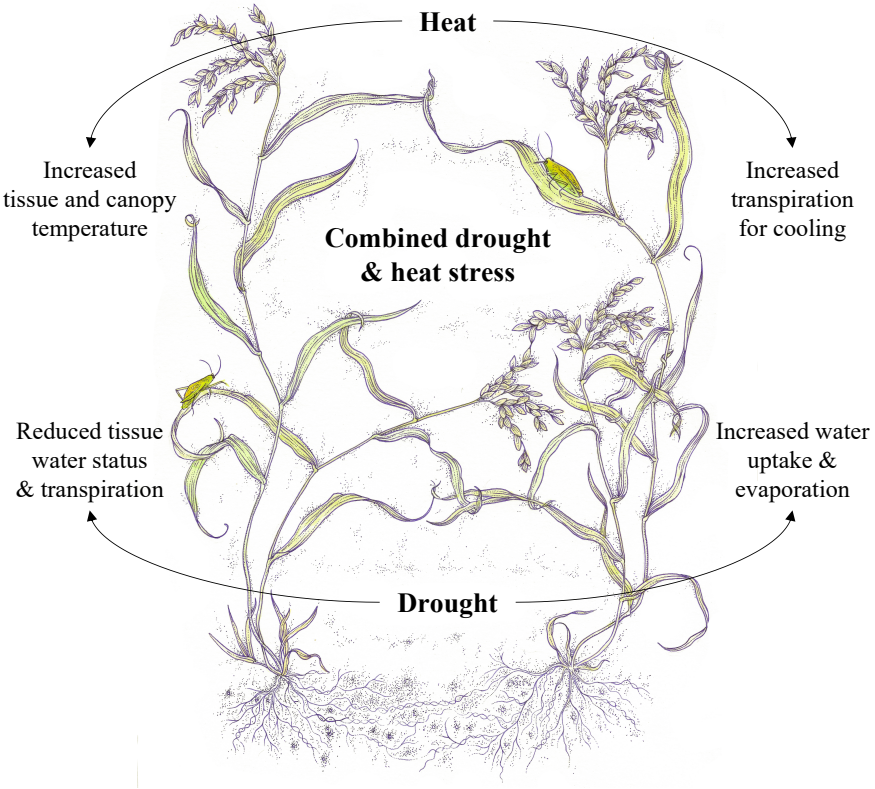


Figure 1.4: Strategies to reduce drought and heat stress are often conflicting. When these stresses are combined, trade-offs in physiological strategies are required to prevent catastrophic damage. The text in the image was taken from Costa et al. (2021).

Drought is often accompanied by high heat and light intensity, but the mechanisms used to reduce these effects can often be conflicting (Fig. 1.4). During heat stress, water loss through transpiration is increased to reduce tissue temperature through evaporative cooling. Whereas during drought, stomata close to reduce transpiration and conserve plant water status. These combined stresses limit photosynthesis through direct cellular damage and a reduction in carbon assimilation via stomatal conductance (*gs*). Plant water use efficiency (WUE) is a term almost ubiquitously mentioned when discussing drought and stomatal conductance. WUE can be analysed at three (Medrano et al., 2015; Medlyn et al., 2017):

1. Instantaneous leaf level - the instantaneous trade-off between carbon assimilation (*A*) for photosynthesis and water loss via transpiration is often shown calculated as *A*/*gs.* At the leaf level, instantaneous measures of WUE can be calculated via the measurement of *A* and *gs*, using an infra-red gas analyser (IRGA) to monitor changes in gas exchange (Lawson and Blatt, 2014; Medrano et al., 2015; Medlyn et al., 2017).
2. Plant level - At the whole plant level, carbon isotope discrimination of $\delta^{13}C$ in plant tissues can be used as a proxy for WUE. Due to the preference for the lighter ^{12}Carbon (^{12}C) during photosynthesis, the ratio of carbon assimilation to stomatal conductance (and therefore WUE) can be estimated using the ratio of ^{12}C to ^{13}C in plant tissues (Adams, Buckley and Turnbull, 2020). This method is used to measure WUE at the plant level and can also be used to estimate WUE over varying time scales, such as monthly or annually (Medlyn et al., 2017; Adams, Buckley and Turnbull, 2020).
3. Groups of individuals - At the largest scale, WUE can be estimated across many individuals, such as the measurement of crops. This method is simply measured as biomass over water loss via evapotranspiration. The latter can be measured in several ways, including the weight of plant pots and the use of soil sensors combined with satellite data (Medrano et al., 2015).

Because the main methods of WUE analysis vary across the scale of the plant, it is difficult to directly compare the estimate of WUE between these methods (Medlyn et al., 2017). The most appropriate measure of WUE will often depend upon the research question, capacity, equipment and funding resource available. Isotope discrimination and the measurement of gas-exchange using IRGAs can be costly, requiring specialist equipment. While the

measurement of WUE via the ratio of biomass to water is cheap but lacks fine detail that may be useful for further analysis.

As WUE improves when stomata close or reduce their aperture, research focusing on the improvement of WUE often surrounds stomata. Stomatal traits, photoprotection and photosynthesis are three interlinked components that can be optimised in the trade-off between water use efficiency and carbon assimilation. These traits and their importance in drought and heat tolerance are described below.

- Stomata

  Stomata are small dynamic pores on the leaf surface that regulate water loss via transpiration and carbon assimilation (*A*) for photosynthesis. Stomata are important gatekeepers between the internal and external environments, managing gas exchange via stomatal conductance to effectively balance water use efficiency (WUE) and productivity (Lawson & Blatt, 2014). Light, CO_2 and vapor pressure deficit (VPD) are the main above ground drivers of stomatal movement, alongside soil status signalling from the root (Foyer & Noctor, 2020). Increasing light, low internal CO_2 and low VPD induces stomatal opening, to acquire CO_2 required for photosynthesis, while the reverse triggers closure (Outlaw, 2003; Moore et al., 2021). However, research has shown a hierarchical signalling response when controlling stomatal movements. For example, during drought stress stomata will remain closed to prevent damaging water loss, even in response to high CO_2 demands and rising leaf temperatures (Moore et al., 2021).

  There is a positive correlation between *gs* and *A*, but a negative correlation between *gs* and WUE, demonstrating plant productivity and WUE are directly at odds with one another (Fig. 1.4 and 1.5). Aperture is the dominant factor but the range of stomatal conductance values a plant can achieve has been shown to be affected by stomatal density and size. A higher maximum capacity *gs* (gs_{max}) is associated with smaller stomata at a greater density (Fig. 1.5), where there is a negative relationship between stomatal size and density across plant taxa (Hetherington & Woodward, 2003; Drake et al., 2013). Recent studies have shown the reduction of stomatal density, through the over expression of the *OsEPF1* (epidermal patterning factor) gene, reduced *gs* and improved drought tolerance in *O. sativa* (Caine et al., 2019a; Mohammed et al., 2019).

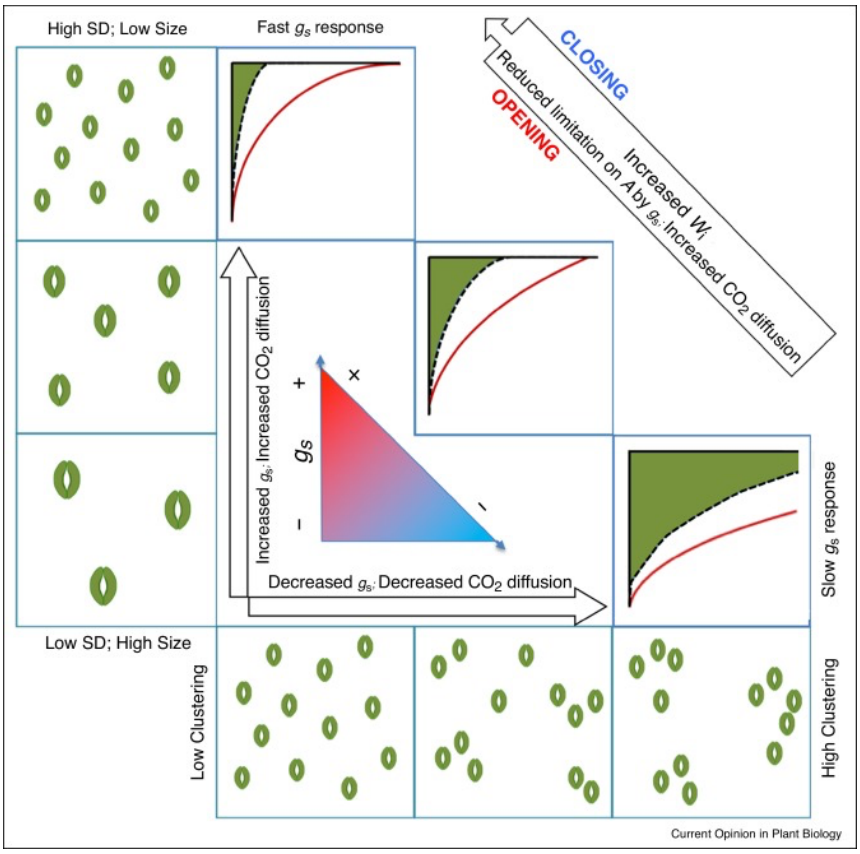


Figure 1.5: The effects of stomata size and density on conductance and dynamics. Arrows in the central box indicate the effect of stomatal size (left panel) and clustering (bottom panel) on stomatal conductance and carbon assimilation. The plots in the three diagonal boxes indicate the effects of stomatal size, density and clustering on the dynamics for stomatal conductance (red line), carbon assimilation (black dashed line) and the limitation of carbon assimilation by stomatal conductance (green area). The top-right arrow indicates the impact of stomatal anatomy on water use efficiency. Image taken from (Faralli et al., 2019).

Recent research has highlighted the importance of stomatal dynamics, that is the speed of stomata opening and closure, on photosynthetic effiency and WUE. Strong correlations between *A* and *gs* are reported for steady state values at maximum capacity (Farquhar & Sharkey, 1982). However, this rarely occurs due to fluctuating fields environments and the requirement for plants to continually balance *A* and water status requirements (Kaiser et al., 2018; Slattery et al., 2018). Due to the physical movement of stomata, dynamic *gs* induction and relaxation reponses lag behind the biochemical reactions of photosynthesis. The effect of delayed stomatal opening has been estimated to limit *A* by 20%, causing a substantial impact on potential crop

productivity (Lawson & Blatt, 2014; Taylor & Long, 2017). Likewise, delayed stomatal closure can result in unncessary water loss and reduce WUE (Vialet-Chabrand et al., 2017; Lawson & Vialet-Chabrand, 2019). Research has proposed subtantial yield gain and enhanced WUE can be obtained through enhancing the rate of stomatal dynamics, causing the close coupling of *gs* and *A* (Faralli et al., 2019). Stomata are an integral layer in a complex system, these microscopic pores have integral roles in pathogen defence, whole plant water status and plant productivity. Compared to most plant taxa, rice have small stomata with rapid dynamics (McAusland et al., 2016). The effect of stomatal dynamics on photosynthetic efficiency and WUE is still a relatively unexplored area in crop species, and suggested as a possible mechanism to improve drought tolerance in rice (Qu et al., 2020). It is clear that with dedicated research on the interplay between stomatal density, size and dynamics, there are substantial gains to be made in photosynthesis, water use efficiency and abiotic stress tolerance.

- Photosynthesis

  Photosynthesis is one of the most important physiological processes on Earth, almost all life in the biosphere depends upon the oxygen and energy-rich carbohydrate products from this biochemical reaction. The importance of this process can be seen in the consequences of the abiotic stress factors discussed above, where photosynthesis and plant growth is affected in all examples given. Photosynthesis is a complex energy transducing process, involving CO_2, water and light energy, to produce an essential source of carbohydrate energy to power plant function and growth. Chloroplasts house the photosynthetic reaction centres, where the light driven reactions in the thylakoid membranes split water into O_2, protons and electrons for the generation of the energy molecules ATP (adenosine triphosphate) and NADPH (Nicotinamide adenine dinucleotide phosphate). These are used to reduce CO_2 in the Calvin-Benson cycle reaction, converting carbon into a carbohydrate energy source in the chloroplast stoma (Johnson, 2016).

  In plants possessing stomata, approximately 98% of CO_2 for photosynthesis is obtained through stomatal conductance and as discussed above, stomatal traits can be exploited to enhance productivity.

However, drought stress can inhibit CO_2 acquisition, when stomata close to prevent water loss (Moore et al., 2021) and it is useful to separate the inhibition of photosynthesis by water deficit into those which directly result from a lack of access to CO_2 (due to stomatal closure) and those which are an impairment of photosynthetic metabolism caused partly by associated stresses.

Stomatal closure also prevents evaporative cooling and can potentially damage delicate photosynthetic organelles and proteins through increased leaf temperature (Caine et al., 2019a). Furthermore, under severe drought stress leaf loss of turgor and adverse water potentials can lead to reduced metabolic processes required for photosynthesis and utilisation of CO_2 (Ashraf & Harris, 2013). Due to the complexities of photosynthesis, the strategies of abiotic stress tolerance are yet to be fully elucidated. However, the upregulated expression of protective proteins (Fig. 1.2) and higher turnover of Rubisco, thylakoid membrane proteins and other machinery related to photosynthesis is suggested to help maintain photosynthetic function and limit the effects of drought (Georgieva et al., 2010; Jarvit et al., 2013). To fully understand and utilise these protective molecular processes, the future characterisation of abiotic stress induction signalling pathways is essential for enhancing photosynthetic tolerance and productivity in challenging conditions (Bailey-Serres et al., 2019).

The targets for enhancing photosynthetic productivity under optimal conditions are relatively better understood than abiotic tolerance mechanisms and tends to mirror those identified for stomatal conductance. Historically, gains in biomass and yield were achieved through the introduction of dwarfing varieties and enhancement of maximum carbon assimilation (A) , which continues to have validity as biomass can be correlated to high levels of A (Carmo-Silva et al., 2017). But under fluctuating field conditions and almost constant changes in stomatal conductance, recent research has also highlighted the importance of the speed of photosynthetic induction as a source of yield improvement (McAusland & Murchie, 2020). Rapid photosynthetic induction utilises fluctuating light more effectively, whereas slow responses can substantially reduce carbon acquisition throughout the day (Vialet-Chabrand et al , 2017). Taylor and Long (2017) estimate 21% loss of carbon assimilation in wheat due to slow photosynthetic induction. Recent work in rice and wheat has suggested that

biochemistry, such as carboxylation rate of Rubisco, places the greatest limitation on photosynthetic induction (Taylor & Long, 2017; Acevedo-Siaca et al., 2020, 2021). Showing that while CO_2 assimilation and stomatal conductance are interlinked, it is not the only limiting factor for productivity and provides other target traits to explore.

- Photoprotection

  Increasingly high light intensities are likely to accompany drought and heat stress, which can have detrimental effects on photosynthetic efficiency and photochemistry. This is particularly relevant in rice growing regions near the equator, which receive especially high levels of solar radiation (Murchie et al., 2015). Under high light intensities photosynthetic machinery cannot utilise the absorbed excess light energy, this can lead to quantum yield reduction and photoinhibition. The latter is often considered to be inactivation of photosystem II because of protein damage (Murchie and Ruban 2020). To limit photoinhibition there are several photoprotective mechanisms that occur in the thylakoid membranes (Müller et al., 2001). Non-photochemical quenching (NPQ) is one of the most effective mechanisms, which acts by releasing the built-up excitation energy in the light-harvesting complexes of photosystem II (LHCII) as heat, thus protecting the photosynthetic antenna from over excitation and photoinhibition (Murchie & Ruban, 2020; Ruban & Wilson, 2021). While NPQ is a photoprotective mechanism, it can also momentarily behave in a photoinhibition-like manner. As part of the mechanism to dissipate excitation energy as heat, NPQ switches the LCHII between a light-harvesting to energy dissipation state. This temporarily reduces quantum yield and therefore photosynthesis, causing a trade-off between photoprotection and photosynthesis (Slattery et al., 2018; Murchie & Ruban, 2020). Relaxation from photoprotection takes minutes to achieve whereas photoinhibition can take days, requiring protein synthesis and a source of energy.

  NPQ has been receiving increasing attention as a potential method to enhance both photoprotection and improve photosynthetic efficiency. Under high light NPQ provides essential photoprotection, but under fluctuating light conditions NPQ relaxation can lag behind the change in light levels. This places a limitation on photosynthetic efficiency and crop yields (Kromdijk et al., 2016). Recent research has shown that enhancing NPQ relaxation responses to low light increases photosynthetic

efficiency and improves productivity (Fig. 1.6; Kromdijk et al., 2016; Hubbart et al., 2018). While NPQ research is still in its infancy, it is clear this is an important trait when developing climate resilient cultivars. Improved NPQ capacity and dynamics could be harnessed to protect rice crops from increasing light intensities and enhance photosynthetic efficiency for much needed yield gains.

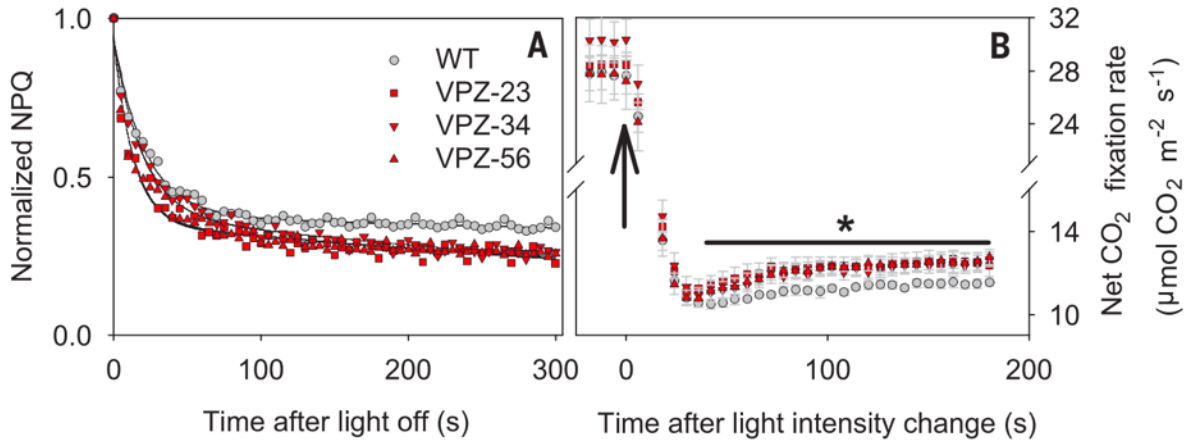


Figure 1.6: CO_2 fixation rates are 9% higher in mutant lines exhibiting a rapid NPQ relaxation rate. Here, the (A) NPQ relaxation rate and (B) CO_2 fixation rate was compared between *Nicotiana tabacum* wild type (WT) and transgenic lines that overexpressed three key NPQ proteins (VPZ-23/34/56). The left y-axis shows the normalised NPQ value and right y-axis net carbon assimilation, the x-axis shows the time (seconds) after (a) the light was switched off and (b) the light intensity was reduced from 2000 – 200 PPFD. Error bars show standard error. Image taken from (Kromdijk et al., 2016).

Stomatal, photosynthetic and photoprotective traits are essential to producing climate resilient, yet productive rice cultivars. However, the natural variation for these traits in rice is relatively unexplored, and therefore underutilised. Characterisation in genetically diverse rice populations will assist identifying useful phenotypes to facilitate crop improvement efforts.

1.4 *O. glaberrima* as a source of natural variation for potential *O. sativa* crop improvement.

The search for abiotic stress resistant rice cultivars has been subject to increasing research focus in recent years (Atwell et al., 2014). Advances have been made through traditional plant breeding methods but recently attention has turned towards the introgression of genes from a range of diverse interspecific material. Including the African rice species, *Oryza*

glaberrima, which demonstrates exceptional resilience to many abiotic stress factors (WARDA, 2006; Bocco et al., 2012).

Oryza sativa and *Oryza glaberrima* are the only two domesticated crop species in the genus *Oryza* (Muthayya et al., 2014) but evolved independently to one another, last sharing a common ancestor approximately 600,000 years ago. *O. glaberrima* was domesticated approximately 3000 years ago independently to *O. sativa*, which underwent domestication around 10,000 years ago (Fig. 1.7; Atwell et al., 2014; Wang et al., 2014). *O. glaberrima* was cultivated along the West-African coast and has not been readily utilised outside of its domestication centre, where it has been steadily replaced by the higher yielding *O. sativa* (Muthayya et al., 2014).

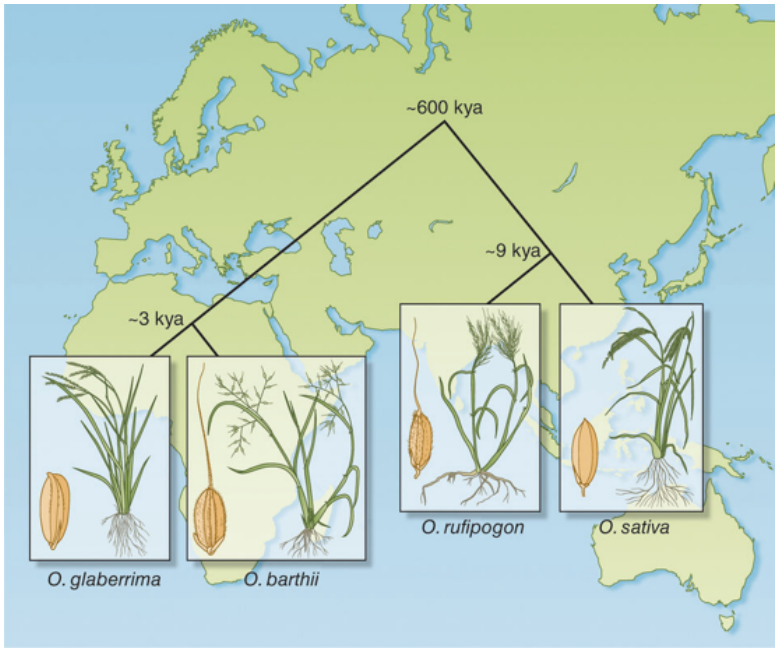


Figure 1.7: The independent evolution and domestication of African and Asian rice species. Modern *O. sativa* and *O. glaberrima* varieties evolved in parallel, independently to one another on separate continents. While both provide a strikingly similar food source, each has evolved in response to independent environmental selective pressures, and this differential selection process could potentially be exploited to improve current commercial cultivars. Image taken from Purugganan, 2014.

The process of crop domestication through artificial selection, and the highly selective pressures of intensive agricultural environments, can cause large genetic changes over

comparatively short spaces of time. *O. glaberrima* appears to have experienced a severe bottleneck during the domestication process and thus is genetically narrow in comparison to *O. sativa* (Nabholz et al., 2014). *O. glaberrima* also exhibits undesirable traits, such as; milling difficulties and lower yields than *O. sativa* due to lodging and grain shattering (Linares, 2002). However, it shows physiological advantages over its more commercial sibling, including out-competing weeds with wide leaves and a swift maturation time, the latter making it an excellent emergency food source (Linares, 2002; Atwell, Wang, & Scafaro, 2014). Despite not being commercially viable, *O. glaberrima* demonstrates substantial resilience to a range of biotic and abiotic stresses, such as; bacteria, viruses, nematodes, iron toxicity, high salinity, nutrient deficiencies and drought (Linares, 2002; Albar et al., 2003; Orjuela et al., 2014). As a result, *O. glaberrima* is an untapped source of natural variation for the improvement of *O. sativa*.

O. sativa and *O. glaberrima* are closely related, and while both possess the AA genome (Xu et al., 2014), interspecific hybridisation between the two species is restricted by a sterility barrier at the S_1 locus. Hybridisation efforts result in 100% F1 hybrid sterility and this has impaired *O. sativa* interspecific breeding efforts (Orjuela et al., 2014; Shen et al., 2015). Consequently there has been valuable research directed into characterising the molecular basis of the sterility barrier between the two species to overcome these challenges (Xu et al., 2014; Shen et al., 2015; Li et al., 2020).

Despite these difficulties, interspecific hybrids have been generated by West Africa Rice Development Association (WARDA). Fertile interspecific hybrids between *O. glaberrima* Steud. and *O. sativa* L. were generated through a combination of traditional breeding techniques and the more complex anther culture and double-haploidisation, followed by embryo rescue and backcrossing (Fig. 1.8; WARDA, 2008). The lines produced from these dedicated breeding efforts are known as the NEw RICe for Africa (NERICA) cultivars (WARDA, 2006; Ikeda et al., 2009).

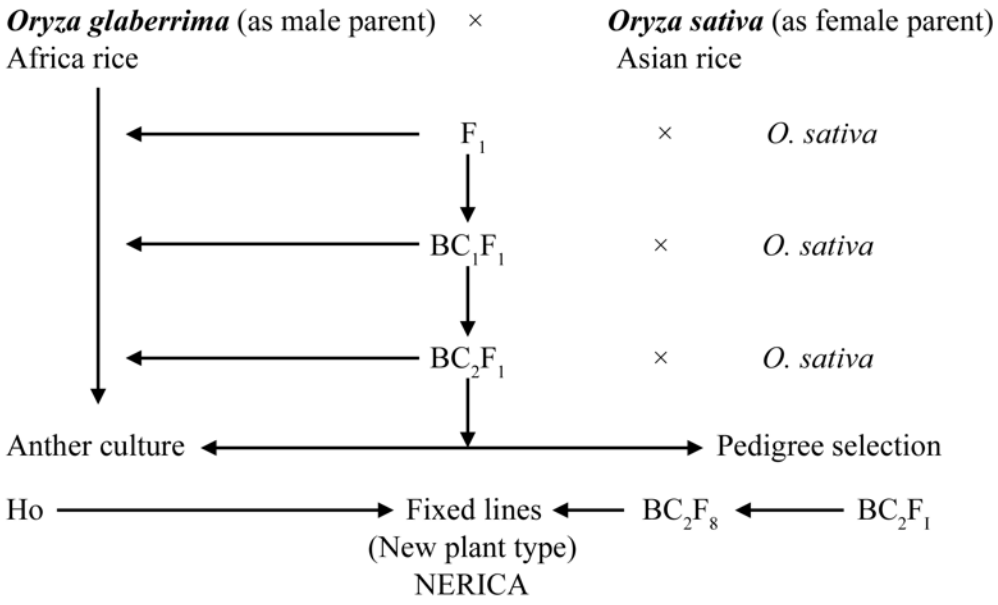


Figure 1.8: Schematic of the generation of NERICA cultivars. As 100% of F_1 progeny produced by the *O. glaberrima* x *O. sativa* cross is sterile, anther culture and embryo rescue is performed to facilitate repeated backcrosses with *O. sativa* to produce fertile interspecific hybrids. Four backcrosses produce ~85-100% fertile, fixed lines which are then selected for desirable traits. The generation of upland NERICA varieties utilised *O. sativa* japonica, which is traditionally a rain fed rice species. Image taken from Nassirou & He, 2011.

NERICA cultivars were initially selected for their tolerance to the dry, rainfed upland ecology, representative of large swathes of West and Central African agricultural regions (WARDA, 2008). Seven NERICA upland cultivars were released in 2000, followed by another 11 in 2006 (Ikeda et al., 2009). Since this time these varieties have been distributed across Africa to local farmers and further breeding efforts have been specifically targeting wet lowland agronomic conditions (WARDA, 2008). Ongoing research on the varieties shows that NERICAs can produce desirable features of both species, including those found in *O. glaberrima* such as; tolerance to pests, nutrient deficiencies and drought under controlled conditions while reflecting the higher yields of *O. sativa* (Orjuela et al., 2014; Kikuta et al., 2017). However, major criticisms of the NERICA varieties include loss of seed viability over time and inconsistency in the performance of the cultivars experienced by farmers in field conditions, to the extent where 50% of farmers adopting it in 2004 had abandoned its use by 2006 (Kijima et al., 2011). As a UN funded humanitarian effort, NERICAs were not developed to replace commercial *O. sativa* crops but provide a reliable food crop for West African farmers and an economic development opportunity (WARDA, 2006). Few studies

have assessed the impact of NERICA on African farmer's livelihoods, however a study conducted by Kijima et al. (2008) suggested that NERICAs can increase income and decrease poverty significantly. Although this is suggested to only be the case when accompanied by appropriate agricultural equipment and cropping patterns, of which many African farmers simply do not have access to or rice specific agricultural knowledge. The problems encountered suggest that NERICAs may not be the solution as was initially hoped. Furthermore, the crop yields and initial purpose of the NERICA varieties is not appropriate for the replacement of high yielding *O. sativa* cultivars.

1.5 Identifying and utilising the natural variation in *O. glaberrima*

O. glaberrima is an underutilised source of genetic diversity and abiotic stress tolerance for *O sativa* improvement. The generation of NERICAs was based on the production of a range of interspecific hybrids using a small number of *O. glaberrima* parent accessions and did not target the characterised genetic variation of beneficial traits in *O. glaberrima* (Monat et al., 2017). In this genomic era, it seems a logical step forward to first identify genes and alleles associated with traits of interest in *O. glaberrima*. Genome editing techniques such as CRISPR/Cas9, would be a direct method to introduce useful variation into the *O. sativa* genome, if trait-related genes of interest are found to be homologous between the two species (Song et al., 2016). However, complex traits associated with abiotic stress tolerance are likely to be quantitative and therefore will require good quality sequenced *O. glaberrima* accessions for gene identification (Tuberosa & Salvi, 2006). Until recently such data was lacking, except in the instance of the few cultivars used in NERICA production and even these had 20-30% gaps in the genome sequence (Orjuela et al., 2014; Monat et al., 2017). The RICE (Rice, Interspecies Comparison and Evolution) group within the Research Institute for Development (IRD), University of Montpellier, have gathered a panel of genetically diverse *O. glaberrima* accessions as a resource for *O. sativa* improvement (Cubry et al., 2018). The re-sequencing efforts of these accessions has provided a high-quality resource to identify trait-related genes within *O. glaberrima*, through analyses such as genome wide association studies (GWAS) (Cubry et al., 2020) and quantitative trait analysis (QTL) (Dufey et al., 2015). QTL analysis and allele mining has been used with *O. glaberrima* to identify loci associated with biotic and abiotic stress tolerance, such as iron toxicity (Dufey et al., 2015) and rice yellow mottle virus (RYMV) resistance (Thiémélé et al., 2010; Pidon et al., 2020).

GWAS has been used with great success in plant science to identify trait-related genes. With the advances in sequencing technology and high throughput phenotyping, GWAS is an important tool that can be used to explore the causal relationship between a phenotype and genotype within a group of organisms. (Wang et al., 2016; Zaidi et al., 2016). In a brief summary, GWAS analysis uses a sample population of a species whose genome has been characterised for polymorphisms, such as high-density single nucleotide polymorphism (SNP) arrays. The chosen group of individuals are then phenotypically characterised for traits of interest, to capture trait variation. An association algorithm is then used to scan all individual genomes, testing for a statistical association between a phenotypic trait and common genetic variation at a locus within the genome (Korte and Farlow, 2013).

Like any analysis, GWAS has its strengths and weaknesses. From a practical perspective, GWAS is a reproducible, cost-effective method that (when shared) produces useful results that the wider scientific community can take forward in several ways (Tam et al., 2019). A particular strength of GWAS over traditional gene mapping methods, is that it can be performed on collections of individuals with unknown ancestry (Korte and Farlow, 2013; Tam et al., 2019). This overcomes the main limitation of QTL analysis, which requires established recombinant inbred line (RIL) families or F2 populations that have co-segregating loci for a trait, which does not represent the natural functional diversity. GWAS can also resolve mapping resolution, which can be a challenge in QTL analysis. However, it is recommended that GWAS and QTL analysis are conducted together, as they mitigate the different limitations in each method (Korte and Farlow, 2013).

GWAS is a powerful method that has been proven to identify novel trait-gene associations and reveal new biological mechanisms across a myriad of species. However, there are several limitations that need to be considered when embarking on an association study. While GWAS is relevant to rare and low-frequency variants, these can be easily missed, and the analysis is better suited to the identification common sources of genetic variation (Korte and Farlow, 2013; Tam et al., 2019). A small sample size can place a limitation on GWAS efficacy, for this reason most GWAS studies will used sample sizes over >300 and may extend to thousands of individuals in human studies. Increasing the sample size increases statistical power, thus the detection of polygenic traits with small effect sizes, and captures a larger proportion of the population that may exhibit low-frequency or rare variants (Hong and Park, 2012). Conversely, large populations exert a higher degree of kinship and population structure, this can inflate the detection of false positives due to allelic fixation through genetic

drift, if not appropriately accounted for. Similarly, significant associations can be missed if these confounding factors are over corrected (Brachi, Morris and Borevitz, 2011; Korte and Farlow, 2012; Tam et al., 2019).

These limitations may seem daunting but can be mitigated through good quality genome sequencing and assembly, appropriate experimental design, large sample sizes and accurately accounting for confounding factors (Korte and Farlow, 2013). In plants species particularly, GWAS is a powerful method that is actively utilised to elucidate trait-gene relationships. GWAS has been shown to successfully to partition complex traits. For example, the genes accelerated cell death6 (ACD6) in *Arabidopsis thaliana* (Todesco et al., 2010) and vacuolar H^+ pyrophosphatase (VVP1) in *Zea mays* (Wang et al., 2016), were identified via GWAS and later functionally validated. GWAS is widely used to elucidate the genetic basis of key agronomic traits across crop species, such as flowering time, plant height and grain filling (Cortes, Zhang and Yu, 2021). Recent studies using GWAS in rice (*O. sativa*) explored the genetic basis of stomatal traits (Chen et al., 2020) and Cubry et al. (2020) identified a key *O. sativa* flowering time gene co-locating in *O. glaberrima*.

As a self-fertilising species, *O. sativa* has proved a highly successful candidate for GWAS studies (Cortes, Zhang and Yu, 2021). However, considering the importance of *O. glaberrima* as a source of heritable natural variation in abiotic stress tolerant traits, the exploration of the phenotypic and genetic diversity in the species is surprisingly limited. However, the recent re-sequencing efforts has opened an exciting new opportunity to mine for abiotic stress trait-related genes using GWAS.

While *O. glaberrima* offers many favourable traits for *O. sativa* improvement, research efforts on identifying trait-related genes to-date have focused on the tolerance to iron toxicity and viral stresses (Albar et al., 2003; Dufey et al., 2015). However, climate change is the biggest challenge facing the future of the planet and providing cultivars with tolerance to abiotic stressors is essential. The traits associated with abiotic stress tolerance and productivity are well documented. Therefore, it seems sensible that future research efforts should be directed towards identifying the trait-related genes in *O. glaberrima*. The investigation of traits, is often focused to one small plant process (Giuliani et al., 2013; Uga et al., 2013; Thi et al.,2016). While this makes the elucidation of specific traits and genes more tractable, it does not take into consideration that such responses are often intertwined. The relationships between traits may be obscured when measured in isolation. Therefore, the search for climate resilience related traits within *O. glaberrima* should ideally be approached

from a broader perspective, exploring the desirable traits mentioned above and their interconnecting associations.

1.6 Statement of research purpose

O. glaberrima possesses many novel abiotic stress tolerant qualities that can be exploited as a source of rice crop improvement. Despite being recognised as a source of novel genetic diversity for decades (Ghesquière et al., 1997; Sarla & Swamy, 2005), *O. glaberrima* remains relatively uncharacterised both phenotypically and genetically. The recent collection and re-sequencing of genetically diverse *O. glaberrima* accessions (Cubry et al., 2018) provides an ideal opportunity to characterise the genes, or define regions within the genome, which are responsible for traits related to abiotic stress tolerance. Taking into consideration the serious impact of climate change and projected desertification in large areas of the world, considerable research efforts should be directed towards investigating the capabilities of *O. glaberrima* as a source for crop improvement.

This project will explore the heritable natural variation within the re-sequenced population of 155 *O. glaberrima* accessions. To achieve this, comprehensive phenotyping efforts will characterise the natural variation across a range of morphological and photosynthetic traits. Using this information, a GWAS will be completed to identify trait-related candidate genes and regions. Further to this, physiological associations between complex traits will be explored to highlight interesting relationships for future research. The research completed here aims to identify useful trait linked genetic markers, for the contribution to future translation and adaptation of *O. sativa* to a changing climate.

CHAPTER TWO

Stomata Detector: High-throughput automation of stomata counting in a population of African rice (*Oryza glaberrima*), using transfer learning.

2.1 Chapter two introductory statement

The manuscript documents the development of a machine learning based software that automates the counting of stomata in a micrograph. The project conception arose from the large volume of stomatal impressions, and subsequent microscopy images (13,110), that were generated from phenotyping (further described in Chapter 3) the *O. glaberrima* population detailed in this thesis. Characterising the variation in *O. glaberrima* stomatal density was important to the objectives of the PhD project, but the manual phenotyping of stomatal density would have been too slow (estimated 9-12 months on top of 3-month image acquisition). There has been recent burgeoning of high throughput stomatal image analysis, but at the time (2018-2019) there was little-to-no open-source software on the subject. One resource was found to be freely available at the time, 'StomataCounter' (Fetter *et al.*, 2019) from University of Vermont and Brown University. However, when tested it did not accurately identify rice stomata, had a slow image processing time, and did not offer a function to batch process large numbers of microscopy images. When repeatedly tested thereafter (2019 - 2021), the StomataCounter resource portal could not process images at all, despite our micrographs meeting the guidance specifications. For this reason, an automated stomatal counting method was developed that would be more accurate than manual methods, with the intention to be further developed as a shared community resource.

This chapter consists of a draft manuscript that has been submitted and accepted to the pre-print repository, bioRχiv (doi.org/10.1101/2021.12.01.469618). We aim to submit this to a peer reviewed journal, pending re-training with micrographs from a wider taxonomic group.

Text in grey was written by Dr. Hamidreza Soltani.

Stomata Detector: High-throughput automation of stomata counting in a population of African rice (*Oryza glaberrima*) using transfer learning.

Sophie B. Cowling[1], Hamidreza Soltani[2], Sean Mayes[1], Erik H. Murchie[1]

bioRχiv
doi: https://doi.org/10.1101/2021.12.01.469618
Posted: 3rd December 2021

[1] Division of Plant and Crop Science, School of Biosciences, University of Nottingham, Sutton Bonington Campus, LE12 5RD
[2]Advanced Data Analysis Centre, University of Nottingham, Sutton Bonington Campus, LE12 5RD

Author contributions:
SBC project conception, acquisition of stomatal peels and microscopy images, image annotation for machine learning training set, statistical analysis and biological advice for HS.
HS generation of automated stomatal counting algorithm and software.
EHM and SM project conception, guidance and advice.

ABSTRACT

Stomata are dynamic structures that control the gaseous exchange of CO_2 from the external to internal environment and water loss through transpiration. The density and morphology of stomata have important consequences in crop productivity and water use efficiency, both are integral considerations when breeding climate change resilient crops. The phenotyping of stomata is a slow manual process and provides a substantial bottleneck when characterising phenotypic and genetic variation for crop improvement. There are currently no open-source methods to automate stomatal counting. We used 380 human annotated micrographs of *O. glaberrima* and *O. sativa* at x20 and x40 objectives for testing and training. Training was completed using the transfer learning for deep neural networks method and R-CNN object detection model. At a x40 objective our method was able to accurately detect stomata (n = 540, $r = 0.94$, $p<0.0001$), with an overall similarity of 99% between human and automated counting methods. Our method can batch process large files of images. As proof of concept, we characterised the stomatal density in a population of 155 *O. glaberrima* accessions using 13,100 micrographs. Here, we present developed Stomata Detector; an open source, sophisticated piece of software for the plant science community that can accurately identify stomata in *Oryza spp.*, and potentially other monocot species.

2.2 INTRODUCTION

The improvement of elite crop varieties is essential in meeting increasing global food demand (Ray *et al*., 2019; FAO *et al*., 2020) and resilience to a changing climate (Kissoudis *et al*., 2016; Gao, 2021). A cornerstone in the development of crop resilience to global warming lies in balancing the trade-off between plant water relations and photosynthesis, in which stomata play a central role (Chaves *et al.*, 2002; Lawson and Blatt, 2014) . Stomata are dynamic microscopic pores in the leaf epidermis, which act as the gatekeepers between a leaf's internal and the external environments. Stomata open and close in response to internal and environmental stimuli, such as light and heat. In doing so, stomata regulate the degree of CO_2 assimilation for photosynthesis (*A*) and water lost through transpiration, both occurring during stomatal conductance (*gs*) (Lawson and Blatt, 2014; Kostaki *et al.*, 2020; Yang *et al.*, 2020). However, there is a trade-off between improved plant water use efficiency (iWUE = *A*/*gs*) and CO_2 assimilated for crop productivity (Blum, 2009; Lawson *et al*., 2010; Lawson and Blatt, 2014). Modifying the density of stomata on the leaf surface (SD) is one method that plants use to balance *A* with WUE, for example increasing SD in *Arabidopsis thaliana* has been shown to enhance photosynthetic rate by 30% (Tanaka, Sugano, Shimada, & Hara-Nishimura, 2013), while in rice (*Oryza sativa*, cultivar IR64) a reduction in SD by greater than 50%, in comparison to the wild type, showed an improvement in WUE without compromising photosynthesis and yield (Caine *et al.*, 2019; Mohammed *et al.*, 2019).
As a functional trait, understanding the effects of stomatal density is also important to several other plant research areas. The speed of stomatal dynamics of opening and closing places a limitation on photosynthetic efficiency under fluctuating light conditions, with consequences for both WUE and productivity and speed may be in part determined by stomata size (Drake *et al*, 2013; Lawson and Blatt, 2014). Stomata are also known to be key players in mediating pathogen resistance, where density and anatomy effect the likelihood of colonisation (Melotto *et al.*, 2008; McKown *et al.*, 2014; Muir, 2020). Consequently, stomatal density and morphology are an important research focus in the search for improved crop productivity and resilience in future climates.

Stomatal density has been shown to be a heritable trait, controlled by suites of genes, such as *EPIDERMAL PATTERNING FACTORs* (*EPFs)*, *EPF-LIKEs* and the *ERECTA*-family (Hara *et al.*, 2007; Hara *et al.*, 2009; Hunt and Gray, 2009; Zoulias *et al.*, 2018). The genetic control of stomatal development and patterning is relatively well-defined in *Arabidopsis thaliana* (Nadeau and Sack, 2002; Chowdhury *et al.*, 2021). Elucidating the genetic control of

stomatal features has only begun in crops relatively recently (McAdam et al., 2021), which is a major barrier when harnessing stomatal traits for crop improvement. Considering the important role that stomata play in carbon acquisition and water use efficiency, the successful introgression of diverse drought and heat resilient plant genotypes is essential to future crop improvement (Qu *et al.*, 2016; Faralli, Matthews and Lawson, 2019; Kimura *et al.*, 2020; Moore *et al.*, 2021). To achieve this, large scale stomatal phenotyping efforts are necessary to capture the diversity across crop species, their wild relatives and, ideally, the wider plant kingdom. This information can then be used to identify desirable stomatal traits, elucidate the underlying genetic control and the genetic diversity through genomic analysis for future breeding programs.

The scientific resources and characterisation of genomic research has rapidly evolved, becoming increasingly precise and high throughout. Whereas rapid, large scale trait measurement pipelines are sorely lacking and recognised to be a major data collection bottleneck. This limits crop breeding progress in global food and nutritional security (Mir *et al.*, 2019). The measurement of stomatal morphology and patterning is a manual process, that is slow, fraught with human error and limits whole population phenotyping efforts. The process comprises of three main steps; 1. acquisition of leaf surface impressions, usually via dental putty (non-destructive) or clear nail varnish (destructive but faster); 2. image acquisition using a microscope, obtaining up to 20 fields of view per peel, based on the microscope objective, peel quality and stomatal size; 3. Measurement of stomatal density and morphology, often completed using the image analysis software IMAGEJ (National Institute of Health, USA). Manual processing is an acceptable technique when working on small numbers of individuals but proves a significant limitation when measuring stomatal traits in a large population, which is necessary for documenting stomatal diversity and genomic analyses. As mentioned above, genomic analyses are increasingly used for identifying trait-related candidate loci and the underlying genetic variation in a trait of interest but require large numbers of individuals for accurate loci detection (Hong & Park, 2012).

There have been recent developments towards the automated analysis of epidermal and stomatal traits (Laga, Shahinnia and Fleury, 2014; Duarte, De Carvalho and Martins, 2017; Jayakody *et al.*, 2017), there are currently no established, open-source automated stomatal analysis methods available to hasten the process. Of the methods that have been developed, none can accurately identify stomata across variable taxa and do not reliably capture the core

crop species (wheat, rice, maize), that will be a focus in securing future food security. There is now an urgent need to develop automated, high-throughput phenotyping of stomatal traits for the scientific community, to benefit both fundamental research and crop improvement. In this project, we aim to substantially reduce the time, cost and error associated with the manual counting of stomata through the development of an automated counting method. This project uses a machine-learning based method to accurately count stomata from a microscopy image. As large volumes of images are produced during a population level phenotyping effort, the method can process a file of multiple images, as well as single micrographs. The accuracy and useability of the method is then tested by phenotyping stomatal density in a population of *O. glaberrima* accessions, consisting of 155 genotypes. Here, we develop a resource that will improve research standards and provide information on stomatal traits at a population level and assist in the identification of molecular markers in the selection of new crop varieties for a sustainable future.

2.3 MATERIALS AND METHODS

2.3.1 Plant material and growth

The seed of 155 *O. glaberrima* accessions was provided by Diversité Adaptation Developpement des plantes (DIADE), IRD-Montpellier, France (Supplementary Table S2.1). Plants were grown and measured in a controlled environment glasshouse (Cambridge HOK, UK) at the Sutton Bonington Campus, University of Nottingham, UK. Conditions were maintained at 28±3 ^{o}C, 50-60% relative humidity and a 12-hour dark:light (07.00 – 19.00 hrs) photoperiod. Light levels were controlled using blackout blinds and metal halide lamps were used to maintain light levels when they fell below 200 μmol m^{-2} s^{-1} photosynthetically active radiation (PAR). Seeds were heat treated to prevent pathogenesis at the primary seedling stage through immersing in water at 55^{o}C for 15 minutes. Seedlings were grown in module trays and transplanted to soil pits (5mx5mx1.25m, LxWxD) within the glasshouse at 2 weeks old. Five replicates of each accession were transplanted in east – west rows (Supplementary Figure S2.1), at 20cm intervals, into high nutrient loam-based soil. Accessions were grown and planted in rotations of twelve genotypes at a time, staggered at 1–2-week intervals. Plants were measured at eight weeks old, and the four healthiest plants for each accession were selected for measurement. Measurements commenced July 2017 and finalised October 2017

(Supplementary Table S2.2). The *O. sativa* variety 'IR64' was used as a reference genotype, as this is well characterised, and planted as a row in every batch.

2.3.2 Stomatal impressions and image collection

Impressions of the leaf epidermis were taken from an approximately 1cm^2 mid-section of the first fully expanded leaf, using fast drying clear nail polish and adhered to a microscope slide. Impressions of the abaxial (basal) and adaxial (upper) leaf surface were taken, as stomatal density and morphology can differ between the two leaf sides (Franks and Farquhar, 2007; Chatterjee *et al.*, 2020). Impressions of each leaf side were obtained from each of the four genotype replicate plants grown. An abaxial and adaxial impression was taken from these replicates, resulting in a total of eight impressions of the leaf surface: four abaxial and four adaxial. This was the case for each of the 155 accessions.

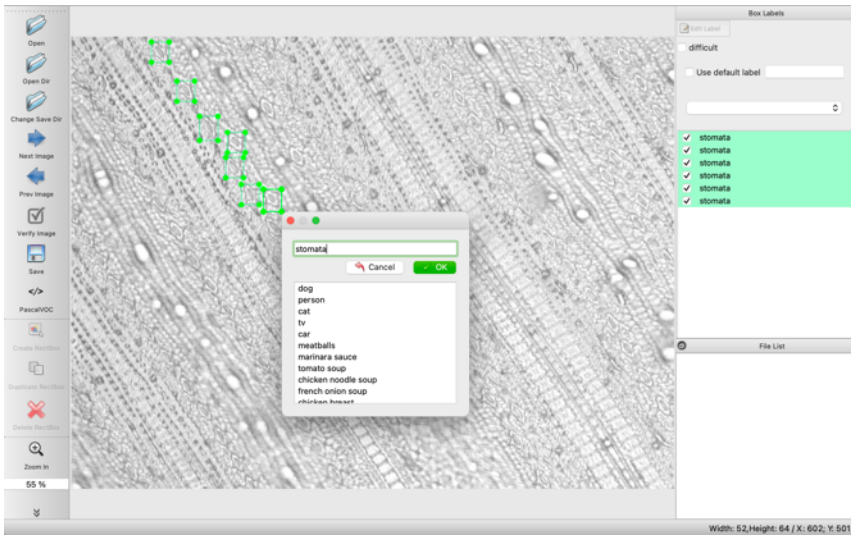
Images were obtained on a Leica DM5000B light microscope at x20 and x40 objectives, using the image settings; 70% brightness, 0.35 gamma and greyscale. Images were saved as a 1728x1944 intermediate quality resolution pdf, this was a compromise between image quality and size for computational processing speed. Microscopy images of impressions for the entire experimental panel of 155 *O. glaberrima* accessions and *O. sativa* 'IR64' replicates were obtained just using the x40 objective, with ten fields of view (FOV) per impression. Eight impressions were obtained for each *O. glaberrima* accession and the *O. sativa* IR64 check genotypes, with ten FOV obtained per impression. This produced a total of 13,110 micrographs to process for phenotypic analysis.

2.3.3 Image annotation

Image annotation is a necessary process in training a computer vision model to accurately identify an object within an image. Our process included the manual definition, using a bounding-box, and a text-based description of a stoma within a micrograph. The generation of annotated stomata microscopy images for training was completed using the open source 'Labellmg' software (Tzutalin, 2015) (Fig. 2.1). 380 images at x20 and x40 objectives were annotated, of which 20% was used for testing and 80% for training.

Post software development, the accuracy of the software and experimental purpose was tested using the *O. glaberrima* population. Testing the accuracy of the software was completed on a subset of microscopy images at x20 and x40 microscope objectives. See Supplementary Figure S2.2 for images of the Stomata Detector software.

(a)



(b)

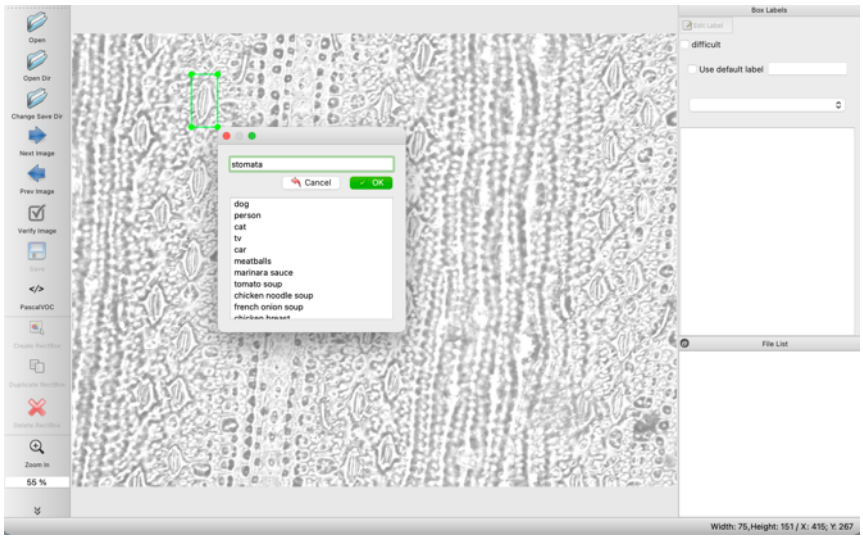


Figure 2.1: Image annotation using the graphical image annotation tool, Labellmg. Microscopy images of the rice species *O. glaberrima* and *O. sativa* were annotated at (a) x20 and (b) x40 objectives using Labellmg software.

2.3.4 Transfer learning

Our method is based on transfer learning for deep neural networks. That is, we have utilised a pre-trained deep model for the different datasets and adapt it for our stomata samples. Here, we briefly review the transfer learning technique and provide the details of the used model in our experiment.

Transfer learning is a machine learning technique in which a learner learns a new task using the experiences of another task but related to the new task. It is also a popular approach in deep learning as it helps to significantly reduce the training time and complexity of designing a deep model resulting in a well-trained model for the target task with good generalisation performance. Transfer learning is used extensively in computer vision and natural language processing. For instance, if the target task is the classification of the CIFAR-10 images, it is better to use a pre-trained model on a very large dataset such as ImageNet as a starting point and adapt this pre-trained model for the new task (e.g., classifying of CIFAR-10 images) by only training a small part of the model. In this way, not only can we use the features learned by a related task, but also, we do not need to design and start training parameters of a model from scratch. This is indeed one of the important features of modern deep learning frameworks such as TensorFlow and PyTorch which provide a rich class of pre-trained models as a built-in API. This helps practitioners and researchers to quickly train an existing model for their own setup to test the performance.

2.3.5 Object Detection Model

Based on the transfer learning approach, we utilise a pre-trained object detection model trained on the standard COCO datasets (Lin et al., 2014). This dataset is a collection of more than 330k images with 80 object categories for large-scale object detection, segmentation, and captioning tasks. Since our goal is detecting and classifying stomata, we use the Faster R-CNN model (Ren *et al.*, 2015) as one of the state-of-the-art methods based on deep neural networks. In particular, we downloaded a pre-trained Faster R-CNN model available in Tensorflow with the Inception-V2 architecture (Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016) as the base model. Inception-V2 is a variation of Inception-V1 also referred to as GoogLeNet was the state-of-the-art architecture at ImageNet competition in ILSRVRC 2014. After loading the pre-trained Faster R-CNN, the last few layers of classification layers are changed to meet the aim of stomata classification and detection. In the next step, the Faster R-CNN with stomata images are trained with different hyper parameters such as learning rate and number of epochs to find out the best parameters to reduce execution time and errors.

2.3.6 Stomata Detector software functionality

A user-friendly interface was created (see Supplementary Figure S2.2 for screenshots) to facilitate automated stomatal detection for end users that are not literate in handling and

running code. Stomata detector requires a threshold value to be set before analysing stomatal micrographs. This is the value the requires an inputted value of 0-1, this is the estimated certainty that the algorithm has correctly identified an object in the micrograph as a stoma (1 = 100% certainty). The user can then decide to load and execute a singular or file of many micrographs. The main interface shows a list of loaded micrographs on the right-hand panel, a summary of the historical and current analyses in the central window. A window is also generated, showing the code in the backend of the software, this can be useful when monitoring the progress when batch processing a file of micrographs. When an analysis is complete, a .csv output file is generated with columns detailing; the micrograph file path name, the chosen threshold value, number of detected stomata, number of stomata not included if they did not pass the defined threshold value, mean and standard deviation.

The micrographs obtained from the *O. glaberrima* accessions and *O. sativa* IR64 were processed through the Stomata Detector software at a threshold value of 0.9. The number of stomata detected in each micrograph was scaled up to give an estimate of SD per mm^2, based on the area of the micrograph (x20 = 0.3 mm^2 and x40 = 0.08 mm^2). The median SD for each genotype was then calculated. All statistical analysis was completed in R-Studio (v. 4.0.5), graphs were generated using the R package ggplot2 (v. 3.3.3).

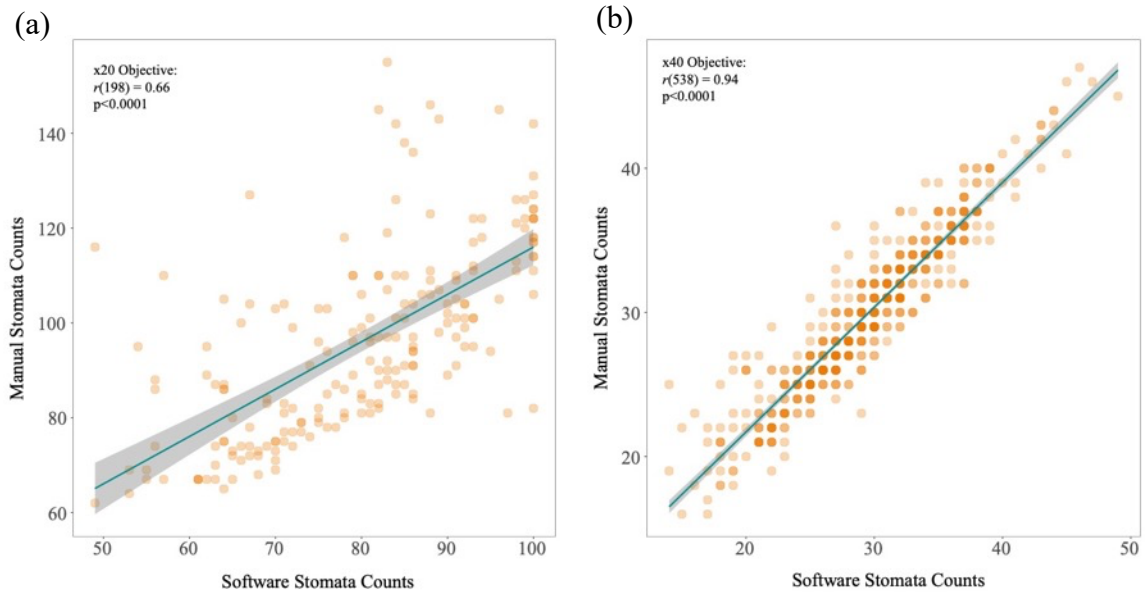
2.4 RESULTS

2.4.1 Stomata Detector can rapidly batch process micrographs

The Stomata Detector software can batch process a folder and singular micrographs. When a single micrograph is analysed by the software, the output includes both a .csv file of results and an image of the original micrograph overlaid with bounding boxes over objects that are identified as stomata (Fig. 2.2c - d and Supp. Fig. S2.2c). When sequentially batch processing a file of images, a .csv file of the results is still produced but to save on processing time and storage capacity, the image overlaid with bounding boxes is not. The software can batch process high numbers of images in a comparably short period of time, for example a folder of all 13,110 stomatal impression micrographs was completed within 24 hours, when left to run for this period. Single images, with accompanied image outputs, are processed almost instantly.

2.4.2 Stomatal detection accuracy

The software was tested for accuracy on images of rice (*O. glaberrima* and *O. sativa*) stomata at x20 and x40 objective, to establish the accuracy at each magnification. The correlation between the automated Stomata Detector and manual counting methods were used as one indication of software accuracy. At x20 magnification a Pearson correlation test showed a significant association between manual and automated stomatal counts (n = 200, $r = 0.66$, $p<0.0001$; Fig. 2.2a and c), with an overall similarity of 83% between the two stomatal counting methods. While a higher correlation was observed at a x40 objective (n = 540, $r = 0.94$, $p<0.0001$; Fig. 2.2b and d), with an overall similarity of 99% between the total sum of two stomatal counting methods. A total of 20 micrographs, 10 for each leaf side, were randomly selected from the pool of 13,110 *O. glaberrima* population micrographs. These were used to further check the accuracy of the automated analysis method in comparison to manual stomata counting (Fig. 2.3a-b and Table 2.1). Counts matched for 12/20 micrographs between the two methods, with an overall similarity of 99.8% for the total stomatal counts ($r = 0.94$, $p<0.0001$). The remainder 8/20 micrographs showed a small difference in stomata count between manual and automated stomatal counting methods (Table 2.1). Of these, stomata were missed in 3/20 and false positives identified in 5/20 micrographs. Examples of false positive and negative stomata identification from these images can be seen in Fig. 2.4 (a-b).



(c) (d)

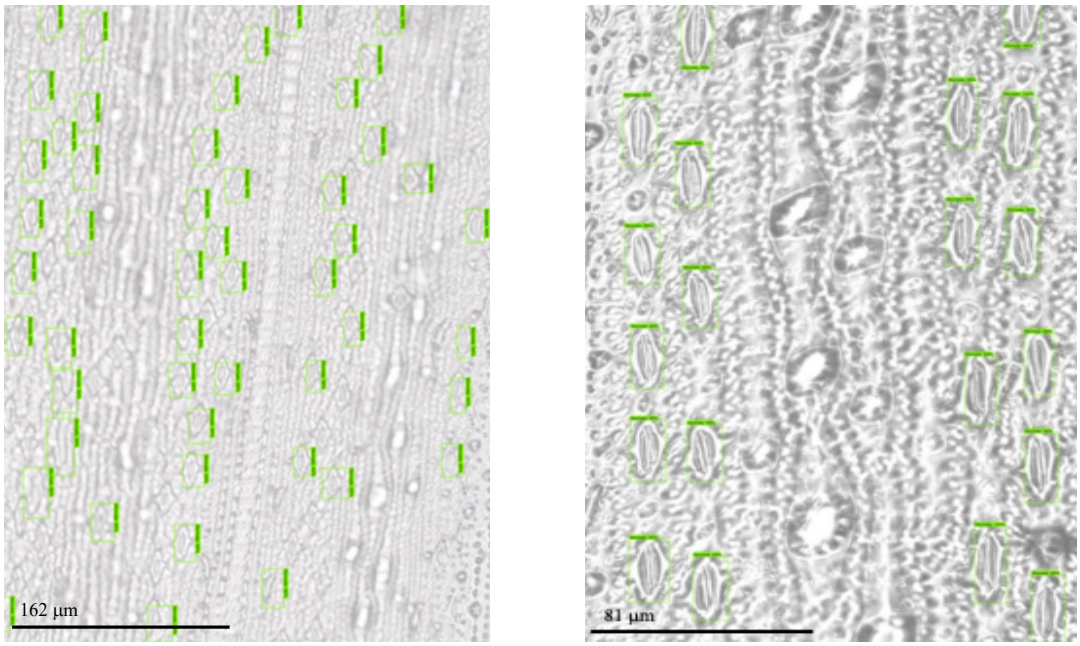
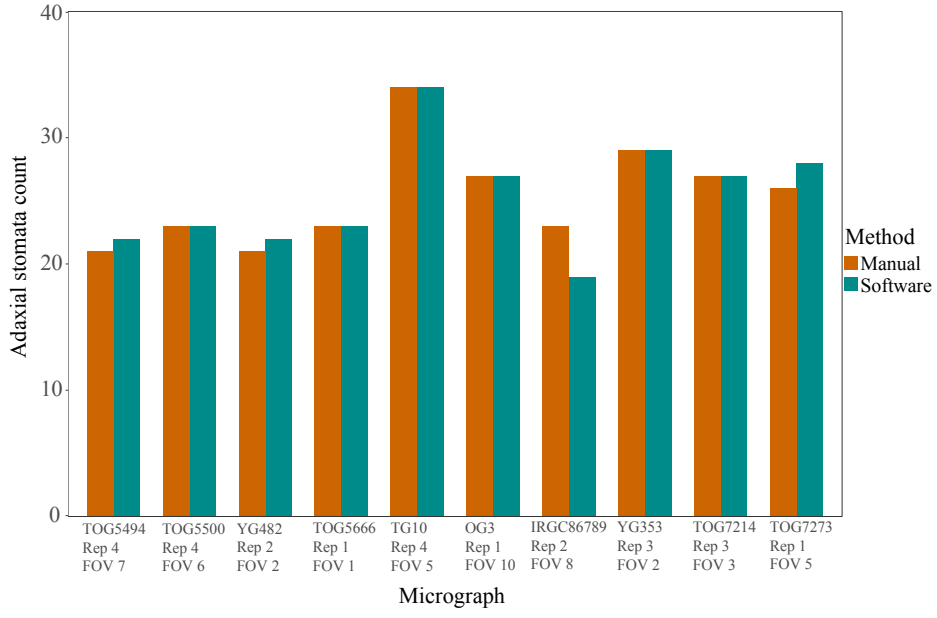


Figure 2.2: Stomata Detector identifies stomata better at higher magnifications. The stomata count obtained from the automated software method was compared to the manual counting method for x20 (a; $r = 0.66$, $p<0.0001$, $n = 198$) and (b; $r = 0.94$, $p<0.0001$, $n = 538$) x40 microscope objectives. The opacity is decreased for individual data points on Fig. 2.2a-b, therefore a greater density of colour indicates overlapping on data points. The blue line shows the line of best fit, banded by a grey 95% confidence interval. Images (c) and (d) show cropped Stomata Detector outputs, detected stomata are shown bound in green boxes.

(a)



(b)

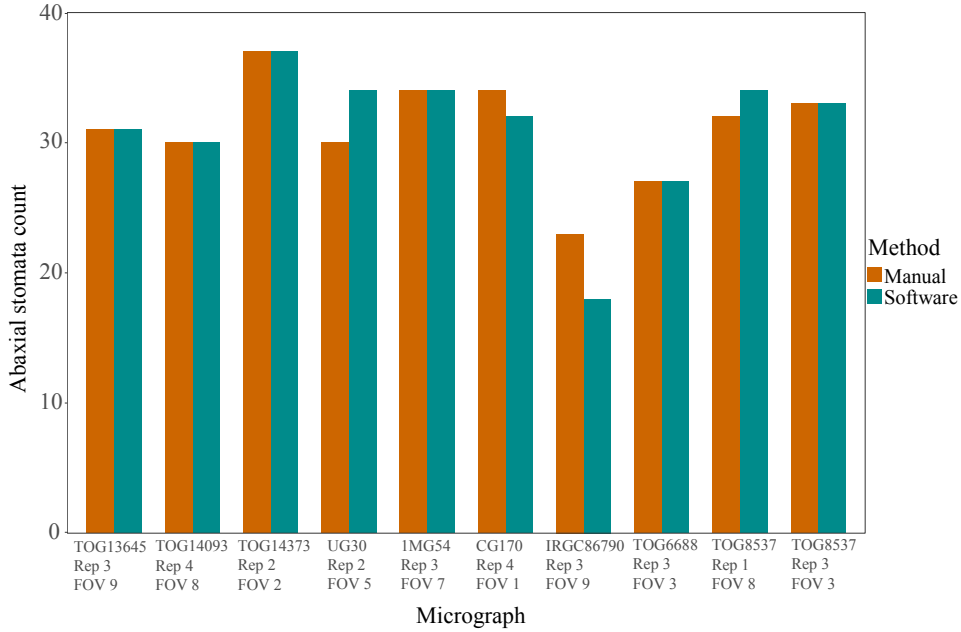


Figure 2.3: Bar charts showing the comparison of manual and automated software stomata counting methods from randomly selected micrographs. 20 outputs, taken from the analysis on 13,100 micrographs from a population of 155 *O. glaberrima* accessions, using the automated stomata counting method were randomly selected ((a) 10 adaxial, (b) 10 adaxial)**.** The micrographs of the randomly selected outputs were manually counted to compare the accuracy between the two methods.

Table 2.1: Randomly selected micrographs showing a difference in stomata count between automated and manual methods. The table shows the micrograph identifying name (accession name, biological replicate (rep) and field of view (FOV)), the stomata count for the automated and manual stomata counting methods, followed by the difference between these two methods.

Accession	Rep	FOV	Leaf side	Automated	Manual	Difference
IRGC_86790	3	9	Abaxial	18	23	5
TOG_8537	1	8		34	32	2
UG30	2	5		34	30	4
CG170	4	1		32	34	2
TOG_5494	4	7	Adaxial	22	21	1
IRGC_86789	2	8		19	23	4
TOG_7273	1	5		28	26	2
YG482	2	2		22	21	1

(a) (b)

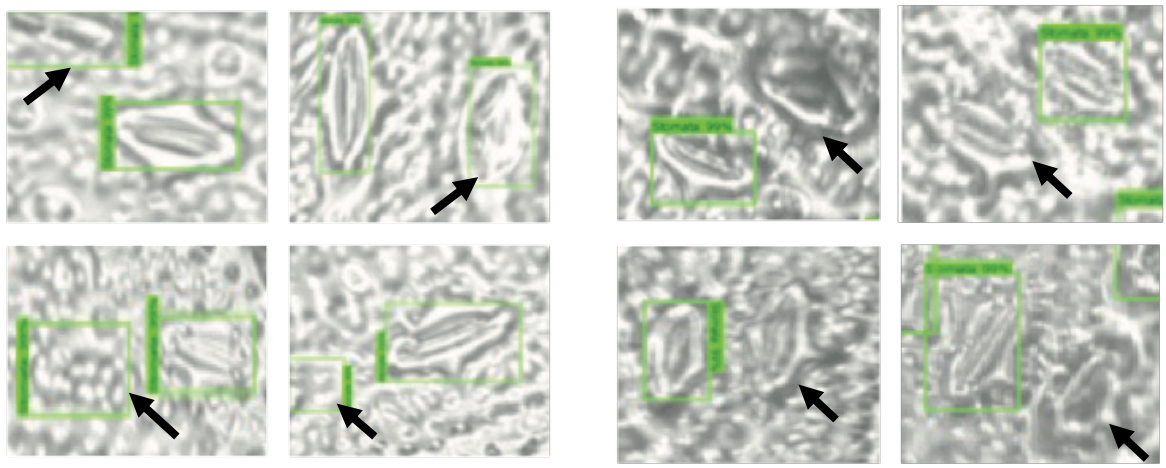


Figure 2.4: Samples showing errors in stomatal detection. False positive (a) identification occurs when features in the micrograph are erroneously detected as stomata. While false negatives (b) are stomata detected by a human during manual counting but not by the automated method. Black arrows indicate a false positive and false negative occurrence.

2.4.3 Quantifying variation in a population of 155 individuals

As proof of concept, we used Stomata Detector to process 13,100 micrographs, collected from a population of 155 *O. glaberrima* accessions and the *O. sativa* cultivar, IR64. This population was previously uncharacterised for stomatal traits. While the micrographs were obtained by hand, the automated analysis was completed within 24 hours. The stomatal density (per mm^2) showed a normal trait distribution, with no spurious outliers (Fig. 2.5). The stomatal density values for the population (abaxial = 377 mm^{-2}, adaxial = 305 mm^{-2}) are in line with that reported by Chatterjee *et al.* (2020) for *O. glaberrima* accession IRGC_103544. Supplementary Figure S2.3 shows boxplots of the stomatal density for each accession, calculated from the outputs of Stomata Detector.

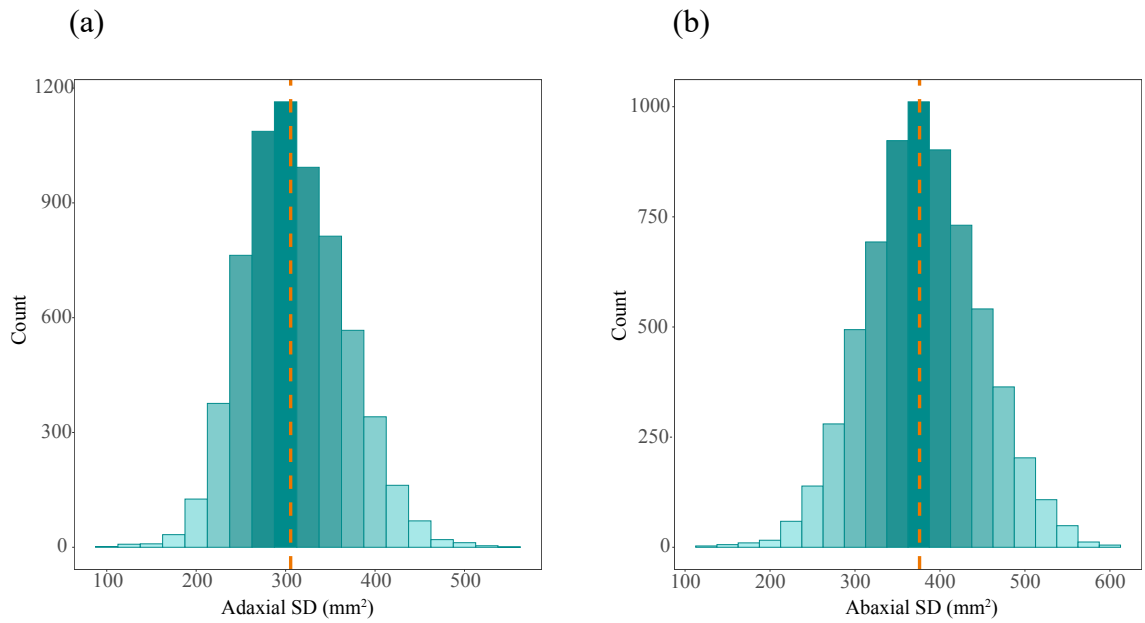


Figure 2.5: Trait distribution of (a) adaxial and (b) abaxial stomatal density for the *O. glaberrima* population, analysed by the automated method. Orange dashed line shows the mean value in the population (n=155).

2.5 DISCUSSION

We have a better understanding than ever that stomatal density and morphology is a critical functional trait for optimising carbon assimilation and minimising abiotic stress (Caine *et al.*, 2019; Faralli, Matthews and Lawson, 2019). As commercial crop species are genetically narrow (Meyer & Purugganan, 2013) and poorly adapted to challenging environmental conditions, the focus is upon underutilised crop species and wild relatives as a source of stomatal genetic diversity (Meyer & Purugganan, 2013). However, manual stomatal phenotyping is slow and thus limits the characterisation of stomatal genetic diversity within populations. The automatic machine-learning based software (Stomata Detector) developed here analyses the stomatal number from micrographs, that would otherwise not be feasible when manually phenotyping a large population.

A prime example of this is the *O. glaberrima* population described here, which was collected as a resource for *O. sativa* crop improvement (Cubry et al., 2020). However, the stomatal diversity in this population remains uncharacterised, which limits the elucidation of the underlying trait genetic diversity. Using the automated counting method developed here, we have been able to characterise the stomatal density in a population of *O. glaberrima*

accessions. This will have a direct impact when exploring the role of stomata on *O. glaberrima* physiology, environmental resilience and elucidate trait-related gene loci.

We have shown the Stomata Detector method can accurately identify stomata, though errors can occur. False positive stomatal identification may happen when there is a structure in the micrograph that is a similar size and shape to a stoma (Fig. 2.3a). Observed false positives included air bubbles and the contours of an epithelial cell, particularly when on the edge of a micrograph. False negatives, that being stomata not detected by the method, typically occur when images are blurry, have low contrast between epithelial and stoma subsidiary cells or very small stomata that can be present on leaf veins. When testing this method, we compromised on image quality in favour of high throughput processing, though this may have caused occurrences of false positive and negatives. Ensuring high quality sample preparation and obtaining micrographs with a greater pixel resolution will minimise these errors, though there will be an increase in processing time. Despite this guidance, perfect stomatal peels are not always possible, and the end user requires a method that is able to accurately identify stomata in a range of peel and micrograph qualities. This can be improved by increasing the sample size, and variability of image quality, used in the training data set. The importance of choosing the optimal magnification for stomatal identification was apparent in Fig. 2.2a-b. This shows that Stomata Detector performed accurately at x40 objective and comparatively poorly at x20, relative to human manual analysis. It is worth noting, the comparison of this automated method to human manual counting (Fig. 2.2a-b) does not account for human error, which may be substantial when analysing hundreds of micrographs and is difficult to quantify. For example, the 20 randomly selected micrographs (Fig. 2.3; Table 2.1) was manually checked three times to ensure accuracy and reduce human error. 5 out of 20 images stomata were mis-counted on the first repetition of manual analysis. Stomata Detector was more accurate at a higher magnification due to the enhanced clarity and size of the leaf epidermis and stomata. Stomata are particularly small in *Oryza spp*., in comparison to other species. At the x20 objective, the leaf epidermal surface area covered is greater but means that the varnish peel across that area undulates more than at a higher magnification, causing many stomata to be out of focus and unidentifiable. Therefore, it is advised that end users dedicate thought and time to establishing the best microscope magnification, image size and quality for the species and peel quality at hand.

The Stomata Detector training set was developed using only a small number of training images, although the software was still found to identify *O. glaberrima* stomata relatively accurately (Fig. 2.2b). Though we will highlight this study did not fully consider the degree of human error, which we recommend quantifying in future use. As the algorithm was trained on a small dataset originating from one experiment, we would advise a degree of caution and quality checking to future users. Ideally, we would like the algorithm to be re-trained on a larger dataset before an open release to the scientific community. Currently our method is enriched with images of *O. glaberrima* and *O. sativa* and does not yet reliably identify other species. There is currently no other resource that accurately identifies *Oryza spp.* stomata, other methods (Fetter et al., 2019) focuses on angiosperm taxa, with large dumbbell shaped stomata. We feel the method demonstrated here will be an asset to the scientific community with training on a larger dataset, especially considering the importance of rice crop improvement as a staple food for billions of people. Though our aim would be to develop a resource to that enables taxonomically diverse stomatal identification. This would require the ongoing retraining of the method from future users.

Our initial work has developed an image processing software which allows accurate stomatal counting of 1000s of stomatal impression in a few days, rather than several months. We would like to extend this to a mobile application, which could be used in field in conjunction with a mobile phone microscope (Orth, Wilson, Thompson, & Gibson, 2018), enabling non-destructive real-time measurements. This would significantly alleviate the current stomata phenotyping bottleneck and alter the possibilities in trait characterisation, even extending to in field analysis of maximal stomatal conductance (Brown and Escombe, 1900; Dow, Bergmann and Berry, 2014). This would contribute to a major step forward in stomatal, and even wider leaf-level phenotyping.

If we are to tackle food insecurity and future sustainability, the integration of differing scientific techniques and fields is essential. Our project developed an open source, sophisticated piece of software that can accurately identify stomata in Oryza, and potentially other monocot, species. Demonstrating that deep machine learning is the ideal partner for narrowing the bottleneck of high-quality plant phenomics and consequently enabling crop improvement.

2.6 CHAPTER TWO SUPPLEMENTARY MATERIAL

2.6.1 Supplementary Tables

Table S2.1: List of *O. glaberrima* accession codes, ecology, country and African region of origin. O. glaberrima germplasm was provided by Diversité Adaptation Developpement des plantes (DIADE), IRD-Montpellier, France. Information on accession ecology and country of origin was provided by AfricaRice. The list of accessions is ordered based on the ecological environment at origin.

Accession code	Ecology	Country of origin	African region
RAM 131	Floating Rice	Mali	West inland
RAM 137	Floating Rice	Mali	West inland
RAM 24	Floating Rice	Guinea	West coast
RAM 48	Floating Rice	Mali	West inland
RAM 55	Floating Rice	Mali	West inland
RAM 77	Floating Rice	Mali	West inland
IRGC_96726	Irrigated lowland	Nigeria	West coast
TOG_5314	Irrigated lowland	Nigeria	West coast
IRGC_96740	Irrigated lowland	Nigeria	West coast
IRGC_86764	Irrigated lowland	Ghana	West coast
IRGC_96790	Irrigated lowland	Nigeria	West coast
IRGC_56785	Irrigated lowland	Liberia	West coast
IRGC_112568	Irrigated lowland	Liberia	West coast
IRGC_86789	Irrigated lowland	Liberia	West coast
IRGC_86790	Irrigated lowland	Liberia	West coast
IRGC_86791	Irrigated lowland	Liberia	West coast
TOG_5969	Irrigated lowland	Nigeria	West coast
TOG_6205	Irrigated lowland	Guinea	West coast
TOG_6206	Irrigated lowland	Zimbabwe	South inland
TOG_6207	Irrigated lowland	Zimbabwe	South inland
TOG_6211	Irrigated lowland	Nigeria	West coast
TOG_6220	Irrigated lowland	Burkina Faso	West inland
TOG_6943	Irrigated lowland	Sierra Leone	West coast

TOG_6951	Irrigated lowland	Sierra Leone	West coast
TOG_7020	Irrigated lowland	Sierra Leone	West coast
TOG_7047	Irrigated lowland	Sierra Leone	West coast
TOG_7106	Irrigated lowland	Mali	West inland
TOG_7108	Irrigated lowland	Mali	West inland
CG10	Irrigated lowland	Senegal	West coast
TOG_7132	Irrigated lowland	Senegal	West coast
TOG_7134	Irrigated lowland	Senegal	West coast
IRGC_103544	Irrigated lowland	Mali	West inland
CG14	Irrigated lowland	Senegal	West coast
TOG_14367	Irrigated lowland	Guinea	West coast
TOG_7214	Irrigated lowland	Upland	West inland
CG171	Irrigated lowland	Senegal	West coast
TOG_7219	Irrigated lowland	Mali	West inland
IRGC_103549	Irrigated lowland	Mali	West inland
TOG_10434	Irrigated lowland	Côte d'Ivoire	West coast
TOG_7255	Irrigated lowland	Chad	North inland
TOG_7273	Irrigated lowland	Cameroon	West coast
IRGC_104589	Irrigated lowland	Burkina Faso	West inland
IRGC_86826	Irrigated lowland	Ghana	West coast
TOG_7451	Irrigated lowland	Burkina Faso	West inland
TOG_7455	Irrigated lowland	Burkina Faso	West inland
TOG_7456	Irrigated lowland	Burkina Faso	West inland
TOG_7455	Irrigated lowland	Côte d'Ivoire	West coast
TOG_7993	Irrigated lowland	Nigeria	West coast
TOG_8049	Irrigated lowland	Nigeria	West coast
TOG_8527	Irrigated lowland	Gambia	West coast
TOG_8537	Irrigated lowland	Gambia	West coast
TOG_8545	Irrigated lowland	Gambia	West coast
TOG_9524	Irrigated lowland	Côte d'Ivoire	West coast
TOG_13645	Irrigated lowland	Guinea	West coast
TOG_13708	Irrigated lowland	Guinea	West coast
TOG_14093	Irrigated lowland	Guinea	West coast

TOG_14373	Irrigated lowland	Guinea	West coast
TOG_14606	Irrigated lowland	Guinea	West coast
TOG_14610	Irrigated lowland	Guinea	West coast
TOG_7190	Irrigated lowland	Côte d'Ivoire	West coast
TG10	Irrigated lowland	Chad	North inland
TG19_G	Irrigated lowland	Chad	North inland
TG25	Irrigated lowland	Chad	North inland
TG57	Irrigated lowland	Chad	North inland
CG45	Irrigated lowland	Senegal	West coast
CG46	Irrigated lowland	Senegal	West coast
CG70	Irrigated lowland	Senegal	West coast
CG150	Irrigated lowland	Senegal	West coast
CG156	Irrigated lowland	Senegal	West coast
CG164	Irrigated lowland	Senegal	West coast
CG170	Irrigated lowland	Senegal	West coast
TOG_5418	Lowland	Nigeria	West coast
TOG_5400	Lowland	Nigeria	West coast
LG33	Lowland	Mali	West inland
TOG_5533	Lowland	Nigeria	West coast
MG04	Lowland	Mali	West inland
EG55	Lowland	Tanzania	East coast
EG85	Lowland	Tanzania	East coast
UG14	Lowland	Cameroon	West coast
UG20	Lowland	Cameroon	West coast
UG26	Lowland	Cameroon	West coast
LG07_S	Lowland	Mali	West inland
LG64	Lowland	Mali	West inland
MG53	Lowland	Mali	West inland
1MG54	Lowland	Mali	West inland
OG1	lowland	Senegal	West coast
OG3	lowland	Senegal	West coast
OG15	lowland	Senegal	West coast
TOG_5321	Rainfed lowland	Nigeria	West coast

TOG_5424	Rainfed lowland	Nigeria	West coast
TOG_5486	Rainfed lowland	Nigeria	West coast
TOG_5494	Rainfed lowland	Nigeria	West coast
TOG_5500	Rainfed lowland	Nigeria	West coast
TOG_5556	Rainfed lowland	Nigeria	West coast
TOG_5666	Rainfed lowland	Nigeria	West coast
TOG_5672	Rainfed lowland	Nigeria	West coast
TOG_5681	Rainfed lowland	Nigeria	West coast
TOG_5814	Rainfed lowland	Liberia	West coast
TOG_5882	Rainfed lowland	Nigeria	West coast
TOG_5953	Rainfed lowland	Nigeria	West coast
TOG_6356	Rainfed lowland	Liberia	West coast
TOG_6603	Rainfed lowland	Liberia	West coast
TOG_6688	Rainfed lowland	Liberia	West coast
TOG_6698	Rainfed lowland	Liberia	West coast
TOG_5286	Rainfed lowland	Nigeria	West coast
TOG_5439	Rainfed lowland	Nigeria	West coast
TOG_5464	Rainfed lowland	Nigeria	West coast
TOG_5566	Rainfed lowland	Nigeria	West coast
TOG_5591	Rainfed lowland	Ghana	West coast
TOG_5639	Rainfed lowland	Nigeria	West coast
TOG_5747	Rainfed lowland	Liberia	West coast
TOG_5775	Rainfed lowland	Liberia	West coast
TOG_7420	Rainfed lowland	Sierra Leone	West coast
IG38	Rainfed lowland	Côte d'Ivoire	West coast
YG353	Rainfed lowland	Guinea	West coast
TOG_12086	Rainfed lowland	Nigeria	West coast
TOG_12160	Rainfed lowland	Nigeria	West coast
TOG_12188	Rainfed lowland	Nigeria	West coast
TOG_12249	Rainfed lowland	Nigeria	West coast
TOG_7406	Rainfed lowland	Ghana	West coast
TOG_12366	Rainfed lowland	Guinea-Bissau	West coast
TOG_12372	Rainfed lowland	Guinea-Bissau	West coast

TOG_12387	Rainfed lowland	Tanzania	East coast
TOG_12388	Rainfed lowland	Cameroon	West coast
YG330	Rainfed lowland	Guinea	West coast
TOG_14116	Rainfed lowland	Liberia	West coast
TOG_14184	Rainfed lowland	Zimbabwe	South inland
YG482	Rainfed lowland	Guinea	West coast
TOG_14361	Rainfed lowland	Guinea	West coast
IG05	Rainfed lowland	Côte d'Ivoire	West coast
IG09	Rainfed lowland	Côte d'Ivoire	West coast
IG14	Rainfed lowland	Côte d'Ivoire	West coast
IG15	Rainfed lowland	Côte d'Ivoire	West coast
IG16	Rainfed lowland	Côte d'Ivoire	West coast
IG19	Rainfed lowland	Côte d'Ivoire	West coast
IG21	Rainfed lowland	Côte d'Ivoire	West coast
IG23	Rainfed lowland	Côte d'Ivoire	West coast
IG35	Rainfed lowland	Côte d'Ivoire	West coast
IG43	Rainfed lowland	Côte d'Ivoire	West coast
IG47	Rainfed lowland	Côte d'Ivoire	West coast
IG324	Rainfed lowland	Côte d'Ivoire	West coast
UG28	Rainfed lowland	Cameroon	West coast
UG30	Rainfed lowland	Cameroon	West coast
YG316	Rainfed lowland	Guinea	West coast
TOG_5326	Shallow forest swamp	Nigeria	West coast
TOG_5453	Shallow forest swamp	Nigeria	West coast
TOG_7274	Shallow Forest Swamp	Cameroon	West coast
TOG_5997	Upland	Nigeria	West coast
TOG_12358	Upland	Côte d'Ivoire	West coast
TOG_12399	Upland	Guinea	West coast
TOG_12401	Upland	Guinea	West coast
TOG_12411	Upland	Guinea	West coast

TOG_12414	Upland	Guinea	West coast
IG36	Upland	Côte d'Ivoire	West coast
YG307	Upland	Guinea	West coast

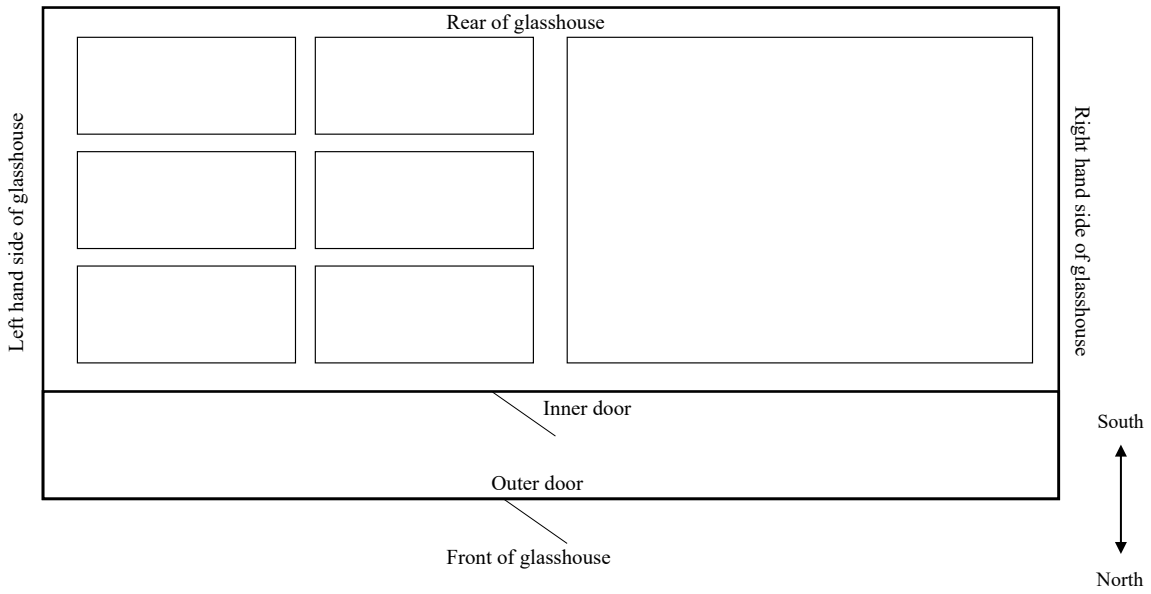
Table S2.2: Table detailing sowing and measurement dates of one hundred and fifty-five *O. glaberrima* accessions. This table details how the *O. glaberrima* accessions were grouped into rotations of twelve, each accession name was assigned a Nottingham line number for ease. The columns show the sowing date, date of the week that measurements commenced (Measurement start date) and date of the week that measurements ceased (Measurement end date) for each rotation of 12 accessions.

***O. glaberrima* line no.**	**Sowing date**	**Measurement start date**	**Measurement end date**
1 - 12	26/04/2017	12/06/2017	16/06/2017
13-24	08/05/2017	26/06/2017	30/06/2017
25-36	22/05/2017	10/07/2017	14/07/2017
37-48	05/06/2017	24/07/2017	28/07/2017
49-60	19/06/2017	07/08/2017	11/08/2017
61-72	03/07/2017	21/08/2017	25/08/2017
73-84	17/07/2017	04/09/2017	08/09/2017
121-132	24/07/2017	11/09/2017	14/09/2017
85-69	31/07/2017	18/09/2017	22/09/2017
133-144	07/08/2017	25/09/2017	28/09/2017
97-108	14/08/2014	02/10/2017	06/10/2017
145-156	21/08/2017	09/10/2017	12/10/2017
109-120	28/08/2017	16/10/2017	20/10/2017

2.6.2 Supplementary Figures

Figure S2.1: *O. glaberrima* phenotyping experimental plan. (a) The left-hand side of the agronomy style glass house was divided into six small beds, three of these were used for the first six batches of measurements, following these all six beds were used until all accessions had been measured. (b) In each small bed twelve accessions and IR64, with five replicates, were grown. The accessions were surrounded by a border of spare seedlings, to minimise for edge effects. Accessions were not randomised to minimise plant damage and mismeasurement during LiCor analysis. Through this process, (c) a batch rotation through the available beds facilitated the growth and analysis of all accessions.

(a)



(b) (c)

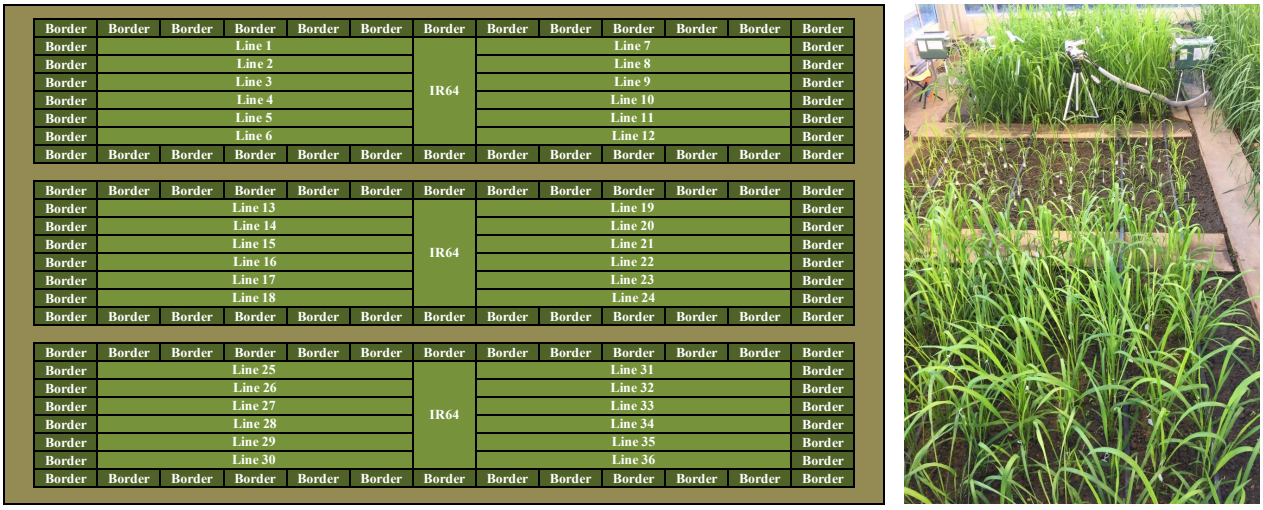
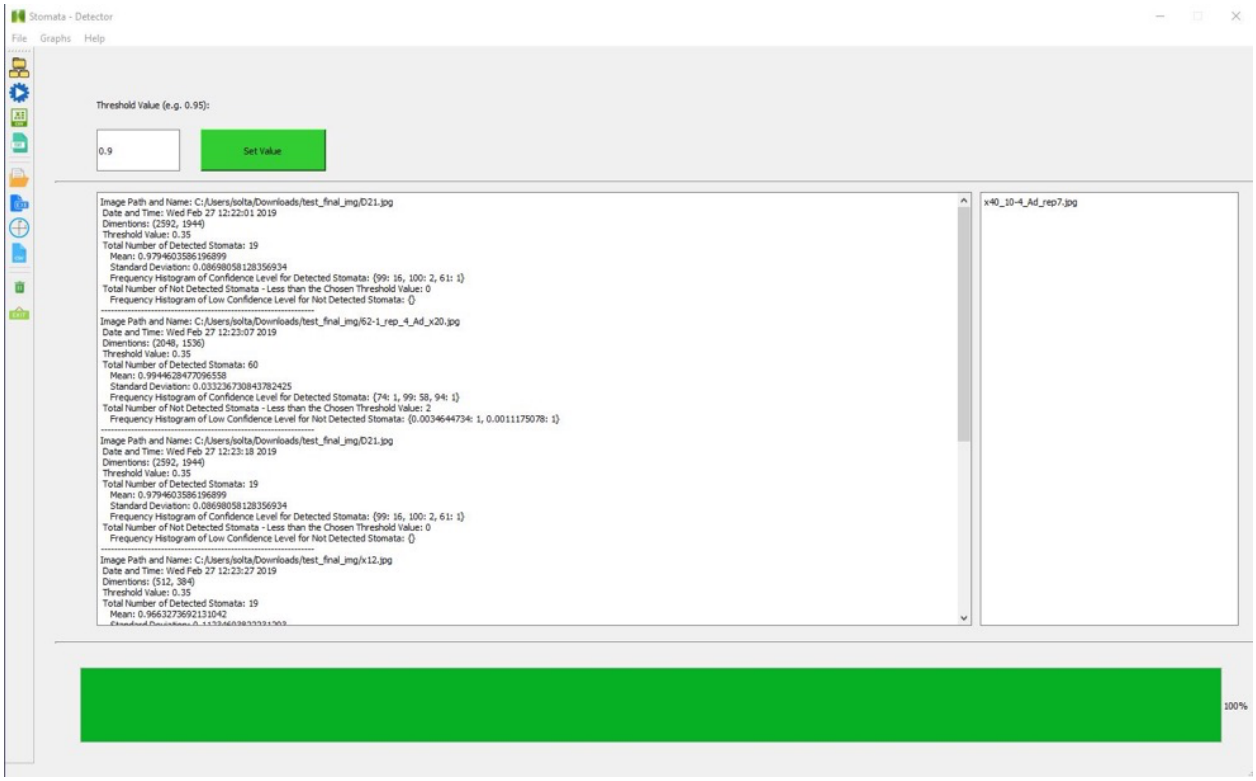
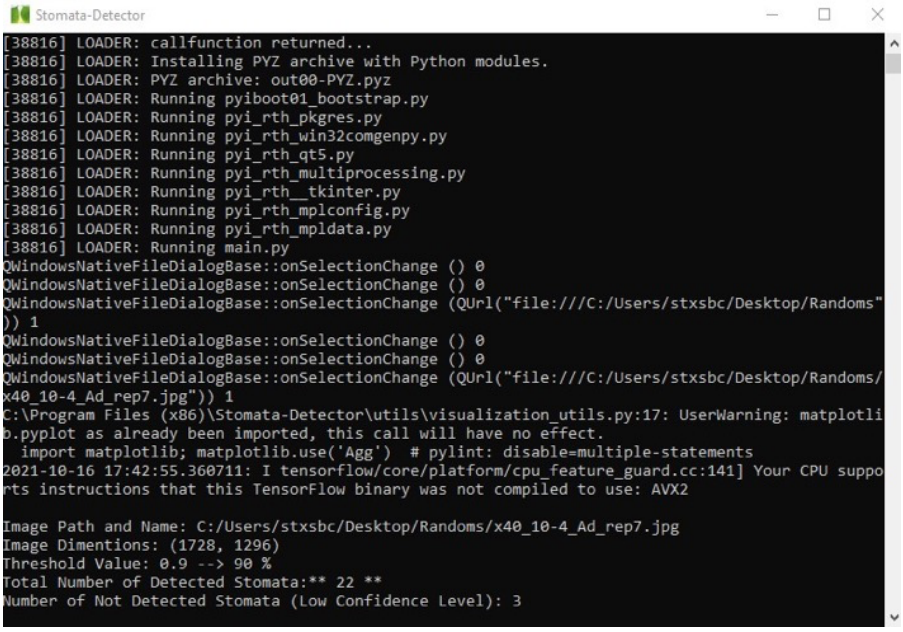


Figure S2.2: Images of Stomata Detector software. Screenshots taken of the Stomata Detector software, showing (a) the graphical user interface, where files and folders are loaded and executed; (b) the back end, which gives a summary of which micrographs are running and the outputs; (c) image outputs from a single micrograph analysis, showing the objects identified as stomata in green bounding boxes, the total number of stomata identified and the micrograph file pathway.

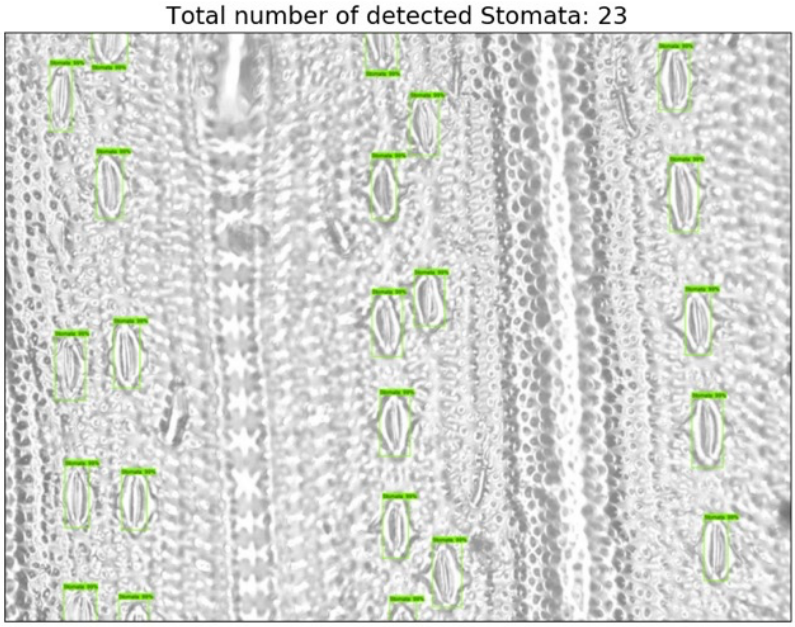
(a)


Stomata - Detector
File Graphs Help
Threshold Value (e.g. 0.95):
0.9
Set Value
Image Path and Name: C:/Users/solta/Downloads/test_final_img/D21.jpg
Date and Time: Wed Feb 27 12:22:01 2019
Dimentions: (2592, 1944)
Threshold Value: 0.35
Total Number of Detected Stomata: 19
Mean: 0.9794603586196899
Standard Deviation: 0.08698058128356934
Frequency Histogram of Confidence Level for Detected Stomata: {99: 16, 100: 2, 61: 1}
Total Number of Not Detected Stomata - Less than the Chosen Threshold Value: 0
Frequency Histogram of Low Confidence Level for Not Detected Stomata: {}
Image Path and Name: C:/Users/solta/Downloads/test_final_img/62-1_rep_4_Ad_x20.jpg
Date and Time: Wed Feb 27 12:23:07 2019
Dimentions: (2048, 1536)
Threshold Value: 0.35
Total Number of Detected Stomata: 60
Mean: 0.9944628477096558
Standard Deviation: 0.033236730843782425
Frequency Histogram of Confidence Level for Detected Stomata: {74: 1, 99: 58, 94: 1}
Total Number of Not Detected Stomata - Less than the Chosen Threshold Value: 2
Frequency Histogram of Low Confidence Level for Not Detected Stomata: {0.0034644734: 1, 0.0011175078: 1}
Image Path and Name: C:/Users/solta/Downloads/test_final_img/D21.jpg
Date and Time: Wed Feb 27 12:23:18 2019
Dimentions: (2592, 1944)
Threshold Value: 0.35
Total Number of Detected Stomata: 19
Mean: 0.9794603586196899
Standard Deviation: 0.08698058128356934
Frequency Histogram of Confidence Level for Detected Stomata: {99: 16, 100: 2, 61: 1}
Total Number of Not Detected Stomata - Less than the Chosen Threshold Value: 0
Frequency Histogram of Low Confidence Level for Not Detected Stomata: {}
Image Path and Name: C:/Users/solta/Downloads/test_final_img/x12.jpg
Date and Time: Wed Feb 27 12:23:27 2019
Dimentions: (512, 384)
Threshold Value: 0.35
Total Number of Detected Stomata: 19
Mean: 0.9663273692131042
x40_10-4_Ad_rep7.jpg
100%

(b)


Stomata-Detector
[38816] LOADER: callfunction returned...
[38816] LOADER: Installing PYZ archive with Python modules.
[38816] LOADER: PYZ archive: out00-PYZ.pyz
[38816] LOADER: Running pyiboot01_bootstrap.py
[38816] LOADER: Running pyi_rth_pkgres.py
[38816] LOADER: Running pyi_rth_win32comgenpy.py
[38816] LOADER: Running pyi_rth_qt5.py
[38816] LOADER: Running pyi_rth_multiprocessing.py
[38816] LOADER: Running pyi_rth__tkinter.py
[38816] LOADER: Running pyi_rth_mplconfig.py
[38816] LOADER: Running pyi_rth_mpldata.py
[38816] LOADER: Running main.py
QWindowsNativeFileDialogBase::onSelectionChange () 0
QWindowsNativeFileDialogBase::onSelectionChange () 0
QWindowsNativeFileDialogBase::onSelectionChange (QUrl("file:///C:/Users/stxsbc/Desktop/Randoms"
)) 1
QWindowsNativeFileDialogBase::onSelectionChange () 0
QWindowsNativeFileDialogBase::onSelectionChange () 0
QWindowsNativeFileDialogBase::onSelectionChange (QUrl("file:///C:/Users/stxsbc/Desktop/Randoms/
x40_10-4_Ad_rep7.jpg")) 1
C:\Program Files (x86)\Stomata-Detector\utils\visualization_utils.py:17: UserWarning: matplotli
b.pyplot as already been imported, this call will have no effect.
import matplotlib; matplotlib.use('Agg') # pylint: disable=multiple-statements
2021-10-16 17:42:55.360711: I tensorflow/core/platform/cpu_feature_guard.cc:141] Your CPU suppo
rts instructions that this TensorFlow binary was not compiled to use: AVX2
Image Path and Name: C:/Users/stxsbc/Desktop/Randoms/x40_10-4_Ad_rep7.jpg
Image Dimentions: (1728, 1296)
Threshold Value: 0.9 --> 90 %
Total Number of Detected Stomata:** 22 **
Number of Not Detected Stomata (Low Confidence Level): 3

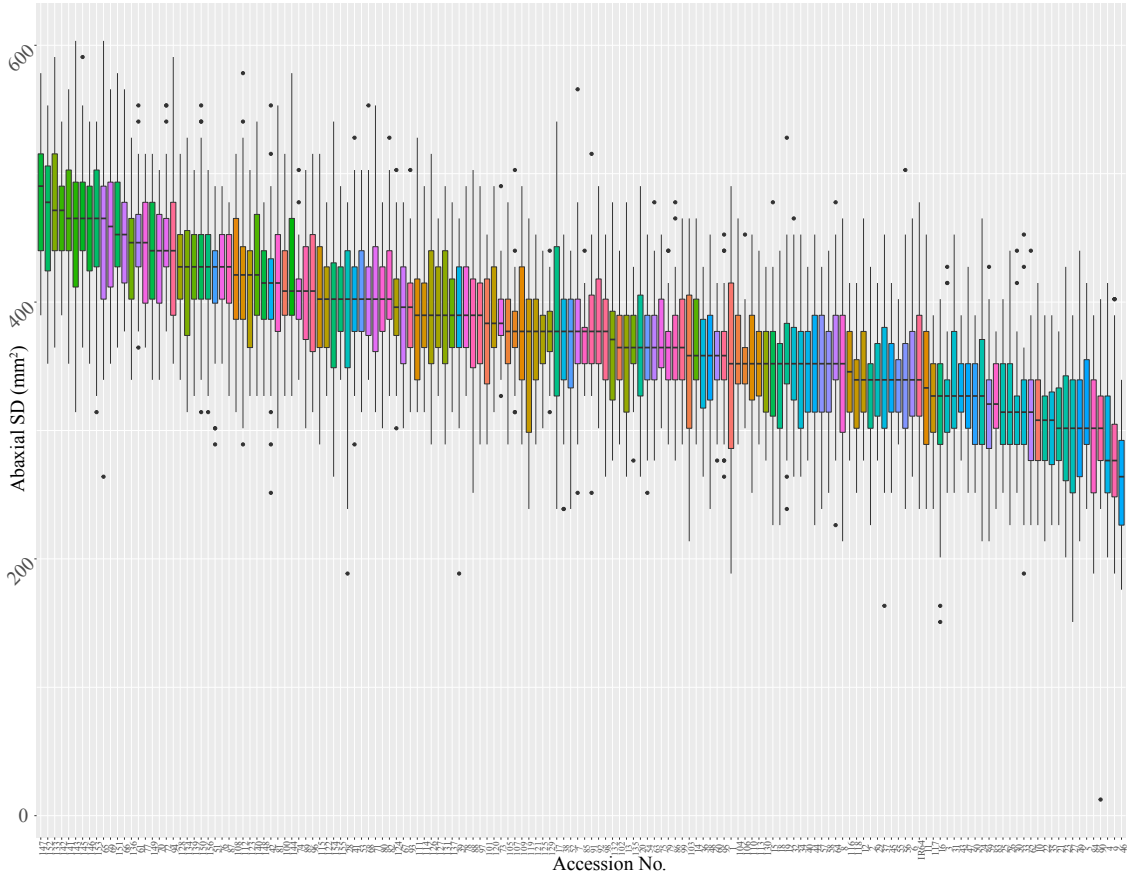
(c)



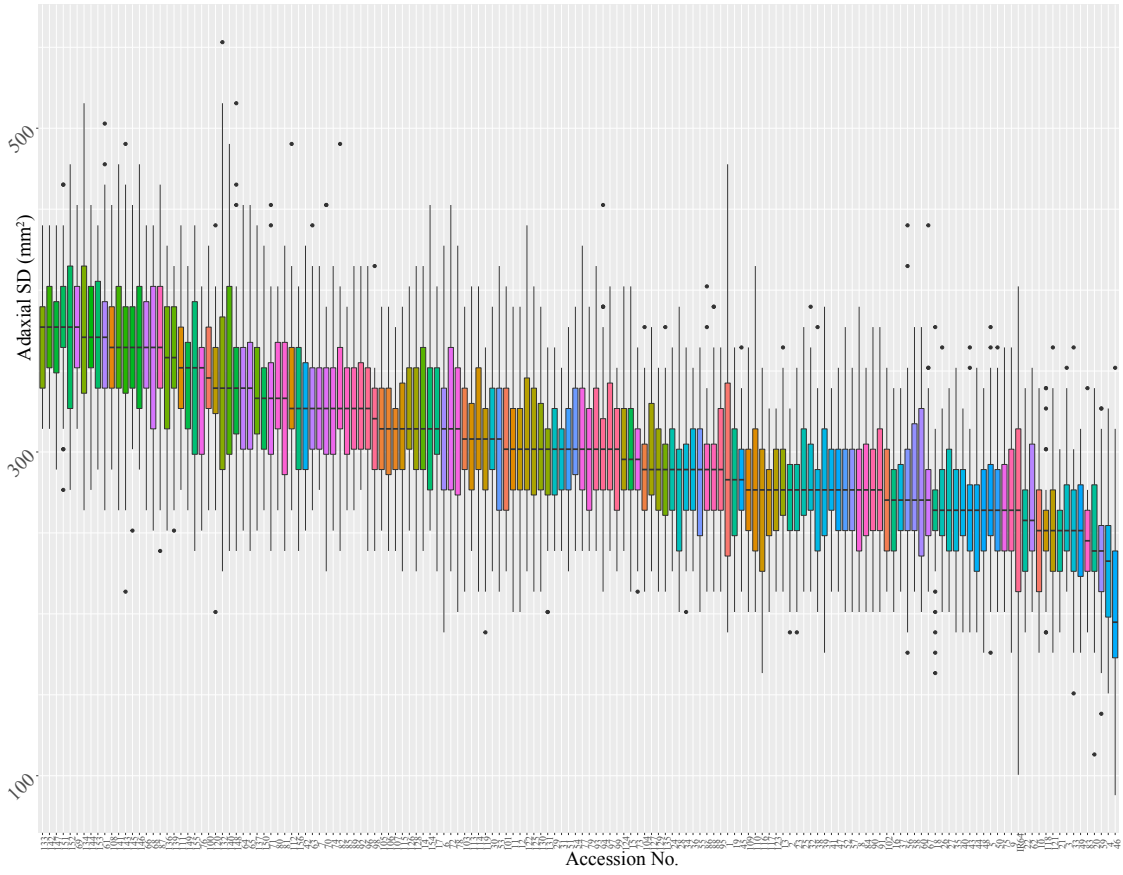
Stomata Path and Name: C:/Users/stxsbc/Desktop/Randoms/x40_11-4_Ad_rep6.jpg
Threshold value: 0.9

Figure S2.3: Boxplot of stomatal density for the *O. glaberrima* population. The boxplots show stomatal density calculated from the automated stomata counts generated by Stomata Detector, showing (a) abaxial and (b) adaxial stomatal density for the 155 *O. glaberrima* accessions and *O. sativa* cultivar, IR64. Boxplot shows the median, interquartile range (IQR) and the whiskers show the minimum (lower quartile – 1.5*IQR) and maximum (upper quartile – 1.5*IQR). Outliers are shown as black data points beyond the whiskers.

(a)



(b)



CHAPTER THREE

Out of Africa: characterising the natural variation in dynamic photosynthetic traits in a diverse population of African rice (Oryza glaberrima).

3. 1 Chapter three introductory statement

This chapter consists of a research publication that was accepted for publication by the Journal of Experimental Botany (doi.org/10.1093/jxb/erab459) in October 2021.

This manuscript focuses on describing the natural variation in 155 *O. glaberrima* accessions across 58 morphological and photosynthetic related traits that are important in abiotic stress tolerance and crop productivity. In conjunction to characterising the variation in the population, this information was used calculate trait heritability and percentage genetic variation to support future breeding efforts.

Recent research on improving water use efficiency and yield gains has highlighted the importance of the optimisation of dynamic traits contributing to increased photosynthetic efficiency. In this chapter I used the response curve data for carbon assimilation, stomatal conductance and non-photochemical quenching to model dynamic induction and relaxation responses due to changes in light levels. This is the first study to complete dynamic analysis for both induction and relaxation for CO_2 assimilation, stomatal conductance and non-photochemical quenching. It was also the largest published study of genotypes for the analysis of photosynthesis dynamic traits.

The key analytical phases of this paper compromise of;

1. Characterising natural variation in physiological and photosynthesis related traits and comparing this to the *O. sativa* cultivar, IR64.
2. Modelling carbon assimilation, non-photochemical quenching and stomatal conductance kinetics and how these interact with wider phenotypic traits
3. Exploring the effect of environmental variables and ecology on trait adaptation

A pdf copy of the publication can be found at the end of the chapter text.

ACCEPTED MANUSCRIPT

Out of Africa: characterising the natural variation in dynamic photosynthetic traits in a diverse population of African rice (*Oryza glaberrima*)

Sophie B Cowling, Pracha Treeintong, John Ferguson, Hamidreza Soltani, Ranjan Swarup, Sean Mayes, Erik H Murchie

Journal of Experimental Botany, erab459, https://doi.org/10.1093/jxb/erab459
Published: 17 October 2021 **Article history** ▾

Received: 06 September 2021
Published: 17 October 2021

Abstract

African rice (*Oryza glaberrima)* has adapted to challenging environments and is a promising source of genetic variation. We analysed dynamics of photosynthesis and morphology in a reference set of 155 *O. glaberrima* accessions. Plants were grown in an agronomy glasshouse to late tillering stage. Photosynthesis induction from darkness and the decrease in low light was measured by gas exchange and chlorophyll fluorescence along with root and shoot biomass, stomatal density and leaf area. Steady state and kinetic responses were modelled. We describe extensive natural variation in *O. glaberrima* for steady state, induction and reduction responses of photosynthesis that has value for gene discovery and crop improvement. Principle component analyses indicated key clusters of plant biomass, kinetics of photosynthesis (CO_2 assimilation, *A)* and photoprotection induction and reduction (measured by Non Photochemical Quenching, NPQ), consistent with diverse adaptation. Accessions also clustered according to countries with differing water availability, stomatal conductance (*gs)*, *A* and NPQ and indicating that dynamic photosynthesis has adaptive value in *O.glaberrima*. Kinetics of NPQ, *A* and *gs* showed high correlation with biomass and leaf area. We conclude that dynamic photosynthetic traits and NPQ are important within *O.glaberrima and* we highlight NPQ kinetics and NPQ under low light.

Keywords: *O. glaberrima*, African rice, photosynthesis, dynamic modelling, natural variation, rice, stomatal conductance, NPQ

Issue Section: Research Paper

3.2 INTRODUCTION

The climate crisis places crop yields under increasing pressure from biotic and abiotic constraints and constitutes a major threat in meeting global food demand (Ray *et al.*, 2019; Franco *et al*., 2020). Substantial yield decreases in key cereal crops are predicted to occur in both vulnerable and productive regions (Black *et al.*, 2008; Challinor *et al.*, 2014). Asian rice (*Oryza sativa*), a dietary staple to a third of the global population, is predicted to experience yield losses up to 37% by the end of the century due to climate change driven drought events (Bocco *et al.*, 2012; Muthayya et al., 2014; Zhao *et al.*, 2017). The development of productive and resilient rice cultivars has been subject of increasing research focus (Atwell *et al*., 2014) and advances have been made through traditional plant breeding methods within the *O. sativa indica* and *japonica* types. But there has also been an interest in the introgression of genes from a range of diverse interspecific material. This includes the African rice species *Oryza glaberrima,* which was domesticated in Africa 2000-3000 years ago, independently to the domestication of Asian rice *Oryza sativa*. *O. glaberrima* retains many properties that are specific to challenging African conditions of soil and climate, including limited water availability, abiotic stress, pest and diseases (Bimpong *et al.*, 2011; Agnoun *et al.*, 2012; Bocco *et al.*, 2012; Orjuela *et al.*, 2014; Cubry et al., 2020). *O. glaberrima* is not suitable for commercial rice production due to lodging, shattering, milling difficulties and low yields in comparison to *O. sativa* (Linares, 2002). However the resilience to a range of abiotic and biotic stresses make *O. glaberrima* an attractive target for gene mining and translation (Fig. 3.1; Sarla and Swamy, 2005), which was one of the motivations for the interspecific New Rice for Africa (NERICA) breeding programme (Wambugu *et al*., 2019). This underlying genetic diversity might allow commercial rice to tolerate increasingly unpredictable climatic conditions. Recent genomic sequencing advances for *O. glaberrima* have now added new possibilities (Cubry et al., 2020).

Photosynthetic efficiency and water use efficiency are important components of productivity and abiotic stress resilience (Zhao et al., 2017a). Stomata are key players in both processes, regulating CO_2 assimilation (*A*; parameter abbreviation list can be found in Supp. Table 3.1) acquired for photosynthesis and the water lost by transpiration via stomatal conductance (*gs*). However, improvements in water use efficiency (WUE) incorporates a trade-off between transpiration rate at the expense of net CO_2 assimilation rate (*A*) (Blum, 2009; Lawson *et al*., 2010; Lawson and Blatt, 2014). The leaf stomatal density (SD) value can affect *gs*; recent work using rice with reduced stomatal density has demonstrated that photosynthesis was not

compromised in well-watered conditions but enhanced WUE in all conditions and improved biomass and yield under water limitation (Caine *et al.*, 2019; Mohammed *et al.*, 2019). Consequently, improved yield in water-limiting environments might be achieved by optimisation of stomatal morphology and density. Dynamics of stomatal aperture alteration have also been increasingly highlighted as playing an essential role in improving photosynthetic efficiency and WUE (Drake *et al*., 2013; Lawson and Blatt, 2014). Stomata can take some time to reach stable *gs* (McAusland et al., 2016). Increasing the speed of stomatal opening and closing, closely coupling to *A* (Fig. 3.1), may be important in conserving water and improving crop yields (Tracy Lawson & Vialet-Chabrand, 2019a).

Historically, light saturated carbon assimilation capacity (A_{max}) (mostly under ambient atmospheric [CO_2]) has been a parameter of interest for photosynthesis improvements (Murchie *et al* 2018). However, recent research now makes it clear that the dynamic responses of photosynthesis and photoprotection (such as non-photochemical quenching, NPQ) to the fluctuating field environment are essential for photosynthetic efficiency-based yield gains (Kromdijk *et al.*, 2016; Taylor and Long, 2017; Murchie *et al.*, 2018; Acevedo-Siaca *et al.*, 2020). Light in plant canopies is transient due to architecture, intermittent cloud cover, solar angle and wind (Burgess et al., 2016). The ability of *A* to rapidly adjust to changes in light levels is limited by two major processes; stomatal dynamics and photosynthetic biochemistry (McAusland *et al.*, 2016; Slattery *et al.*, 2018; Acevedo-Siaca *et al.*, 2020). In wheat, slow induction dynamics were estimated to cost 21% of carbon assimilation acquisition (Taylor & Long, 2017). Further dynamic leaf photosynthetic efficiency can be improved through the rapid relaxation of photoprotection (Kromdijk *et al.*, 2016; Hubbart *et al.*, 2018). Under high light NPQ dissipates excess excitation energy as heat. However, in fluctuating light conditions NPQ dynamics can lag behind shifts in light level, limiting photosynthesis. On this basis, it is clear that elucidating photosynthesis related dynamics are an essential focus of improving crop yields and improving abiotic stress tolerance, whereby plants can utilise light and CO_2 with increased efficiency.

Variation in photosynthetic, NPQ and stomatal traits has been examined in *O. sativa*, however there is no comprehensive analysis which compares both induction and decline. We hypothesise that due to the origins within the diverse African climates, substantial variation for dynamic photosynthesis traits may exist within the genome of *O. glaberrima* and we have used a new, whole-genome re-sequenced, resource of 155 *O. glaberrima* accessions

(Wambugu *et al*., 2019; Cubry *et al.*, 2020) to characterise 58 phenotypic traits for photosynthesis and leaf WUE. This includes the use of automated machine-learning to describe stomatal density and gas-exchange methods to facilitate the modelling of *A*, NPQ and *gs* induction and decline dynamics across a large population of individuals. Furthermore, as the effect of environment driven trait adaptation is central to *O. glaberrima*'s novelty, we explore the effect of twenty climatic variables and ecological niche upon trait adaptation within the population. Here, we describe an African rice population with broad heritable variation in a range of useful traits and we provide evidence that dynamic and steady state photosynthesis and photoprotective traits are linked to whole plant growth. To our knowledge this is the largest survey of dynamic photosynthesis for a species in the *Oryza* genus to date. This further highlights the importance of *O. glaberrima* as an essential source of variation for crop improvement and providing a solid base for future research to elucidate physiological processes and pursue trait-related gene identification.

3.3 MATERIALS AND METHODS

3.3.1 Plant material and growth conditions

The seeds of 155 *O. glaberrima* accessions were provided by the Interspecies Comparison & Evolution (RICE) team within Diversité Adaptation Developpement des plantes (DIADE), IRD-Montpellier, France. A table of information presenting the plant material is provided in Supp. Table 3.2.

Plants were grown, measured and processed at the Sutton Bonington Campus, University of Nottingham, UK. Plants were sown and grown in the controlled environment agronomy style glasshouse (Cambridge HOK, UK). Conditions were maintained at a 12-hour dark:light (07.00 – 19.00 hrs) photoperiod, controlled using blackout blinds, temperature of 28±3 ^{o}C and 50-60% relative humidity. Metal halide lamps were used to maintain light levels when they fell below 200 mmol m^{-2} s^{-1} photosynthetically active radiation (PAR). Seeds were heat treated to prevent pathogenesis at the primary seedling stage through immersing in water at 55^{o}C for 15 minutes. Seeds were germinated in a module tray for two weeks before being transplanted to soil pits (5mx5mx1.25m, LxWxD) within the glasshouse. Five replicates of each accession were transplanted in east – west rows, at 20cm intervals, into high nutrient loam-based soil in 2 x 5m concrete pits. Plants were irrigated by drip tapes twice a day, to provide a soil water content close to field capacity.

Due to the size of the population accessions, planting was staggered at 1–2-week intervals. Accessions were grown in rotations of 12 genotypes at a time, with 5 biological replicates, 4 of which were selected for measurement. Plants were measured at 8 weeks old when they were approximately in the mid to late tillering stage. After the measurement of each rotation of 12 genotypes, the top ~12 inches of soil was removed and replaced with fresh loam-soil before commencing the next rotation. Measurements commenced July 2017 and finalised October 2017. The elite *O. sativa* variety 'IR64' was used as a reference genotype and planted as a row in every batch. IR64 was not included in the first four rotations of measurements, as a check genotype was not initially considered, and was measured from 7th August 2017 onwards.

3.3.2 Gas-exchange measurements

An IRGA (infra-red gas analyser, Li-Cor 6400XT, Lincoln, NB, USA) was used on the uppermost fully expanded leaf. A light-induction programme was used: leaves were dark adapted for one hour, the sample leaf was then placed in the leaf cuvette and allowed to achieve steady state in darkness before being subject to a PPFD of 1500 µmol m^{-2} s^{-1}, from in-built red and (10%) blue LED lights, from 0-900 secs, reducing to 100 µmol m^{-2} s^{-1} from 900-1200 secs. A graphical representation of the induction assay can be seen in Fig. 3.1. The leaf cuvette conditions were maintained at a block temperature of 30 ^{o}C, 400 $mmol^{-1}$ mol^{-1} CO_2, flow rate 500 ml min^{-1} and 50-65% humidity. Gas exchange data was logged every 10 s. Measurements were collected between 09.00 and 16.00 hrs. Chlorophyll fluorescence parameters were collected simultaneously, by applying a single saturating pulse before the application of actinic light to attain Fo and Fm and then at intervals of 60s following this for the calculation of ΦPSII (PSII operating efficiency in the light), qP (Photochemical quenching) and NPQ (non-photochemical quenching: measurement of a photoprotective process that estimates the rate constant for PSII heat loss) (Murchie & Lawson, 2013). Intrinsic water use efficiency (iWUE) was calculated post-data collection as CO_2 assimilation rate (*A*) / stomatal conductance (*gs*). We calculate that vapour pressure deficit in the cuvette was approximately 1.51 – 2.10 kPa. Saturation or near-saturation was achieved within this timescale. Raw data for *A*, *gs* and NPQ is shown graphically in Supplementary fig. S3.4 as individual replicates and means per accession. Supplementary figures and data are deposited at Zenodo repository: https://doi.org/10.5281/zenodo.5555930: Murchie., 2021).

3.3.3 Stomatal density and automated stomatal counting

The methodology for the leaf surface impressions, image acquisition, development of an automated stomatal counting method and calculation of the stomatal density can be found in chapter two, method sections 2.3.2 to 2.3.6.

3.3.4 Morphological traits

Plant height, leaf area, root and shoot dry biomass were taken at 8 weeks post-germination, after the completion of gas exchange measurements. Each plant was dug up and care taken to preserve the root system. The shoot area was measured using LiCor LI-3100 area meter (Lincoln, NB, USA). The root ball was soaked and carefully washed to preserve root structure, as described by York *et al.* (2018). The shoot and root material were then placed in a drying oven at 70ºC for 72 hrs before weighing for dry biomass.

3.3.5 Data analysis

All data analyses were performed using R-Studio (v. 4.0.1).

To reduce the temporal and spatial effects of measuring the accessions in batches, a linear mixed-effects model ('lme4' package, v.1.1-26) was used to calculate best linear unbiased predictions (BLUPs) and predicted means, considering the random effects of; accession, sowing date, measurement date, location within the glasshouse (1|block and the interaction of 1|set:block) and, if relevant, IRGA machine (Supp. Fig S3.2 at Zenodo). In the statistical R coding language, this was written as:

```
lmer(trait_of_interest ~ (1|genotype) + (1|sowing_day) + (1|measurement_date)
     + (1|LiCOR) + (1|set:block) + (1|block), data = data)
```

BLUPs are commonly used to account for the random effects that accompany measuring large populations in fluctuating environments (Robinson, 1991) Merk et al., 2012; Zendonadi et al., 2021). After genotype, the largest source of variation was attributed to sowing date, while the position in the glasshouse showed the least variation in the model outputs. The coefficients of the mixed effects model were also used to estimate broad sense heritability (H^2). All results reported here use the adjusted means data generated from the mixed effects model. Normality was tested using the Shapiro-Wilk test. A 0.01 a value was used, as the Shapiro-Wilk test tends to report false negatives in sample sizes exceeding 50 individuals. All data for IR64 were found to be normally distributed, whereas 25 out of 57

traits in the *O. glaberrima* panel showed a deviation from a statistically normal distribution. In the results, and *O. glaberrima* descriptive statistics Table 1, non-normal traits report values for median and interquartile range (IQR) (median (IQR)), whereas normally distributed traits will report mean and standard deviation (mean value (sd value)). A full breakdown of IR64 and *O. glaberrima* normality statistics, box and distribution plots can be found in Supp. Files S2 - S4 at Zenodo.

A bespoke Python pipeline was written to identify the data point at 95% of the maximum and extract values within the induction side of a curve (Supp. File S5 at Zenodo).

The correlation analyses were completed using a Pearson correlation coefficient in the 'Corrplot' package (v. 0.84), with a correlation significance threshold of $p<0.1005$.

The percentage genetic variation (PGV) was calculated as follows; $[(x_{max} - x_{min})/x)] \times 100$. Where x_{max}, x_{min} and x respectively denote the maximum, minimum and mean values for a trait in the population (Gu et al., 2014). This measure is used to quantify the genetic variation of a trait within the combined population. Values more than 100% signify where the range is greater than the mean and representing particularly high underlying genetic variation.

3.3.6 Kinetic modelling

Dynamic modelling of *A,* NPQ and *gs* was performed using a dose-response curve fitting method, previously used to model stomatal responses by Barratt et al. (2021) . The 'drc' (v. 3.0) package was used to analyse and extract several useful parameters for both curve induction and reduction responses (Ritz *et al*., 2015), denoted by $_i$ and $_r$ respectively. The measured parameters are detailed in Table 3.1 and include curve slope ($_{i/r\ slope}$), lower limit ($_{i/r\ min}$), upper limit ($_{i/r\ max}$) and the time taken to reach a defined percentage of the dependant variable, in this case 10 ($_{i/r\ 10}$), 50 ($_{i/r\ 50}$) and 90 ($_{i/r\ 90}$) %. A representation of these parameters on *A* and *gs* response curves can be found in Fig. 3.1. The LL.4 (log-logistic 4-parameter) model was chosen as the best fit for both the *A* induction, *gs* induction and reduction responses. The LL.3 (log-logistic 3-parameter) model was used for NPQ induction and the W2.4 (4-parameter Weibull2) model for NPQ relaxation. The comparison of 8 different models, followed by Akaike's information criterion analysis was used to select the best model fit. All 155 accessions were analysed for *A* induction curves, 24 accessions were removed for *gs* induction and *gs* decline curve fitting, due to unusable curve measurements.

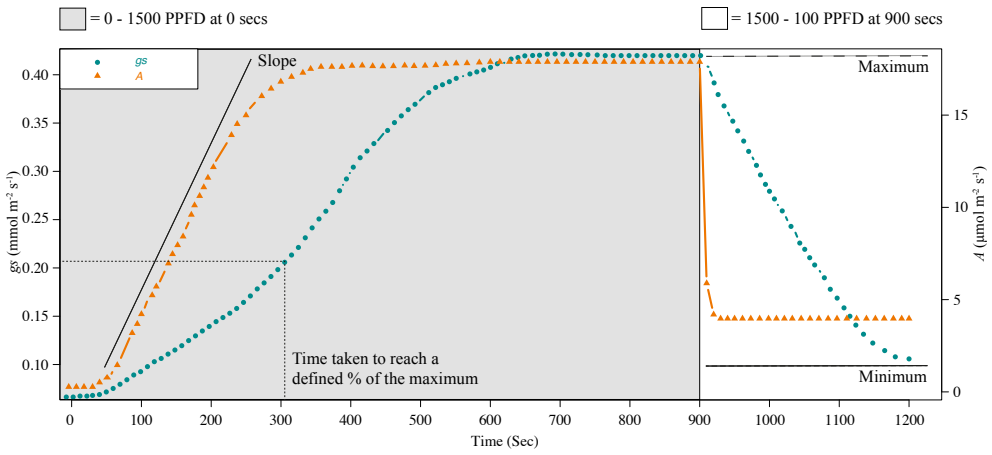


Figure 3.1: Schematic showing example induction and reduction responses to changes in light intensity during the IRGA gas exchange assay used in this study. At zero seconds, dark adapted leaves are exposed to a light change from 0-1500 PPFD until 900 seconds (grey area), followed by a decrease in light from 1500-100 PPFD between 900-1200 seconds (white area). This plot shows real examples of *A* (right hand axis) and *gs* (left hand axis) gas-exchange measurements, and where the parameters; minimum, maximum, slope and time to reach a defined percentage of the curve maximum are captured.

Due to the volume of data, the best fitting model was selected for the induction and reduction curve for each parameter (*gs*, *A*, NPQ) and then applied to all data (e.g. for *gs* induction, a LL.4 model was applied to all 155 genotypes). To ensure that the model selection process captured the variation that may occur in the population, 5 genotypes were randomly chosen for model selection and the consensus model was used. The estimated parameters were generated from the model (min, max, ed50 and slope) was manually cross referenced to the raw data, to ensure these outputs closely represented the raw data. To evidence the fit of the selected models to the raw data, we include a table showing the model fit of a randomly selected accession for each parameter (Supp Table S3.3 at Zenodo). We note the large SE for $A_{r\ slope}$, likely due to the rapid and steep drop in *A* that occurs between two data points, so we attribute less confidence in this parameter.

3.3.7 Multivariate and Climatic Analysis

The multivariate analyses, principal component analysis (PCA) and hierarchical clustering (H-clustering) methods require a complete dataset, with no missing values. Consequently, missing phenotype data was imputed using the missMDA package (v. 1.18). PCA and H-

clustering was performed using the FactoMineR (v. 2.4) package, the H-clustering was performed using the HCPC method. Details of the FactoMineR package and HCPC algorithm can be found in Lê *et al*. (2008) .

The PCA and H-clustering of the phenotypic dataset contained all 155 *O. glaberrima* accessions and the IR64 *O. sativa* representative. The analysis included 64 quantitative trait variables and 4 qualitative variables, including ecological niche from the collection location, country of origin and African region for each accession.

Agroecological niche and geographical coordinates of the collection sites for each *O. glaberrima* accession was provided by AfricaRice (Fig. 3.2).

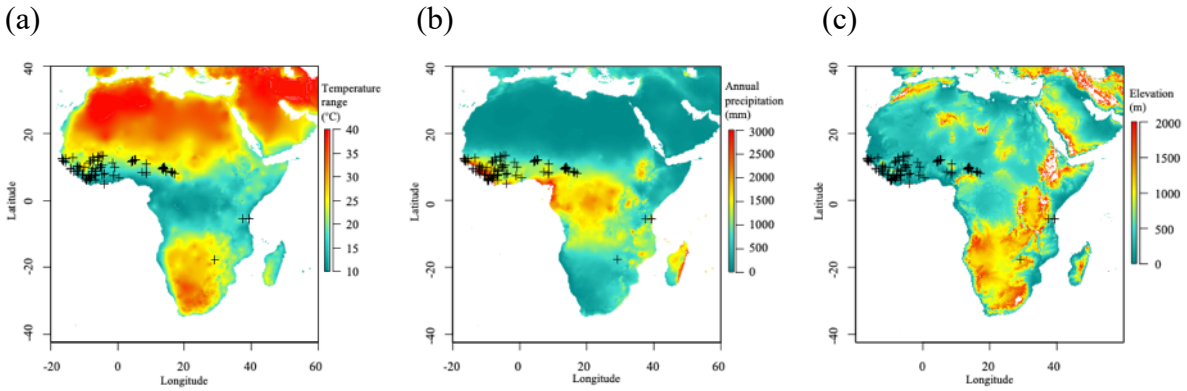


Figure 3.2: Map showing the geographical collection locations for 105 *O. glaberrima* accessions and the annual (a) temperature range, (b) precipitation and (c) elevation range across Africa. Collection locations were provided by AfricaRice and climatic data obtained from the WorldClim bioclimatic variables resource.

We have complete ecological information for all 155 accessions, and geographical coordinates for 105 accessions (Fig.3.2a-c). 19 variables for temperature and precipitation, at the collection site of each accession, were obtained using the BIOCLIM data set (Hijmans *et al*., 2005). Information on the elevation above sea level was obtained using the elevatr package (v. 0.3.1). PCA and H-clustering analysis, and subsequent climate – trait correlations, was performed on the subset of 105 accessions for which we had geographical co-ordinates.

FactoShiny was used to generate summary reports of the PCA and H-clustering analyses on both the phenotype data and climate data, these can be found in Supp. Files S6-S9 at Zenodo.

3.4 RESULTS

3.4.1 Phenotypic analysis of morphology and steady state photosynthesis

Significant variation and high levels of percentage genetic variation (PGV) was identified between accessions across all morphological, gas-exchange and fluorescence traits measured in this *O. glaberrima* panel (Table 3.1, Supp. File S2 and S4 at Zenodo).

Table 3.1: Table showing the range of natural variation and broad-sense heritability (H^2) within a population of diverse *O. glaberrima* accessions across dynamic and static traits. Normally distributed traits report the trait mean and standard deviation (sd), whilst the median and interquartile range (IQR) is given for non-normally distributed traits. PGV is the percentage of genetic variation. Sig. refers to the ANOVA test between two mixed effects models, where the accession is present as an effect in one model and not in another. H^2 was calculated using a linear mixed effects model, as the variation attributed to genotype over total trait variation. A significant result suggests that the accession genotype has an effect and therefore the trait is heritable. *** p<0.0001, ** p<0.001, * p<0.01.

Trait	Min	Max	Mean (sd)	Median (IQR)	PGV	Sig.	H^2
Steady state							
A_{max}	17.92	22.42	20.16 (0.87)		22.30	***	0.11
ETR_{max}	104.10	144. 30	123.00 (7.60)		32.69	***	0.22
gs_{max}	0.26	0.46	0.34 (0.03)		58.82	***	0.17
$iWUE_{max}$	58.34	71.81		62.63 (3.08)	21.40	***	0.07
NPQ_{max}	1.98	2.30	2.13 (0.06)		14.88	***	0.12
$\Phi PSII_{max}$	0.16	0.22	0.19 (0.01)		32.70	***	0.22
qP_{max}	0.43	0.51	0.47 (0.02)		18.71	***	0.17
$Trmmol_{max}$	4.13	5.75	4.81 (0.27)		29.84	***	0.12
VPD_{max}	1.43	1.52	1.47 (0.01)		6.38	***	0.00
Morphological							
Shoot:Root	3.36	9.50		5.55 (1.13)	73.22	***	0.12
Shoot biomass	2.36	7.85		4.40 (0.89)	123.33	***	0.14
Shoot area	279.07	1068.97		652.73 (126.27)	118.51	***	0.16
Root biomass	0.36	1.98		0.77 (0.18)	203.75	***	0.23
Plant height	61.79	93.86	78.77 (6.20)		40.72	***	0.18
Adaxial SD	260.16	353.50	314.08 (18.04)		29.72	***	0.21
Abaxial:Adaxial	1.05	1.40	1.24 (0.06)		18.87	***	0.15
Abaxial SD	324.31	435.99	388.83 (22.54)		28.72	***	0.21

Dynamic							
$gs_{i\ slope}$	-3.90	-1.90		-2.44 (0.38)	78.43	***	0.14
$gs_{i\ max}$	0.35	0.49	0.42 (0.03)		31.88	*	0.11
$gs_{i\ min}$	0.06	0.14		0.08 (0.01)	112.50	**	0.12
$gs_{i\ 10}$	65.50	207.09		108.19 (29.45)	126.83	***	0.18
$gs_{i\ 50}$	179.53	324.80		223.71 (48.40)	64.17	*	0.11
$gs_{i\ 90}$	477.70	684.41		539.21 (48.40)	37.53	NS	0.08
$gs_{i\ rate}$	0.0005	0.0008		0.0006 (<0.01)	45.45	*	0.09
$A_{i\ slope}$	-2.42	-1.50		-1.78 (0.17)	50.84	**	0.10
$A_{i\ max}$	20.44	27.83	23.43 (1.47)		31.53	***	0.20
$A_{i\ min}$	-1.32	-1.09	-1.22 (0.04)		18.85	NS	0.01
$A_{i\ 10}$	49.01	140.97		63.79 (13.96)	136.43	***	0.15
$A_{i\ 50}$	189.56	334.26		217.84 (24.73)	64.97	**	0.12
$A_{i\ 90}$	652.34	849.54		718.78 (52.10)	27.14	NS	0.06
$A_{i\ rate}$	0.03	0.04	0.03 (0.002)		33.33	NS	0.05
$NPQ_{i\ slope}$	-3.48	-2.32		-2.72 (0.22)	42.65	***	0.12
$NPQ_{i\ max}$	2.12	2.44	2.24 (0.05)		14.29	***	0.08
$NPQ_{i\ 10}$	19.57	28.12	23.80 (1.53)		35.75	***	0.11
$NPQ_{i\ 50}$	50.98	56.44	53.65 (1.02)		10.16	NS	0.05
$NPQ_{i\ 90}$	118.70	132.30	125.58 (2.71)		10.82	NS	0.03
$NPQ_{i\ rate}$	0.017	0.02	0.02 (<0.01)		15.00	**	0.07
$gs_{r\ slope}$	6.48	10.49		8.05 (0.80)	49.43	NS	0.07
$gs_{r\ min}$	-0.30	0.02		-0.13 (0.08)	228.57	**	0.11
$gs_{r\ max}$	0.38	0.61	0.48 (0.05)		48.96	**	0.15
$gs_{r\ 10}$	908.91	956.10		922.35 (8.56)	05.11	NS	0.07
$gs_{r\ 50}$	1061.65	1385.05	1208.21 (76.16)		26.77	***	0.26
$gs_{r\ 90}$	1370.52	2116.04		1640.46 (266.31)	44.56	***	0.23
$gs_{r\ rate}$	0.0006	0.0101		0.001 (<0.01)	339.28	***	0.25
$A_{r\ slope}$	-332.60	-322.20		-328.45 (2.15)	3.17	NS	0.01
$A_{r\ min}$	2.88	3.60	3.22 (0.11)		24.05	***	0.11
$A_{r\ max}$	17.09	21.72	19.22 (0.96)		22.36	NS	0.04
$A_{r\ 10}$	901.90	902.20	902.09 (0.04)		0.03	NS	0.03
$A_{r\ 50}$	905.30	905.70		905.45 (0.05)	0.04	NS	0.03
$A_{r\ 90}$	910.60	911.00		910.77 (0.05)	0.04	NS	0.02
$A_{r\ rate}$	0.79	2.36		1.51 (0.09)	103.97	*	0.08
$NPQ_{r\ slope}$	-49.71	-36.27	-41.83 (2.65)		34.35	***	0.21

$NPQ_{r\ min}$	0.56	0.66	0.61 (0.02)	23.00	***	0.16
$NPQ_{r\ max}$	1.99	2.30	2.13 (0.06)	14.55	***	0.10
$NPQ_{r\ 10}$	921.31	923.66	922.60 (0.39)	0.27	***	0.19
$NPQ_{r\ 50}$	946.63	954.76	950.23 (1.41)	0.97	***	0.19
$NPQ_{r\ 90}$	985.20	1008.16	995.50 (4.13)	02.50	***	0.19
$NPQ_{r\ rate}$	0.013	0.020	0.016 (<0.01)	43.75	***	0.14

Root biomass, shoot biomass and shoot area showed a five, three and four-fold variation respectively. Plant height showed a 1.3-fold variation across the *O. glaberrima* population. Significant ($p<0.001$) positive correlations were identified between shoot biomass, shoot:root ratio, shoot area and plant height. While a negative correlation was found between root biomass and shoot:root ratio but a positive correlation between root biomass, against shoot biomass and area. (Fig. 3.3a and c).

Even though key steady state photosynthesis traits showed a relatively narrow distribution (typically between 15 and 40 %), both shoot biomass and shoot area showed ($p<0.01$ - 0.05; Fig. 3.3a and Supp. File S10 at Zenodo), positive correlations to A_{max}, qP_{max}, ETR_{max} and $\Phi PSII_{max}$, providing confidence that steady state photosynthesis is linked to biomass production. gs_{max} showed an almost two-fold variation across *O. glaberrima* accessions. PGV for steady-state traits ranged from 6.38% to 58.82% (Table 3.1), with most traits in the 20 – 30% range, including key photosynthetic traits. All key steady state photosynthetic traits showed significant ($p<0.0001$) positive correlations to one another (Supp. Fig. S3.3; Supp. File S10 at Zenodo). As expected, and detailed in the literature (Lawson and Blatt, 2014), $iWUE_{max}$ was highly correlated with gs_{max} (Fig. 3.3b) (and $Trmmol_{max}$) but not A_{max}, indicating stomatal limitation of *A*.

Stomatal morphology did not show a clear relationship with conductance. A relatively modest 1.3-fold accession dependent variation in the abaxial SD and adaxial SD was observed. The abaxial SD was 1.24 fold greater than adaxial (Fig. 3.4). PGV showed that all SD traits were highly significant ($p<0.0001$, Table 3.1) across the *O. glaberrima* accessions, showing that SD has a genetic basis. However, no association between any SD traits and $iWUE_{max}$, or gs_{max} were detected. Unexpectedly the adaxial SD showed a significant but weak correlation to NPQ_{max}, similarly abaxial SD showed a negative association with plant height (Fig. 3.3d). The SD ratio also showed correlations with NPQ_{max} and plant height, along with A_{max}, ETR_{max}, $\Phi PSII_{max}$, qP_{max} (Fig. 3.3c), the reasons for which are unclear.

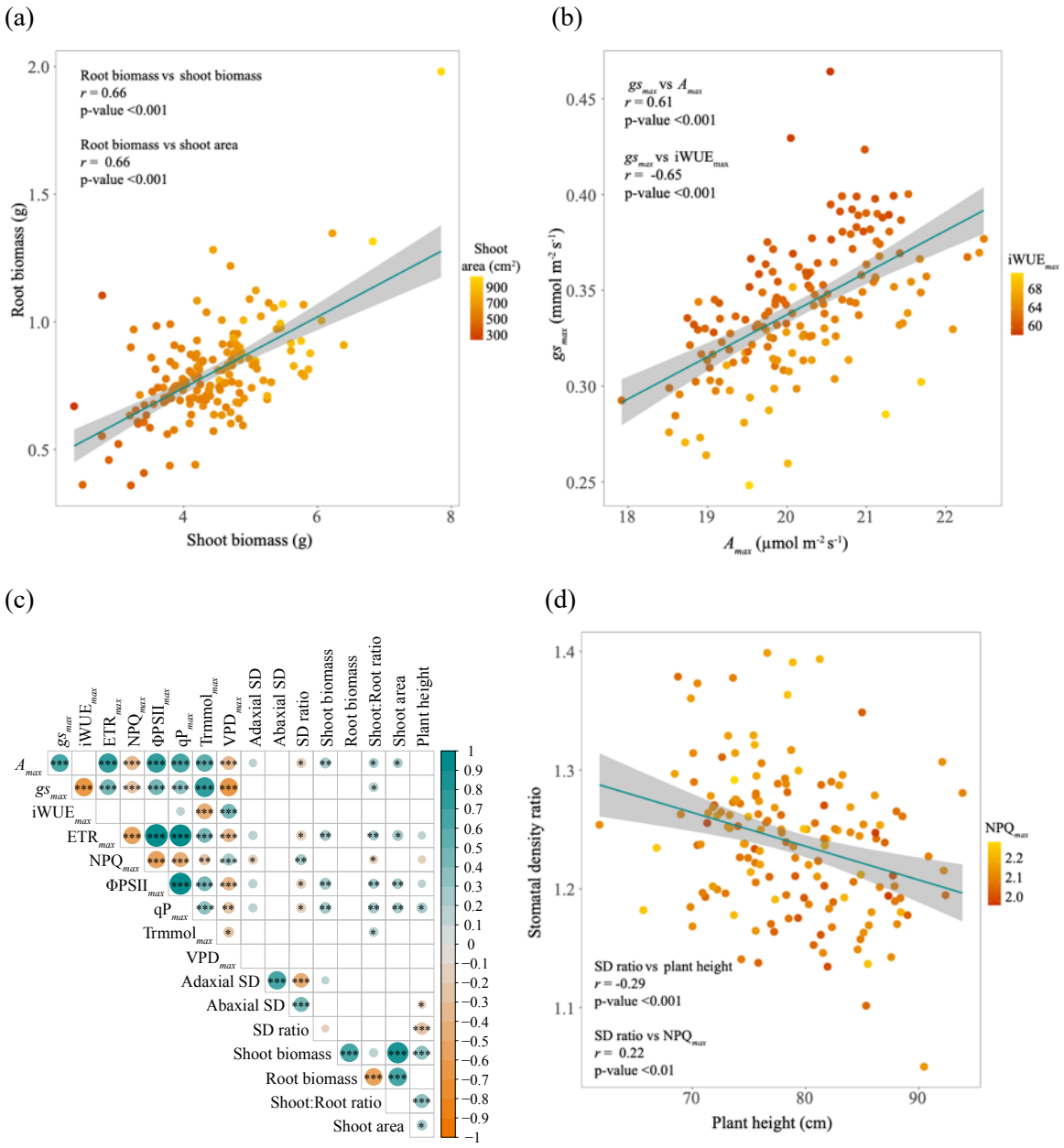
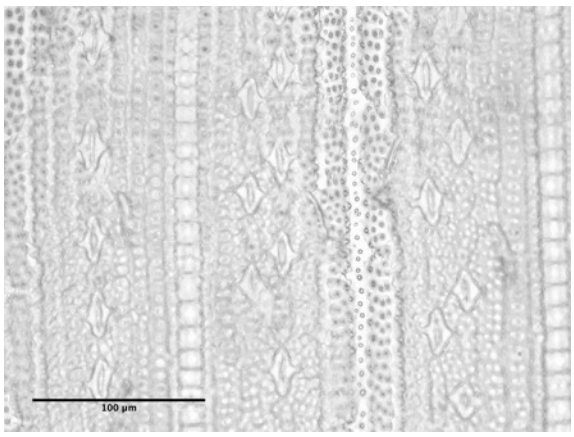


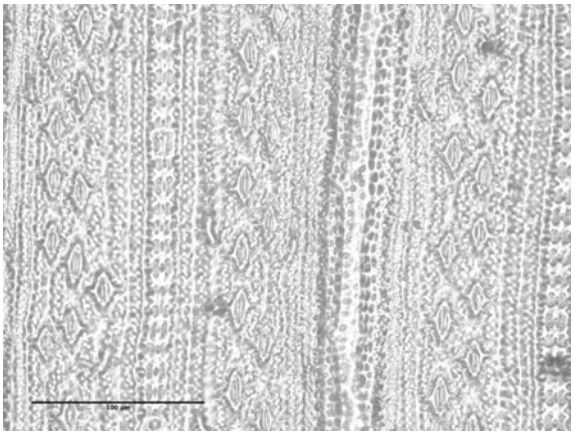
Figure 3.3: *O. glaberrima* shows a range interesting morphological and steady state photosynthetic trait correlations. The colour gradient shows a second correlation against the y-axis variable. (a) Positive correlation between root and shoot biomass (r=0.66, p<0.001, n=155), the second correlation shows root biomass and shoot area (r=0.66, p<0.001, n=155). (b) The effect of gs_{max} on A_{max} (r=0.61, p<0.001) and the second correlation of gs_{max} against iWUE$_{max}$ (r=-0.65, p<0.001, n=131),. (c) Pearson correlation matrix showing associations between morphological and steady-state gas exchange traits, filtered to show trait associations at a p<0.1005 significance threshold. Correlations are scaled by colour, shown in the right-hand scale bar, stars indicate significance between traits (p<0.001***, p<0.01**, p<0.05*). (d) Negative correlation between SD ratio and plant height (r=-0.29, p<0.001,

n=155),, while SD ratio against NPQ_{max} (r=0.22, p<0.01, n=155), shows a positive correlation. Line of best fit is shown with a grey band showing the 95% confidence interval.

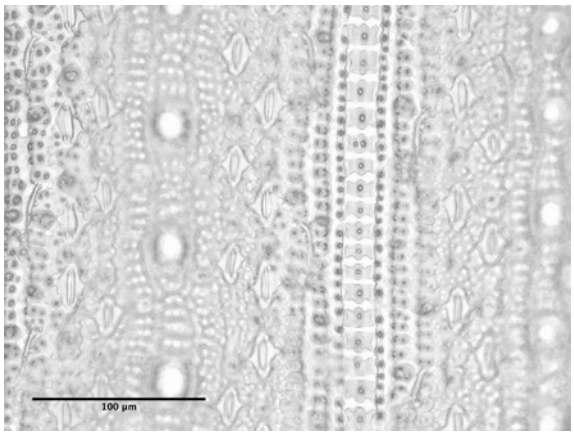
(a) TOG_5464 adaxial



(c) TOG_14116 adaxial



(b) TOG_5464 abaxial



(d) TOG_14116 abaxial

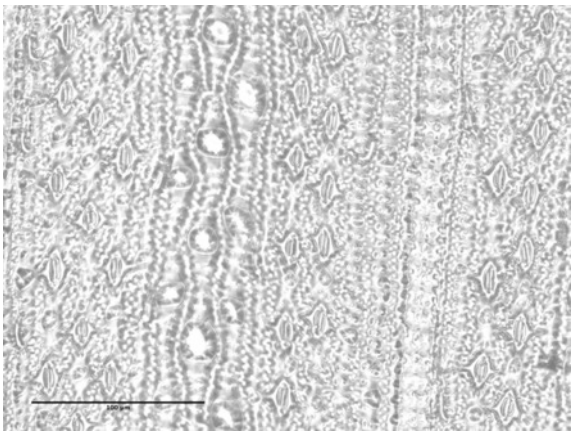


Figure 3.4: Microscope images showing examples of the *O. glaberrima* accessions with highest (TOG_14116) and lowest (TOG_5464) recorded SD within a population of 155 individuals. These images demonstrate the extent of SD variation in the population and the qualitative correlation between high SD and small stomatal size, (a) TOG_5464; Adaxial SD = 260 mm^{-2}, (b) TOG_5464; Abaxial SD = 325 mm^{-2} , (c) TOG_14116; Adaxial SD = 345 mm^{-2} , (d) TOG_14116; Abaxial SD = 426 mm^{-2}

3.4.2 Phenotypic analysis of dynamic photosynthesis

Dynamic responses are now recognised as important determinants of photosynthetic productivity. Responses of gas exchange, fluorescence and photoprotection to light shifts were modelled and show significant variation in 29 traits (Table 3.1, column 'Sig'; Fig. 3.5 (a-d) and Fig. 3.6 (a-b) and Supp. Fig. S3.4(a-f) at Zenodo). A full list of correlation coefficients and significance can be found in Supplementary Data File S3.10.

The well-documented divergence between the induction of *gs* and *A* was observed, where a lag in *gs* induction and reduction occurs relative to *A* (Fig. 3.1 and 3.5a-c). The mean upper limit estimates for *A* induction and reduction curves ($A_{i\ max}$ and $A_{r\ max}$) and *gs* induction and reduction ($gs_{r\ max}$ and $gs_{i\ max}$) curves were similar ($p<0.0001$, Supp. Fig. S3.5a-b at Zenodo) to measured values. The estimated averages for the mean lower limits of the *A* induction ($A_{i\ min}$), *gs* induction and reduction ($gs_{i\ min}$ and $gs_{r\ min}$) curves are close to zero (Table 3.1).

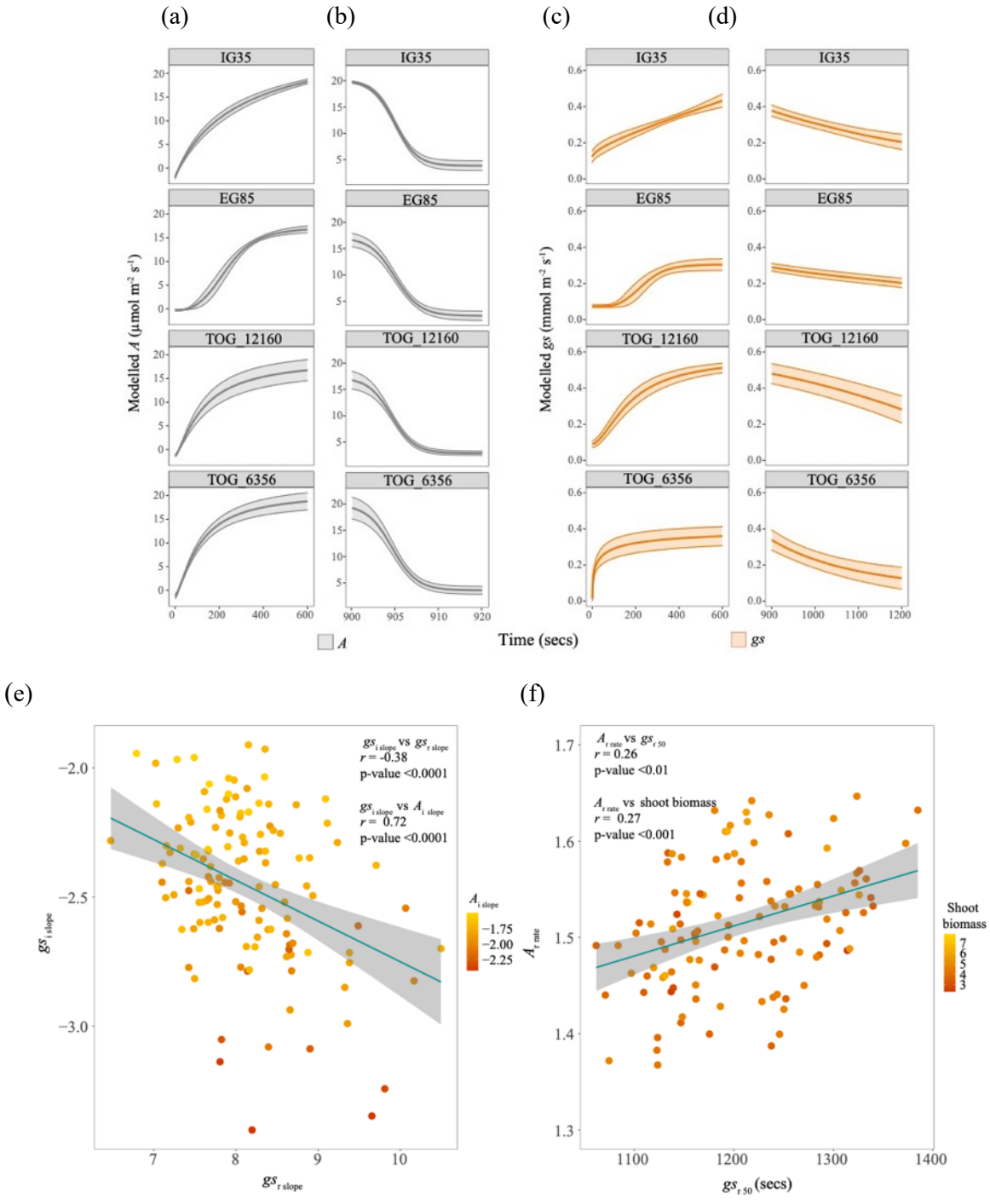


Figure 3.5: Demonstrating the variation of (a, b) *A* and (c, d) *gs* dynamic responses to light intensity changes within the *O. glaberrima* population using four example accessions. IG35 and TOG_6356 were used as example of a 'slow' and 'fast' responding accession respectively, whereas EG85 and TOG_12160 are used to demonstrate intermediate gradient of responses in the population. (e) $gs_{\text{i slope}}$ was correlated with $gs_{\text{r slope}}$ (r=-0.38, p<0.0001, n=131) and $A_{\text{i slope}}$ (r=0.72, p<0.0001, n=131); during induction a greater negative value indicates a steeper slope, this relationship is reversed for the decrease. (f) $A_{\text{r rate}}$ shows positive associations to $gs_{\text{r 50}}$ (r=-0.26, p<0.01, n=131) and shoot biomass (r=-0.27, p<0.001, n=155).

The average time taken to reach 10% of the maximum induction curve was significantly less for CO_2 assimilation ($A_{\text{i 10}}$) than *gs* ($gs_{\text{i 10}}$) whereas the time taken to reach 50% of the induction curve for $A_{\text{i 50}}$ and $gs_{\text{i 50}}$, did not significantly differ. However, the average time to reach induction to 90% of the maximum ($A_{\text{i 90}}$), was significantly longer than that of $gs_{\text{i 90}}$.

Interactions between stomatal and CO_2 assimilation indicate co – dependence, although correlation coefficients are not as strong as expected considering the known stomatal limitation of photosynthesis (Fig. 3.5 e-f). Notably the steepness of the *gs* induction slope ($gs_{\text{i slope}}$) shows consistent positive correlation across with key induction traits, such as $gs_{\text{i 90}}$, $A_{i\text{ slope}}$, $A_{\text{i 90}}$, $A_{\text{i rate}}$ and NPQ_{islope}. The negative correlations between the rate of $gs_{\text{i slope}}$ with $gs_{\text{i rate}}$ and $A_{\text{i rate}}$ indicates the association between a steeper *gs* induction slope and faster rate of induction for both *gs* and *A*.

The rate of induction in high light was positively correlated with rates of decline in low light (Fig. 3.5). There are positive correlations between both $gs_{\text{i rate}}$ vs $gs_{\text{r rate}}$ and $A_{\text{i rate}}$ vs $A_{\text{r rate,}}$ and a negative correlation between $gs_{\text{i slope}}$ vs $gs_{\text{r slope}}$, these consistent correlations between traits suggest that accessions exhibiting rapid stomatal opening also close at a greater rate (Fig. 3.5c-d and Supp. Fig S3.5 at Zenodo). It was also identified that rapid stomatal closure was correlated with enhanced $iWUE_{max}$. A steeper $gs_{\text{r slope}}$ was found to be negatively correlated to $iWUE_{max}$ and accessions with a faster relaxation time ($gs_{\text{r 10}}$, $gs_{\text{r 50}}$ and $gs_{\text{r 90}}$) showed negative correlations to improved $iWUE_{max}$.

Like steady state traits, *A* and *gs* dynamics were also linked to plant biomass and morphology in these data further supporting the role of photosynthesis in determining growth. Greater $A_{\text{i rate}}$ was positively correlated with total plant biomass and shoot biomass. $A_{\text{r rate}}$ showed

positive associations to total plant biomass, shoot biomass, shoot:root ratio and shoot area. $A_{r\ slope}$ had negative associations to shoot biomass, shoot:root ratio and plant height, while a more rapid $A_{r\ 90}$ was correlated to a greater shoot biomass, shoot:root ratio and shoot area. $gs_{i\ rate}$ showed positive associations to total plant biomass, shoot biomass root biomass and shoot area.

Again, there were fewer links with stomatal morphology; a negative association was identified between SD ratio and $A_{r\ rate}$. Only upper leaf SD was also found to have positive relationships to $gs_{i\ 50}$ and $A_{r\ rate}$, negative to $gs_{r\ min}$.

3.4.3 Non-photochemical quenching dynamics

NPQ is of particular interest here because it showed multiple relationships with photosynthesis and biomass. The model estimation of the NPQ induction and relaxation curve upper limit ($NPQ_{i\ max}$ and $NPQ_{r\ max}$) was close to the measured value for $NPQ_{max,}$ providing confidence in the method (Table 3.1; Supp. Fig. S3.5c at Zenodo)

We observed limited significance between the kinetics of NPQ relaxation and kinetics of *A*. Importantly consistent negative correlations between the *A* reduction curve lower limit ($A_{r\ min}$) achieved under 100 PPFD, and $NPQ_{r\ slope}$, $NPQ_{r\ 50}$ and $NPQ_{r\ 90}$ suggest that *A* maintains a higher value under low light conditions when NPQ relaxes rapidly (Kromdijk et al., 2016). Additionally, $NPQ_{i\ slope}$, $NPQ_{i\ rate}$ and the time taken to induce 90% of the maximum ($NPQ_{i\ 90}$), showed negative correlations with A_{max}.

Speed of induction was not closely related to NPQ capacity: only the time taken to reach 90% of the NPQ curve upper limit ($NPQ_{i\ 90}$) positively correlated to a greater NPQ_{max}. Like gas exchange traits NPQ induction and relaxation traits were positively correlated ($NPQ_{i\ slope}$ vs $NPQ_{r\ slope}$ and $NPQ_{i\ rate}$ vs $NPQ_{r\ rate}$).

Interestingly, NPQ and *gs* dynamic traits also showed numerous correlations. A positive correlation was identified between $gs_{i\ slope}$ and $NPQ_{i\ slope}$, and negative correlations between $gs_{i\ slope}$ and $NPQ_{i\ 10}$, $NPQ_{i\ 50}$ and $NPQ_{i\ 90}$. $gs_{i\ rate}$ was positively related to $NPQ_{i\ slope}$ and $NPQ_{i\ 90}$. Accessions with steeper $gs_{r\ slope}$ were also found to have a greater $NPQ_{r\ rate}$ (Fig. 3.6d). These associations highlight a complex interdependent relationship between *gs*, *A* and NPQ and the recent link noted between underlying control of NPQ by PsbS and the dynamics of

stomatal conductance and gas exchange (Fig. 3.6c-d) (Kromdjik et al 2016; Glowacka et al 2018).

Further NPQ relaxation traits were related to morphological and SD traits, indicating that photoprotection has a role in determining growth. $NPQ_{i\ slope}$ and $NPQ_{i\ 90}$ (Fig. 3.6e) showed negative correlations with shoot biomass and shoot area. $NPQ_{i\ rate}$ positively correlated to shoot biomass and shoot area. A more pronounced set of associations was observed during NPQ relaxation; shoot biomass and shoot area, (respectively) showed negative correlations to $NPQ_{r\ slope}$, $NPQ_{r\ 50}$ and $NPQ_{r\ 90}$ (Fig. 3.6f) and positive correlations to $NPQ_{r\ 10}$ and $NPQ_{r\ rate}$. Root biomass showed similar, but not as strong, association to $NPQ_{r\ slope}$, $NPQ_{r\ 10}$, $NPQ_{r\ 90}$ and $NPQ_{r\ rate}$.

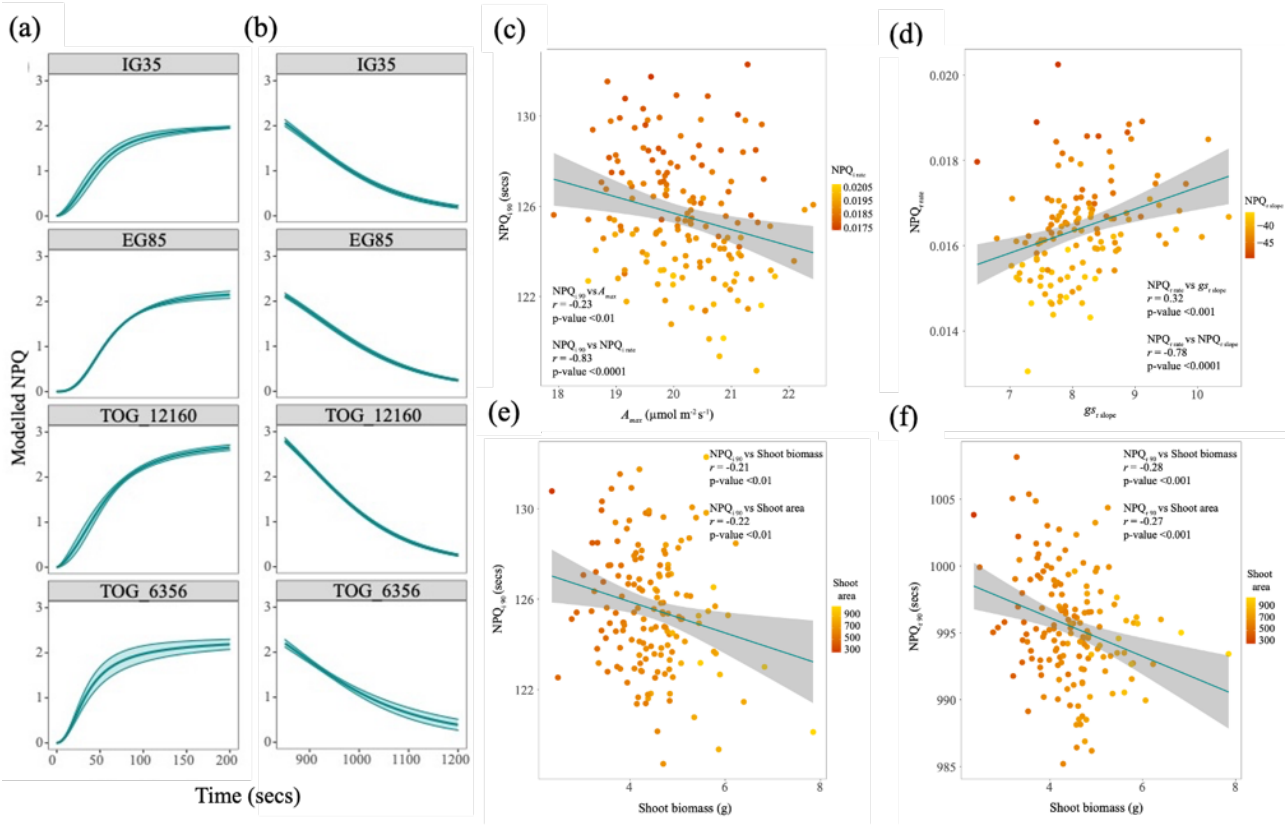


Figure 3.6: Demonstrating the variation of NPQ (a) induction and (b) relaxation responses to light intensity changes within the *O. glaberrima* population using four example accessions, as explained in Fig. 3.5. (c) Negative correlations were identified with $NPQ_{i\ 90}$ against A_{max} (r=0.32, p<0.001) and $NPQ_{i\ rate}$ (r=-0.83, p<0.0001). (d) $NPQ_{r\ rate}$ showed a positive relationship to $gs_{r\ slope}$ (r=-0.83, p<0.0001), where a high value indicates a steeper slope. A negative correlation between $NPQ_{r\ rate}$ and $NPQ_{r\ slope}$ (r=-0.78, p<0.0001), where for this model a more negative value suggests a steeper

relaxation slope. $NPQ_{i\ 90}$ (e) and $NPQ_{r\ 90}$ (f) both showed associations to shoot biomass (r=-0.21, p<0.01; r=-0.28, p<0.01) and shoot area (r=-0.22, p<0.01; r=-0.27, p<0.01).

3.4.4 Trait and ecological comparison between *O. glaberrima* and *O. sativa*

It is informative to compare the *O. glaberrima* trait variation to that of the elite Asian *O. sativa* cultivar, IR64 (Supp. File S3 at Zenodo and Table 3.1) even though caution should be observed using just one genotype. We highlight the slower induction rates of photosynthesis of IR64.

IR64 had a slightly smaller shoot than *O. glaberrima* but a greater root biomass, reflected in the lower shoot:root ratio of IR64, suggesting a greater investment in roots. IR64 height was lower. IR64 displayed a greater SD on the abaxial leaf side than *O. glaberrima* and IR64 had a lower SD ratio.

IR64 did not differ from *O. glaberrima* for A_{max} and NPQ_{max}. However, average ETR_{max} and $\Phi PSII_{max}$ were higher in IR64. IR64 showed a slightly lower gs_{max} and greater $iWUE_{max}$ in comparison to *O. glaberrima*. The latter is likely a direct result of the higher levels of gs_{max} observed in *O. glaberrima*. Clear differences were found in dynamics of *A*, *gs* and NPQ between the two species.

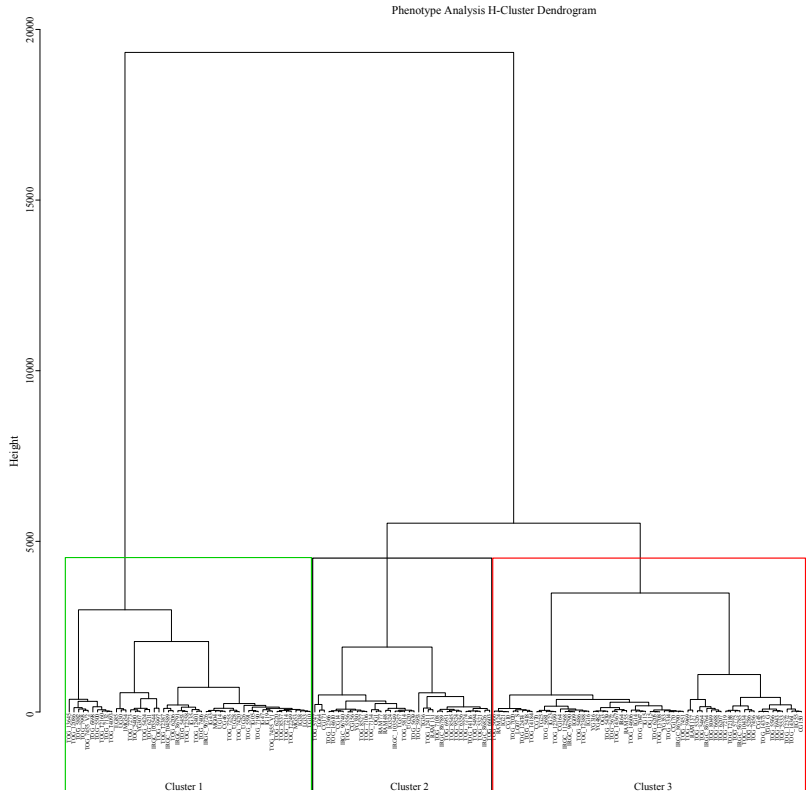
Induction appears to be slower in *O. sativa* (IR64) than *O. glaberrima*, according to mean comparisons in $gs_{i\ 10}$, $gs_{i\ 50}$, $A_{i\ 10}$, $A_{i\ 50}$, $A_{i\ 90}$, $A_{i\ rate}$, $NPQ_{i\ 10}$ and $NPQ_{i\ 50}$ (Supp. Fig. S3.4 at Zenodo). This implies IR64 had a longer *gs* and NPQ lag phase. The initial rapidity of the *gs* induction curve may facilitate the significantly faster *A* response observed in *O. glaberrima*, suggesting that *O. glaberrima* may be able to respond better to the onset of high light than IR64. During the decrease, IR64 and *O. glaberrima* did not significantly differ for $gs_{r\ 10}$, $gs_{r\ 50}$, $gs_{r\ 90}$, $g_{r\ rate}$, $A_{r\ 10}$, $A_{r\ 90}$, $A_{r\ rate}$, $NPQ_{r\ 10}$ and $NPQ_{r\ rate}$. IR64 was found to have a faster reduction response for $A_{r\ 50}$, $NPQ_{r\ 50}$ and $NPQ_{r\ 90}$ in comparison to *O. glaberrima* (Table 3.1 in comparison to Supp. File S3 at Zenodo).

3.4.5 Impact of country and ecology of origin on *O. glaberrima* trait adaptation

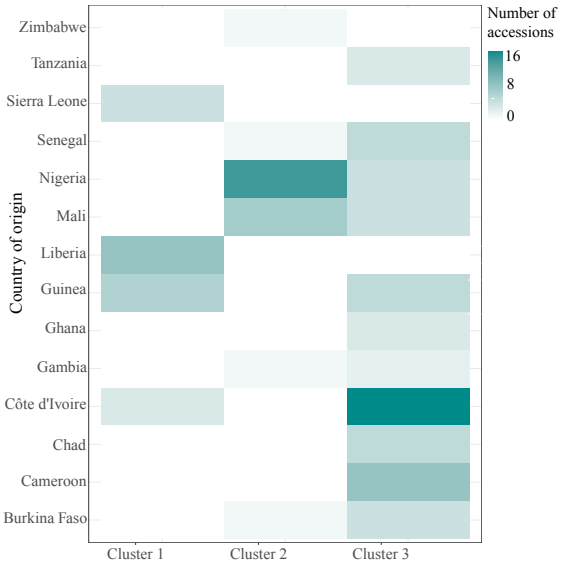
An important aspect of *O. glaberrima*'s novelty is the independent evolution and phenotypic variation to *O. sativa*. We used principal component analyses (PCA) and hierarchical (H)-clustering to explore natural trait variation and the adaptive effect of environmental climatic

variables. Here we identify phenotypic trends which cluster according to country and environment, indicating adaptation and possibly variation in growth strategy.

(a)



(b)



(c)

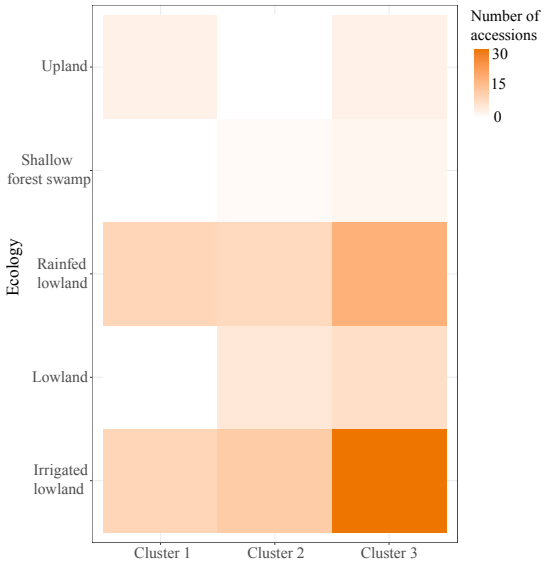


Figure 3.7: Hierarchical clustering on principal components was completed on 155 *O. glaberrima* accessions (a) for 64 phenotypic traits. K-means were used to decide at which point to cut the tree (k=4). Plots were generated detailing the frequency of individuals assigned to each (b) country of origin and (c) ecological niche with the clusters (1-3) identified from hierarchical clustering analysis on climatic data.

The PCA and H-clustering were separated into two grouped analyses for phenotypic and climatic variables. For the PCA of phenotypic traits (Supp. Fig. S3.6b at Zenodo), 12 PCs were selected as they explain 95% of the variance (Supp. Fig. S3.6 at Zenodo). The H-clustering analysis identified three clusters (Fig. 3.7a) with common sources of trait variation (Supp. File S3.7 at Zenodo).

The accessions in cluster 1 are characterised by a slow *gs* reduction time ($gs_{r\ 10/50/90}$), rapid *A* and NPQ induction time ($A_{i\ 50}$ and $NPQ_{i\ 90}$), steep *A* reduction curve ($A_{r\ slope}$), rapid *A* reduction time ($A_{r\ 50/90}$), high values for gs_{max}, $A_{r\ rate}$, $NPQ_{i\ rate}$, shoot:root ratio and low values for root biomass, VPD_{max} and $iWUE_{max}$. Accessions present in cluster 2 demonstrate *gs* reduction curves with a steep slope and rapid reduction times ($gs_{r\ 10/50/90}$), high trait values for NPQ_{max}, VPD_{max}, $iWUE_{max}$ and low values for plant biomass, shoot biomass, shoot area, gs_{max}, ETR_{max} and $\Phi PSII_{max}$. Accessions in the largest group, cluster 3, show high trait values for total biomass, shoot biomass, shoot area, root biomass, A_{max}, low levels of NPQ ($NPQ_{r\ min}$) under reduced light (100 PPFD) and rapid *gs* reduction time ($gs_{r\ 50/90}$). Cluster 3 is the group where IR64 can be found and consists mostly of lowland type accessions.

Adaptation to different environments was explored during the multivariate analyses. Figure 3.8a-b shows axes PC1 and 2 overlaid with ecological niche and country of origin. *O. glaberrima* accessions cluster separately dependent upon their ecological origin, in particular upland or lowland types (Fig. 3.8b). While most accessions originate from lowland-type ecologies, it can still be seen that upland and lowland types show trait variation to one another. Accessions also show a high degree of trait variation due to country of origin that have contrasting climates (Fig. 3.8a). For example, distinct clustering can be seen between landlocked Burkina-Faso, which borders the Sahara, and coastal Gambia. A categorical analysis was performed to establish if the accessions that occupy each cluster of the H-clustering analysis share similar origins (Supp. Fig. 3.7b-c at Zenodo). While there is no obvious relationship, a greater proportion of upland accessions occupy cluster 1, whereas a

large proportion of accessions originating from lowland ecologies are present in cluster 3 (Supp. Fig. S3.7c at Zenodo).

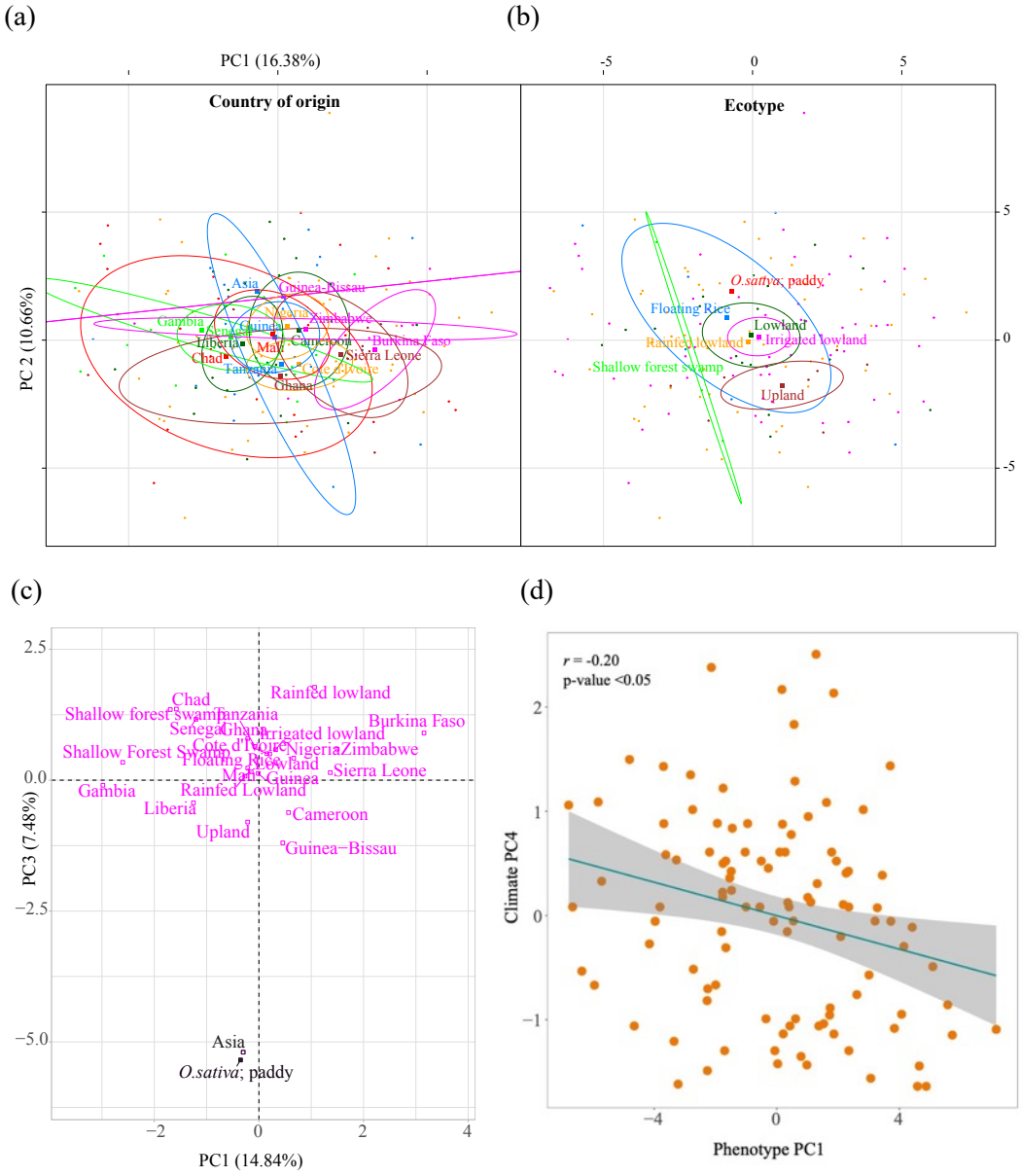


Figure 3.8: Graphical PCA outputs exploring the effects and associations of climate, country and ecological origin for 155 *O. glaberrima* accessions. The PCs 1 and 2 generated from the PCA analysis of phenotypic data are overlaid with 95% confidence ellipses for the *O. glaberima* accessions categorical variables of (a) country and (b) ecological origin. (c) PCs 1 and 3 from the PCA of phenotypic data show the separate clustering of *O. sativa* (IR64), based on country of origin and ecology categories, from *O. glaberrima.* (d) PC1 from the PCA on phenotypic data was found to be a function of PC4 of climatic data PCA analysis.

The diversity of climates and elevations (Fig. 3.2a-c) are likely to have directly impacted trait adaptation and resilience. A PCA focused on climatic traits explored the relationship between climate and phenotype. The first 4 PCs explain 90% of trait variation in the population (Supp. File S8 at Zenodo). H-clustering identified three distinct clusters of accessions with common sources of variation in climatic variables (Supp. Files S9 and S11 at Zenodo). A categorical analysis of ecological niche and country of origin, for the accessions present in each cluster showed a clear distinction of climate-based clustering due to country of origin (Fig. 3.7c). Cluster 1 contains all accessions that originate from the neighbouring countries of Liberia and Sierra Leone. Cluster 2 contains all accessions from Zimbabwe and most accessions originating from Nigeria. While cluster 3, which contains the largest number of accessions, contains all individuals originating from Cameroon, Chad, Ghana, Tanzania and the majority of accessions from Côte d'Ivoire and Senegal.

With the extensive phenotypic and climatic variables reduced to a small number of components, we completed a correlation analysis between the phenotypic and climatic trait PCs to identify groups of climatic drivers on trait adaptation. A significant positive association was identified between trait PC1 and climatic PC4 (r=-0.20, $p<0.05$; Fig. 3.8d), suggesting key traits contributing to phenotypic traits PC1, which includes photosynthetic traits and shoot biomass, have adapted in response to precipitation related variables that contribute to climate PC4 loadings. Other significant associations were identified between phenotype PC5 and climate PC4 (r=0.25, $p<0.05$), phenotype PC8 and climate PC2 (r=0.24, $p<0.05$), phenotype PC11 and climate PC3 (r=0.25, $p<0.05$) (Supp. File S12 and Supp. Fig. S3.8 at Zenodo).

During the multivariate analyses, we observed that *O. sativa* IR64 values cluster separately from *O. glaberrima* for both ecology and country of origin. This can be seen most clearly when plotting PCs1 and 3, where the two species cluster distinctly for the Asian country of origin and paddy field ecology (Fig. 3.8c).

3.5 DISCUSSION

Crop production in future climates has the challenge of increasing productivity whilst retaining resilience. To do so, optimising interactions and trade-offs between carbon assimilation, photoprotection and water loss will be essential. However, we do not yet have

complete understanding of the genetic basis of the co-regulation of the interlinked processes and components (light harvesting, photoprotection, electron transport, carbon assimilation and stomatal conductance) involved. Recent progress shows that crop productivity and water use efficiency is only partly dependent upon 'steady state' maximum values of A_{max} and gs_{max}. SD, stomatal conductance and photoprotection dynamics have been identified as critical traits to optimise carbon assimilation and minimise abiotic stress (Kromdijk *et al.*, 2016; Caine *et al.*, 2019; Faralli *et al*., 2019). However, elite gene pools may be genetically narrow and poorly adapted to challenging environmental conditions. Attention is increasingly focused upon underutilised crop species and wild relatives as a source of genetic diversity to improve resilience in commercial species (Draic *et al*., 2011). Whilst the variation for photosynthesis induction has been partly characterised in *O. sativa*, this is not true of *O. glaberrima* (Acevedo-Siaca *et al.*, 2020; Acevedo-Siaca *et al.*, 2021) . The *O. glaberrima* association panel used here was developed as a resource for crop improvement, which may have diversity not available in *O. sativa* (Agnoun *et al*., 2012). For the first-time a comprehensive analysis of photosynthesis and morphology- related traits has been completed in *O. glaberrima*. Our novel approach uses a large pool of accessions, with a large range of heritable natural variation to explore the natural variation and relationships in these traits. While we cannot make a meaningful comparison between *O.glaberrima and O.sativa* we observed key differences with the former showing faster photosynthesis induction. Though a caveat should be made this observation is only against one *O. sativa* cultivar and demonstrates the need for a comprehensive comparison between the two species in future. This may be an indication of adaptation to drier soils and air generally, requiring faster opening and closure of stomata (Lawson and Vialet-Chabrand 2019).

Here we have described extensive natural variation in *O. glaberrima* for steady state, induction and relaxation / reduction responses for *A* and g*s*. This suggests underlying genetic diversity to these traits that could be identified and exploited. We identified indications of heritability (H^2) and underlying genetic variation (PGV) in many of these traits (Table 3.1). Trait heritability values were comparable to estimates of similar traits from maize (Choquette et al., 2019) but they are marginally lower than those previously demonstrated in *O. sativa* (Qu et al., 2017) though a genetic component still indicated. A calculation of heritability using genomic data would provide a more accurate estimation (Zhu & Zhou, 2020). This would be useful when selecting traits for genetic introgression or characterisation. The large number of accessions used here (155) permits a statistical comparison that was not possible

in related studies on dynamic photosynthesis in *O.sativa* where fewer lines were analysed (Acevedo-Siaca *et al.*, 2020).

A global PCA and clustering analysis showed a distinction between clusters of high biomass (cluster 3), low biomass (cluster 2) and low root biomass (cluster 1). The fast *gs* decrease, low gs_{max}, high NPQ_{max} and high $iWUE_{max}$ of cluster 2 would suggest a conservative type geared toward water conservation, whilst the high total biomass of cluster 3 is consistent with a fast growth type displaying rapid *gs* decrease, low NPQ and a higher A_{max}. The association of cluster 3 with wetter lowland environments is consistent with higher productivity. We therefore see a general consistency in these 2 clusters with photosynthetic, water-use and biomass production 'strategy'. It is also notable that steady state A_{max} correlates well with biomass, suggesting that capacity for higher photosynthesis is still important. Increases in photosynthetic capacity are known to improve light responses in rice (Sun *et al*., 2016).

3.5.1 Extensive natural variation identified in dynamic photosynthetic traits

In recent years there has been a shift in photosynthesis related research towards dynamic responses in place of steady state values (Kromdijk *et al.*, 2016; Taylor and Long, 2017; Murchie *et al.*, 2018; Acevedo- Siaca *et al.*, 2020). It is now recognised that irradiance fluctuations in field conditions, and the ability of stomatal and photosynthetic responses to respond instantaneously can substantially affect plant productivity (Taylor & Long, 2017). To enable greater productivity in dynamic environments such as a crop canopy, one would anticipate that all components of photosynthesis would respond rapidly to 'track' light closely. Each component has a different effect, thus fast activation of the Calvin cycle and CO_2 assimilation during induction is beneficial, while rapid reduction of NPQ and fast stomatal closure at transition to low light enables the attainment of improved CO_2 efficiency and iWUE at low light.

It is clear that we see some independence of dynamic traits, but interesting associations appear which indicate a link with biomass. Recent research suggests that major yield gains can be made by enhancing photoprotection capacity and NPQ dynamic responses (Kromdijk *et al.*, 2016; Hubbart *et al.*, 2018). Rapid NPQ relaxation can remove the limitation on quantum yield of CO_2 assimilation, allowing a quicker recovery of photosynthetic efficiency upon *A* reduction (Kromdijk *et al.*, 2016; Murchie and Ruban, 2020). Our findings support this: NPQ relaxation dynamics were the only group found to have ubiquitous associations to

increased shoot biomass and area. Notably, we also observed that values for *A* under low light were greater in those accessions that exhibited rapid NPQ relaxation and those that have lower NPQ capacity under low light ($NPQ_{r\ min}$). It is also hypothesised that faster induction of CO_2 assimilation may reduce the need for photoprotection during induction (McAusland & Murchie, 2020), however we found no association between *A* induction traits and NPQ dynamic or steady state values. We did find that faster NPQ induction is associated with greater photosynthetic capacity, shoot area and biomass.

Whilst no associations were identified between NPQ and *A* reduction dynamics, we found strong positive correlations between the speed of *gs* and NPQ dynamics. This may highlight the importance of the key NPQ protein, Photosystem II Subunit (PsbS), on stomatal conductance, as shown by Głowacka *et al.* (2018). Whereby PsbS overexpression, which increases both NPQ capacity and NPQ dynamic rate (Kromdijk *et al.*, 2016; Głowacka *et al.*, 2018; Hubbart *et al.*, 2018), it also reduces the extent of stomatal opening in tobacco. This may be reflected here by the negative correlation between NPQ_{max} and gs_{max}, also that *gs* induction rate was lower when NPQ induction was faster. This highlights the need to further explore the associations between NPQ and *gs* dynamics: these have not been elucidated although there is a general principle that limitations imposed by *gs* or Rubisco activation state would result in a further reduction of electron transport and an enhanced NPQ. We suggest that in *O.glaberrima* NPQ may be a key factor in regulating the degree of *gs*, through the effect of PsBs on stomatal opening, and *A* reduction dynamics. Akin to the relation between *A* and *gs*, there is a trade off in NPQ as it reduces photosynthetic quantum yields under low irradiance.

No association was identified between the water use related traits, *gs* and $iWUE_{max}$ and stomatal density, this may be because the variation was less than that needed to produce changes in gas exchange traits (Caine *et al.*, 2019; Mohammed *et al.*, 2019). It is also possible this highlights the importance of stomatal size and morphology, rather than density, on these traits. Smaller stomata have been shown to have improved water use efficiency, maximal stomatal conductance and dynamics (Chatterjee *et al.*, 2020; Drake *et al.*, 2013; Dittberner *et al.*, 2018; Lawson and Vialet-Chabrand, 2019). However, the positive correlations we identified between SD ratio, NPQ_{max} and the level of NPQ achieved under low light ($NPQ_{r\ min}$) is unusual. The significant negative association between SD ratio and $A_{r\ rate}$ has no direct interpretation but may indicate SD ratio is a trait worthy of further work.

Upper leaf SD had positive relationships to $gs_{i\ 50}$ and $A_{r\ rate}$, negative to $gs_{r\ min}$ also indicates that distinction between the leaf surfaces may be important.

We note that some relationships highlighted in the Pearson pairwise correlation analysis showed significant but relatively low correlation coefficient (*r*) values (Fig. 3.6e-f), but those noted here in the text were found to have consistent relationships. Therefore, we consider these findings to be biologically relevant, particularly as they support prior findings in the literature (Lawson and Blatt, 2014; Kromdijk et al., 2016; Glowacka et al., 2018; Taylor and Long, 2018). However, we suggest these findings are followed up with an improved experimental design and a larger number of biological replicates, which may provide further clarity to the initial findings here. The further elucidation between the interplay of photoprotective, stomatal and assimilation dynamics should include detailed morphological characterisation (Ohsumi *et al.*, 2007; Drake *et al*., 2013; McAusland *et al.*, 2016), including the associated mesophyll conductance (Campany *et al.*, 2016; Deans *et al.*, 2019). The proportion that photosynthetic dynamics are limited by stomata or biochemistry seems species dependent (Tinoco-Ojanguren and Pearcy, 1993; Taylor and Long, 2017; De Souza *et al.*, 2020). *O. sativa* photosynthetic induction has been shown to be predominantly limited by biochemistry (Acevedo-Siaca *et al.*, 2020; Acevedo-Siaca *et al.*, 2021) and the same assumption might be extended to *O. glaberrima* due to a similar genomic composition (Stein et al., 2018), however we conclude from our data that stomatal limitations may be more pronounced in *O. glaberrima*.

We recognise limitations in the glasshouse experimental design used here. A linear mixed effects model and the calculation of BLUPs was used to mitigate the effects of temporal and spatial factors, where variation may have been overrepresented in certain genotypes due to the lack of randomised design and measurement of accessions across the summer - autumn seasonal transition. Furthermore, while *O. sativa* IR64 was used as a check-genotype, it was not introduced until the fifth rotation of accessions and could not be used to normalise the dataset against. A repeat experiment would be recommended to confirm our associations identified here. A randomised design, ideally with all genotypes sown at once in subsequent repetitions, would be more statistically robust and may have yielded clearer trait associations. We also recommend the inclusion of multiple *O. sativa* genotypes in each measurement

rotation that could be used to check for environmental effects, to normalise against and as an opportunity to complete a greater comparison between the two species.

3.5.2 Accessions have adapted to variable ecological and environmental regimes in different countries

No comprehensive studies exist that tease apart the ecological and environmental variables that correlate to specific trait adaptation in *O. glaberrima*. This information is useful from an evolutionary perspective but may be essential in the selection of cultivars for abiotic stress tolerance and trait-related genetic characterisation.

Of note, we identified a significant association between the climate PC4 and phenotype PC1 (Fig. 3.8d; Supp. File. S12 at Zenodo). This relationship suggests that key photosynthetic traits contributing to PC1 have adapted in response to climatic pressures associated with PC4, such as elevation and the combined effect of temperature and precipitation. However, these are broad observations for climatic-trait correlations across the African continent, lacking resolution that can be seen in studies on a discrete geographical area (Wolfe & Tonsor, 2014). For the selection of abiotic stress tolerance adapted cultivars, the H-clustering analyses would be of particular use, as this generated three distinct clades of *O. glaberrima* accessions stemming from similar climatic and phenotypic variables. These clusters of individuals should be treated with caution, and we recommend re-analysis with equal numbers of individuals representing the categories discussed here. This will ensure the clusters identified here are not because of unequal weighting. Furthermore, the climatic H-clustering demonstrated clear grouping of accessions due to country of origin (Fig. 3.7b), suggesting that a higher resolution analysis of environmental effect on trait adaptation would be beneficial. Re-analysis using equal representation from each country or ecological group would be possible within this panel, though would substantially reduce the overall sample size as lowland accessions predominate. Alternatively, a greater number of accessions, representing each factor discussed here, could be obtained from the AfricaRice seed bank.

We identified adaptation based upon ecological origin in the PCA analysis (Fig. 3.8b), supporting a known distinction between *O. glaberrima* upland and lowland phenotypes (Ghesquière, 1997). However, there is no comprehensive description in the literature of the physiological differences that are associated with each ecological origin. Though due to the

unequal representation of accessions from each ecological niche in this analysis it is difficult to obtain a clear indication of the effect of ecology on trait adaptation.
The environmental analysis completed here produces useful information of accessions displaying similar phenotypic qualities because of environmental adaptation. Equally, this also highlights the requirement for a dedicated study to truly elucidate the environmental and ecological trait adaptation of *O. glaberrima*, utilising equally represented accessions from a range of ecological niches and assessing physiological adaptation to climatic variables at a range of spatial scales.

4.3.2 Conclusions

Here, we have demonstrated that *O. glaberrima* has broad, heritable natural variation in a range of important traits, which are likely to aid in the improvement of *O. sativa*. This is the first study to describe photosynthetic, photoprotection and dynamic traits in *O. glaberrima*, the size of which is not matched in panels of *O. sativa* accessions. The phenotyping efforts compiled here will provide a basis for the identification of interesting traits for physiology research, aid in the selection of accessions for crop improvement efforts and information for genetic characterisation.

Journal of Experimental Botany
https://doi.org/10.1093/jxb/erab459 Advance Access Publication 17 October, 2021
This paper is available online free of all access charges (see https://academic.oup.com/jxb/pages/openaccess for further details)

RESEARCH PAPER

Out of Africa: characterizing the natural variation in dynamic photosynthetic traits in a diverse population of African rice (*Oryza glaberrima*)

Sophie B. Cowling[1], Pracha Treeintong[1], John Ferguson[1,*], Hamidreza Soltani[2], Ranjan Swarup[1], Sean Mayes[1] and Erik H. Murchie[1,†]

[1] Division of Plant and Crop Science, School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD, UK
[2] Advanced Data Analysis Centre, University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD, UK
* Present address: Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK.
† Correspondence: erik.murchie@nottingham.ac.uk

Received 6 September 2021; Editorial decision 14 October 2021; Accepted 15 October 2021

Editor: Tracy Lawson, University of Essex, UK

Abstract

African rice (*Oryza glaberrima)* has adapted to challenging environments and is a promising source of genetic variation. We analysed dynamics of photosynthesis and morphology in a reference set of 155 *O. glaberrima* accessions. Plants were grown in an agronomy glasshouse to late tillering stage. Photosynthesis induction from darkness and the decrease in low light was measured by gas exchange and chlorophyll fluorescence along with root and shoot biomass, stomatal density, and leaf area. Steady-state and kinetic responses were modelled. We describe extensive natural variation in *O. glaberrima* for steady-state, induction, and reduction responses of photosynthesis that has value for gene discovery and crop improvement. Principal component analyses indicated key clusters of plant biomass, kinetics of photosynthesis (CO_2 assimilation, *A*), and photoprotection induction and reduction (measured by non-photochemical quenching, NPQ), consistent with diverse adaptation. Accessions also clustered according to countries with differing water availability, stomatal conductance (g_s), *A*, and NPQ, indicating that dynamic photosynthesis has adaptive value in *O. glaberrima*. Kinetics of NPQ, *A*, and g_s showed high correlation with biomass and leaf area. We conclude that dynamic photosynthetic traits and NPQ are important within *O. glaberrima*, and we highlight NPQ kinetics and NPQ under low light.

Keywords: African rice, dynamic modelling, natural variation, NPQ, *O. glaberrima*, photosynthesis, rice, stomatal conductance.

Introduction

The climate crisis places crop yields under increasing pressure from biotic and abiotic constraints and constitutes a major threat in meeting global food demand (Ray *et al.*, 2019). Substantial yield decreases in key cereal crops are predicted to occur in both vulnerable and productive regions (Black *et al.*, 2008; Challinor *et al.*, 2014). Asian rice (*Oryza sativa*), a dietary

stable to a third of the global population, is predicted to experience yield losses up to 37% by the end of the century due to climate change-driven drought events Bocco *et al.*, 2012; Zhao *et al.*, 2017). The development of productive and resilient rice cultivars has been the subject of increasing research focus (Atwell *et al.*, 2014), and advances have been made through traditional plant breeding methods within the *O. sativa indica* and *japonica* types. However, there has also been an interest in the introgression of genes from a range of diverse interspecific material. This includes the African rice species *Oryza glaberrima*, which was domesticated in Africa 2000–3000 years ago, independently of the domestication of Asian rice *O. sativa*. *Oryza glaberrima* retains many properties that are specific to challenging African conditions of soil and climate, including limited water availability, abiotic stress, pests, and diseases (Bimpong *et al.*, 2011; Agnoun *et al.*, 2012; Bocco *et al.*, 2012; Cubry *et al.*, 2020).

Oryza glaberrima is not suitable for commercial rice production due to lodging, shattering, milling difficulties, and low yields in comparison with *O. sativa* (Linares, 2002). However, the resilience to a range of abiotic and biotic stresses makes *O. glaberrima* an attractive target for gene mining and translation (Fig. 1; Sarla and Swamy, 2005), which was one of the motivations for the interspecific New Rice for Africa (NERICA) breeding programme (Wambugu *et al.*, 2019). This underlying genetic diversity might allow commercial rice to tolerate increasingly unpredictable climatic conditions. Recent genomic sequencing advances for *O. glaberrima* have now added new possibilities (Cubry *et al.*, 2020).

Photosynthetic efficiency and water use efficiency (WUE) are important components of productivity and abiotic stress resilience (Zhao *et al.*, 2017). Stomata are key players in both processes, regulating CO_2 assimilation (A; the parameter abbreviation list can be found in Supplementary Table S1 and the water lost by transpiration via stomatal conductance (g_s). However, improvements in WUE incorporate a trade-off between transpiration rate at the expense of net CO_2 assimilation rate (A) (Blum, 2009; Lawson *et al.*, 2010; Lawson and Blatt, 2014). The leaf stomatal density (SD) value can affect g_s; recent work using rice with reduced SD has demonstrated that photosynthesis was not compromised in well-watered conditions but enhanced WUE in all conditions and improved biomass and yield under water limitation (Caine *et al.*, 2019; Mohammed *et al.*, 2019). Consequently, improved yield in water-limiting environments might be achieved by optimization of stomatal morphology and density. Dynamics of stomatal aperture alteration have also been increasingly highlighted as playing an essential role in improving photosynthetic efficiency and WUE (Drake *et al.*, 2013; Lawson and Blatt, 2014). Stomata can take some time to reach stable g_s (McAusland *et al.*, 2016). Increasing the speed of stomatal opening and closing, closely coupling to A (Fig. 1), may be important in conserving water and improving crop yields (Lawson and Vialet-Chabrand, 2019).

Historically, light-saturated carbon assimilation capacity (A_{max}) (mostly under ambient atmospheric [CO_2]) has been a parameter of interest for photosynthesis improvements (Murchie *et al.*, 2018). However, recent research now makes it clear that the dynamic responses of photosynthesis and photoprotection [such as non-photochemical quenching (NPQ)] to the fluctuating field environment are essential for photosynthetic efficiency-based yield gains (Kromdijk *et al.*, 2016; Taylor and Long, 2017; Murchie *et al.*, 2018; Acevedo-Siaca *et al.*, 2020). Light in plant canopies is transient due to architecture, intermittent cloud cover, solar angle, and wind (Burgess *et al.*, 2016). The ability of A to rapidly adjust to changes in light levels is limited by two major processes: stomatal dynamics and photosynthetic biochemistry (McAusland

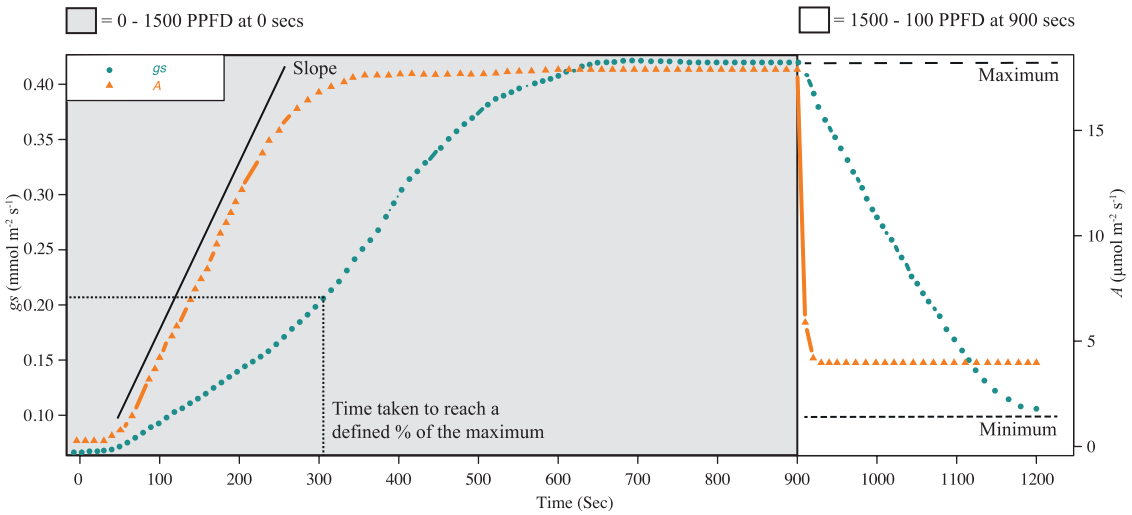


Fig. 1. Schematic showing example induction and reduction in response to changes in light intensity during gas exchange measurements. These examples of raw A and g_s gas exchange measurements show the modelled dynamic response parameters; minimum, maximum, slope, and time to reach a defined percentage of the curve maximum.

et al., 2016; Slattery *et al.*, 2018; Acevedo-Siaca *et al.*, 2020). In wheat, slow induction dynamics were estimated to cost 21% of carbon assimilation acquisition (Taylor and Long, 2017). Further dynamic leaf photosynthetic efficiency can be improved through the rapid relaxation of photoprotection (Kromdijk *et al.*, 2016; Hubbart *et al.*, 2018). Under high light, NPQ dissipates excess excitation energy as heat. However, in fluctuating light conditions, NPQ dynamics can lag behind shifts in light level, limiting photosynthesis. On this basis, it is clear that elucidating photosynthesis-related dynamics is an essential focus of improving crop yields and improving abiotic stress tolerance, whereby plants can utilize light and CO_2 with increased efficiency.

Variation in photosynthetic, NPQ, and stomatal traits have been examined in *O. sativa*; however, there is no comprehensive analysis which compares both induction and decline. We hypothesize that due to the origins within the diverse African climates, substantial variation for dynamic photosynthesis traits may exist within the genome of *O. glaberrima* and we have used a new, whole-genome re-sequenced, resource of 155 *O. glaberrima* accessions (Wambugu *et al.*, 2019; Cubry *et al.*, 2020) to characterize 58 phenotypic traits for photosynthesis and leaf WUE. This includes the use of automated machine learning to describe SD and gas exchange methods to facilitate the modelling of A, NPQ, and g_s induction and decline dynamics across a large population of individuals. Furthermore, as the effect of environment-driven trait adaptation is central to the novelty of *O. glaberrima*, we explore the effect of 20 climatic variables and ecotype upon trait adaptation within the population. Here, we describe an African rice population with broad heritable variation in a range of useful traits and we provide evidence that dynamic and steady-state photosynthesis and photoprotective traits are linked to whole-plant growth. To our knowledge, this is the largest survey of dynamic photosynthesis for a species in the *Oryza* genus to date. This further highlights the importance of *O. glaberrima* as an essential source of variation for crop improvement and providing a solid base for future research to elucidate physiological processes and pursue trait-related gene identification.

Materials and methods

Plant material and growth conditions

The seeds of 155 *O. glaberrima* accessions were provided by the Interspecies Comparison & Evolution (RICE) team within Diversité Adaptation Developpement des plantes (DIADE), IRD-Montpellier, France. A table of information presenting the plant material is provided in Supplementary Table S2.

Plants were grown, measured, and processed at the Sutton Bonington Campus, University of Nottingham, UK. Plants were sown and grown in a controlled-environment agronomy-style glasshouse (Cambridge HOK, UK). Conditions were maintained at a 12 h dark:light (07.00–19.00 h) photoperiod, controlled using blackout blinds, temperature of 28±3 °C, and 50–60% relative humidity. Metal halide lamps were used to maintain light levels when they fell below 200 μmol m^{-2} s^{-1} photosynthetically active radiation (PAR). Seeds were heat treated to prevent pathogenesis at the primary seedling stage through immersing in water at 55 °C for 15 min. Seeds were germinated in a module tray for 2 weeks before being transplanted to soil pits (5 m×5 m×1.25 m, L×W×D) within the glasshouse. Five replicates of each accession were transplanted in east–west rows, at 20 cm intervals, into high nutrient loam-based soil in 2×5 m concrete pits. Plants were irrigated by drip tapes twice a day, to provide a soil water content close to field capacity. Soil top layers were replaced every 2 weeks from the same batch.

Due to the size of the population accessions, planting was staggered at 1–2 week intervals. Accessions were grown in rotations of 12 genotypes at a time, with five biological replicates, four of which were selected for measurement. Plants were measured at 8 weeks old when they were approximately in the mid to late tillering stage. Measurements commenced in July 2017 and ended in October 2017. The elite *O. sativa* variety 'IR64' was used as a reference genotype and planted as a row in every batch (see 'Data analysis' below).

Gas exchange measurements

An IRGA (infra-red gas analyser; Li-Cor 6400XT, Lincoln, NE, USA) was used on the uppermost fully expanded leaf. A light induction programme was used: leaves were dark adapted for 1 h, the sample leaf was then placed in the leaf cuvette and allowed to achieve steady state in darkness before being subject to a photosynthetic photon flux density (PPFD) of 1500 μmol m^{-2} s^{-1}, from in-built red and (10%) blue LED lights, from 0 s to 900 s, reducing to 100 μmol m^{-2} s^{-1} from 900 s to 1200 s. A graphical representation of the induction assay can be seen in Fig. 1. The leaf cuvette conditions were maintained at a block temperature of 30 °C, 400 μmol^{-1} mol^{-1} CO_2, flow rate 500 ml min^{-1}, and 50–65% humidity. Gas exchange data were logged every 10 s. Measurements were collected between 09.00 h and 16.00 h. Chlorophyll fluorescence parameters were collected simultaneously, by applying a single saturating pulse before the application of actinic light to attain F_o and F_m and then at intervals of 60 s following this for the calculation of ϕPSII (PSII operating efficiency in the light), qP (photochemical quenching), and NPQ (non-photochemical quenching: measurement of a photoprotective process that estimates the rate constant for PSII heat loss) (Murchie and Lawson, 2013). Intrinsic water use efficiency (iWUE) was calculated post-data collection as CO_2 assimilation rate (A)/stomatal conductance (g_s). We calculate that vapour pressure deficit (VPD) in the cuvette was ~1.51–2.10 kPa. Saturation or near-saturation was achieved within this time scale. Raw data for A, g_s, and NPQ are shown graphically in Supplementary Fig. S1 as individual replicates and means per accession.

Stomatal density and automated stomatal counting

Stomatal impressions were taken from the same area of the first fully expanded leaf where the IRGA measurements were obtained. A ~1 cm^2 negative impression of the abaxial (basal) and adaxial (upper) leaf surface was taken using fast-drying clear nail polish and adhered to a microscope slide. Impressions were obtained after all other measurements had been taken.

Images were obtained on a Leica DM5000B light microscope at ×40 objective with 10 fields of view per impression. Due to the volume of images (13 110), a bespoke machine learning-based software was created to automatically calculate the number of stomata in each image. The software can reliably identify *O. glaberrima* and *O. sativa* stomata, showing a high correlation (r=0.94; n=540 images per counting method) between software and manual stomatal counts. Our method was based on transfer learning for deep neural networks: we have utilized a pre-trained deep model for the different datasets and adapt it for user-annotated rice stomata samples. Based on the transfer learning approach, we utilize a pre-trained object detection model trained on the

standard COCO datasets (Lin *et al.*, 2014). Since our goal is detecting and classifying stomata, we use the Faster R-CNN model (Ren *et al.*, 2015) as one of the state-of-the-art methods based on deep neural networks. We used the Faster R-CNN model available in Tensorflow with the Inception-V2 architecture (Szegedy *et al.*, 2016) as the base model. Inception-V2 is a variation of Inception-V1, also referred to as GoogLeNet, which was the state-of-the-art architecture at the ImageNet competition in ILSVRC 2014. After loading the pre-trained Faster R-CNN, the last few layers of classification layers are changed to meet the aim of stomata classification and detection. In the next step, the Faster R-CNN with stomata images are trained with different hyperparameters such as learning rate and number of epochs to find out the best parameters to reduce execution time and errors. Further information on our methodology is located in File S1 at the Zenodo repository; Murchie, 2021).

Morphological traits

Plant height, leaf area, and root and shoot dry biomass were taken at 8 weeks post-germination, after the completion of gas exchange measurements. Each plant was dug up and care taken to preserve the root system. The shoot area was measured using a LiCor LI-3100 area meter. The root ball was soaked and carefully washed to preserve root structure, as described by York (2018). The shoot and root material were then placed in a drying oven at 70 °C for 72 h before weighing for dry biomass.

Data analysis

All data analyses were performed using R-Studio (v. 4.0.1).

To reduce the temporal and spatial effects of measuring the accessions in batches, a linear mixed-effects model ('lme4' package, v.1.1-26) was used to calculate best linear unbiased predictions (BLUPs) and predicted means, considering the effects of accession, sowing date, measurement date, location within the glasshouse, and, if relevant, IRGA machine (Supplementary Fig. S2). BLUPs are commonly used to account for the random effects that accompany measuring large populations in fluctuating environments (Robinson, 1991; Merk *et al.*, 2012). The coefficients of the mixed-effects model were also used to estimate broad sense heritability (H^2). All results reported here use the adjusted means data generated from the mixed-effects model. Normality was tested using the Shapiro–Wilk test. A 0.01 α value was used, as the Shapiro–Wilk test tends to report false negatives in sample sizes exceeding 50 individuals. All data for IR64 were found to be normally distributed, whereas 25 out of 57 traits in the *O. glaberrima* panel showed a deviation from a statistically normal distribution. In the results, and *O. glaberrima* descriptive statistics Table 1, non-normal traits report values for median and interquartile range (IQR), whereas normally distributed traits will report mean and standard deviation. A full breakdown of IR64 and *O. glaberrima* normality statistics, box, and distribution plots can be found in Files S2–S4 at Zenodo.

A bespoke Python pipeline was written to identify the data point at 95% of the maximum and extract values within the induction side of a curve (File S5 at Zenodo).

The correlation analyses were completed using a Pearson correlation coefficient in the 'Corrplot' package (v. 0.84), with a correlation significance threshold of $P<0.1005$.

The percentage genetic variation (PGV) was calculated as follows; $[(x_{max}–x_{min})/x)]\times100$. Where x_{max}, x_{min}, and x, respectively, denote the maximum, minimum, and mean values for a trait in the population (Gu *et al.*, 2014). This measure is used to quantify the genetic variation of a trait within the combined population. Values >100% signify where the range is greater than the mean and represent particularly high underlying genetic variation.

Kinetic modelling

Dynamic modelling of A, NPQ, and g_s was performed using a dose–response curve-fitting method, previously used to model stomatal responses by Barratt *et al.* (2021). The 'drc' (v. 3.0) package was used to analyse and extract several useful parameters for both curve induction and reduction responses (Ritz *et al.*, 2015), denoted by $_i$ and $_r$, respectively. The measured parameters are detailed in Table 1 and include curve slope ($_{i/r\ slope}$), lower limit ($_{i/r\ min}$), upper limit ($_{i/r\ max}$), and the time taken to reach a defined percentage of the dependent variable, in this case 10 ($_{i/r\ 10}$), 50 ($_{i/r\ 50}$), and 90% ($_{i/r\ 90}$). A representation of these parameters on A and g_s response curves can be found in Fig. 1. The LL.4 (log-logistic 4-parameter) model was chosen as the best fit for both the A induction, and g_s induction and reduction responses. The LL.3 (log-logistic 3-parameter) model was used for NPQ induction, and the W2.4 (4-parameter Weibull2) model for NPQ relaxation. The comparison of eight different models, followed by Akaike's information criterion analysis, was used to select the best model fit. All 155 accessions were analysed for A induction curves; 24 accessions were removed for g_s induction and g_s reduction curve fitting due to unusable curve measurements.

Due to the volume of data, the best fitting model was selected for the induction and reduction curve for each parameter (g_s, A, and NPQ) and then applied to all data (e.g. for g_s induction, a LL.4 model was applied to all 155 genotypes). To ensure that the model selection process captured the variation that may occur in the population, five genotypes were randomly chosen for model selection and the consensus model was used. The estimated parameters generated from the model (min, max, ed50, and slope) were manually cross-referenced to the raw data, to ensure these outputs closely represented the raw data. To evidence the fit of the selected models to the raw data, we include a table showing the model fit of a randomly selected accession for each parameter (Supplementary Table S3). We note the large SE for $A_{r\ slope}$, probably due to the rapid and steep drop in A that occurs between two data points, so we attribute less confidence in this parameter.

Multivariate and climatic analysis

The multivariate analyses, principal component analysis (PCA), and hierarchical clustering (H-clustering) methods require a complete dataset, with no missing values. Consequently, missing phenotype data were imputed using the missMDA package (v. 1.18). PCA was performed using the FactoMineR (v. 2.4) package, and the H-clustering was performed using the HCPC method. Details of the FactoMineR package and HCPC algorithm can be found in Lê *et al.* (2008).

The PCA and H-clustering of the phenotypic dataset contained all 155 *O. glaberrima* accessions and the IR64 *O. sativa* representative. The analysis included 64 quantitative trait variables and four qualitative variables, namely narrow ecotype, broad ecotype, country of origin, and African region.

Agroecological niche and geographical coordinates of the collection sites for each *O. glaberrima* accession were provided by AfricaRice. We have complete ecological information for all 155 accessions, and geographical coordinates for 105 accessions (Fig. 2A–C). Nineteen variables for temperature and precipitation, at the collection site of each accession, were obtained using the BIOCLIM dataset (Hijmans *et al.*, 2005). Information on the elevation above sea level was obtained using the elevatr package (v. 0.3.1). PCA and H-clustering analysis, and subsequent climate–trait correlations, were performed on the subset of 105 accessions for which we had geographical coordinates.

Table 1. The range of natural variation and broad-sense heritability (H^2) within a population of diverse *O. glaberrima* accessions across dynamic and static traits

Trait	Min	Max	Mean (SD)	Median (IQR)	PGV	Sig.	H^2
Steady state							
A_{max}	17.92	22.42	20.16 (0.87)		22.30	***	0.11
ETR_{max}	104.10	144.30	123.00 (7.60)		32.69	***	0.22
g_{smax}	0.26	0.46	0.34 (0.03)		58.82	***	0.17
$iWUE_{max}$	58.34	71.81		62.63 (3.08)	21.40	***	0.07
NPQ_{max}	1.98	2.30	2.13 (0.06)		14.88	***	0.12
$\phi PSII_{max}$	0.16	0.22	0.19 (0.01)		32.70	***	0.22
qP_{max}	0.43	0.51	0.47 (0.02)		18.71	***	0.17
$Trmmol_{max}$	4.13	5.75	4.81 (0.27)		29.84	***	0.12
VPD_{max}	1.43	1.52	1.47 (0.01)		6.38	***	0.00
Morphological							
Shoot:root	3.36	9.50		5.55 (1.13)	73.22	***	0.12
Shoot biomass	2.36	7.85		4.40 (0.89)	123.33	***	0.14
Shoot area	279.07	1068.97		652.73 (126.27)	118.51	***	0.16
Root biomass	0.36	1.98		0.77 (0.18)	203.75	***	0.23
Plant height	61.79	93.86	78.77 (6.20)		40.72	***	0.18
Adaxial SD	260.16	353.50	314.08 (18.04)		29.72	***	0.21
Abaxial:adaxial	1.05	1.40	1.24 (0.06)		18.87	***	0.15
Abaxial SD	324.31	435.99	388.83 (22.54)		28.72	***	0.21
Dynamic							
$g_{si\ slope}$	–3.90	–1.90		–2.44 (0.38)	78.43	***	0.14
$g_{si\ max}$	0.35	0.49	0.42 (0.03)		31.88	*	0.11
$g_{si\ min}$	0.06	0.14		0.08 (0.01)	112.50	**	0.12
$g_{si\ 10}$	65.50	207.09		108.19 (29.45)	126.83	***	0.18
$g_{si\ 50}$	179.53	324.80		223.71 (48.40)	64.17	*	0.11
$g_{si\ 90}$	477.70	684.41		539.21 (48.40)	37.53	NS	0.08
$g_{si\ rate}$	0.0005	0.0008		0.0006 (<0.01)	45.45	*	0.09
$A_{i\ slope}$	–2.42	–1.50		–1.78 (0.17)	50.84	**	0.10
$A_{i\ max}$	20.44	27.83	23.43 (1.47)		31.53	***	0.20
$A_{i\ min}$	–1.32	–1.09	–1.22 (0.04)		18.85	NS	0.01
$A_{i\ 10}$	49.01	140.97		63.79 (13.96)	136.43	***	0.15
$A_{i\ 50}$	189.56	334.26		217.84 (24.73)	64.97	**	0.12
$A_{i\ 90}$	652.34	849.54		718.78 (52.10)	27.14	NS	0.06
$A_{i\ rate}$	0.03	0.04	0.03 (0.002)		33.33	NS	0.05
$NPQ_{i\ slope}$	–3.48	–2.32		–2.72 (0.22)	42.65	***	0.12
$NPQ_{i\ max}$	2.12	2.44	2.24 (0.05)		14.29	***	0.08
$NPQ_{i\ 10}$	19.57	28.12	23.80 (1.53)		35.75	***	0.11
$NPQ_{i\ 50}$	50.98	56.44	53.65 (1.02)		10.16	NS	0.05
$NPQ_{i\ 90}$	118.70	132.30	125.58 (2.71)		10.82	NS	0.03
$NPQ_{i\ rate}$	0.017	0.02	0.02 (<0.01)		15.00	**	0.07
$g_{sr\ slope}$	6.48	10.49		8.05 (0.80)	49.43	NS	0.07
$g_{sr\ min}$	-0.30	0.02		-0.13 (0.08)	228.57	**	0.11
$g_{sr\ max}$	0.38	0.61	0.48 (0.05)		48.96	**	0.15
$g_{sr\ 10}$	908.91	956.10		922.35 (8.56)	05.11	NS	0.07
$g_{sr\ 50}$	1061.65	1385.05	1208.21 (76.16)		26.77	***	0.26
$g_{sr\ 90}$	1370.52	2116.04		1640.46 (266.31)	44.56	***	0.23
$g_{sr\ rate}$	0.0006	0.0101		0.001 (<0.01)	339.28	***	0.25
$A_{r\ slope}$	-332.60	-322.20		-328.45 (2.15)	3.17	NS	0.01
$A_{r\ min}$	2.88	3.60	3.22 (0.11)		24.05	***	0.11
$A_{r\ max}$	17.09	21.72	19.22 (0.96)		22.36	NS	0.04
$A_{r\ 10}$	901.90	902.20	902.09 (0.04)		0.03	NS	0.03
$A_{r\ 50}$	905.30	905.70		905.45 (0.05)	0.04	NS	0.03
$A_{r\ 90}$	910.60	911.00		910.77 (0.05)	0.04	NS	0.02
$A_{r\ rate}$	0.79	2.36		1.51 (0.09)	103.97	*	0.08

Table 1. Continued

Trait	Min	Max	Mean (SD)	Median (IQR)	PGV	Sig.	H^2
$NPQ_{r\ slope}$	–49.71	–36.27	–41.83 (2.65)		34.35	***	0.21
$NPQ_{r\ min}$	0.56	0.66	0.61 (0.02)		23.00	***	0.16
$NPQ_{r\ max}$	1.99	2.30	2.13 (0.06)		14.55	***	0.10
$NPQ_{r\ 10}$	921.31	923.66	922.60 (0.39)		0.27	***	0.19
$NPQ_{r\ 50}$	946.63	954.76	950.23 (1.41)		0.97	***	0.19
$NPQ_{r\ 90}$	985.20	1008.16	995.50 (4.13)		02.50	***	0.19
$NPQ_{r\ rate}$	0.013	0.020	0.016 (<0.01)		43.75	***	0.14

Normally distributed traits report the trait mean and standard deviation, whilst the median and interquartile range is given for non-normally distributed traits. PGV is the percentage of genetic variation. Sig. refers to the ANOVA test between two mixed-effects models, where the accession is present as an effect in one model and not in another. A significant result suggests that the accession genotype has an effect and therefore the trait is heritable. ***P<0.0001, **P<0.001, *P<0.01

FactoShiny was used to generate summary reports of the PCA and H-clustering analyses on both the phenotype data and climate data; these can be found in Files S6–S9 at Zenodo.

Results

Phenotypic analysis of morphology and steady-state photosynthesis

Significant variation and high levels of PGV were identified between accessions across all morphological, gas exchange, and fluorescence traits measured in this *O. glaberrima* panel (Table 1; Files S2, S4 at Zenodo).

Root biomass, shoot biomass, and shoot area showed a 5-, 3-, and 4-fold variation, respectively. Plant height showed a 1.3-fold variation in PGV. Significant (P<0.001) positive correlations between all plant growth traits were found (Fig. 3A, C).

Even though key steady-state photosynthesis traits showed a relatively narrow distribution (typically between 15% and 40%), both shoot biomass and shoot area showed significant (P<0.01–0.05; Fig. 3A; File S10 at Zenodo), positive correlations to A_{max}, qP_{max}, ETR_{max}, and $\phi PSII_{max}$, providing confidence that steady-state photosynthesis is linked to biomass production. g_{smax} showed an almost 2-fold variation across *O. glaberrima* accessions. PGV for steady-state traits ranged from 6.38% to 58.82% (Table 1), with most traits in the 20–30% range, including key photosynthetic traits. All key steady-state photosynthetic traits showed significant (P<0.0001) positive correlations to one another (Fig. S3). Some unexpected relationships were apparent, for example between NPQ_{max} and VPD_{max}. $iWUE_{max}$ was highly correlated with g_{smax} (Fig. 3B) (and $Trmmol_{max}$) but not A_{max}, indicating stomatal limitation of A.

Stomatal morphology (Fig. 4) did not show a clear relationship with conductance. A relatively modest 1.3-fold accession-dependent variation in the abaxial SD and adaxial SD was observed. The abaxial SD was 1.24-fold greater than the adaxial SD. PGV showed that all SD traits were highly significant (P<0.0001, Table 1) across the *O. glaberrima* accessions, revealing that SD has a genetic basis. However, no association between any SD traits and $iWUE_{max}$, or g_{smax} was detected. Unexpectedly the adaxial SD showed a significant negative correlation to NPQ_{max}, while abaxial SD showed a negative association with plant height (Fig. 3D). The SD ratio, however, showed significant associations with multiple traits (Fig. 3C); A_{max}, ETR_{max}, $\phi PSII_{max}$, qP_{max}, NPQ_{max}, and plant height, the reasons for which are unclear.

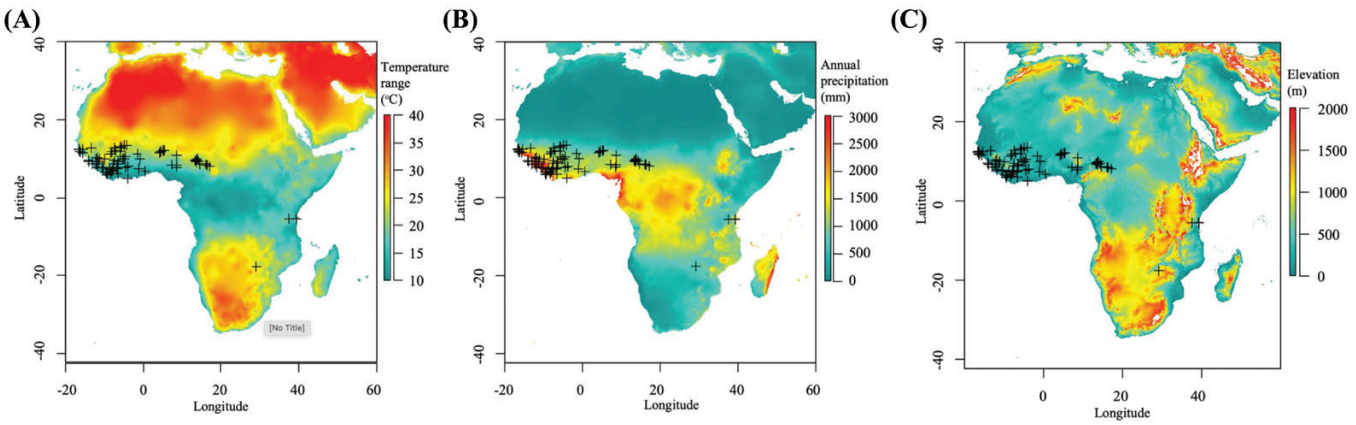


Fig. 2. Map showing the geographical collection locations of *O. glaberrima* accessions used in this study. The annual range of (A) temperature, (B) annual precipitation, and (C) elevation across Africa.

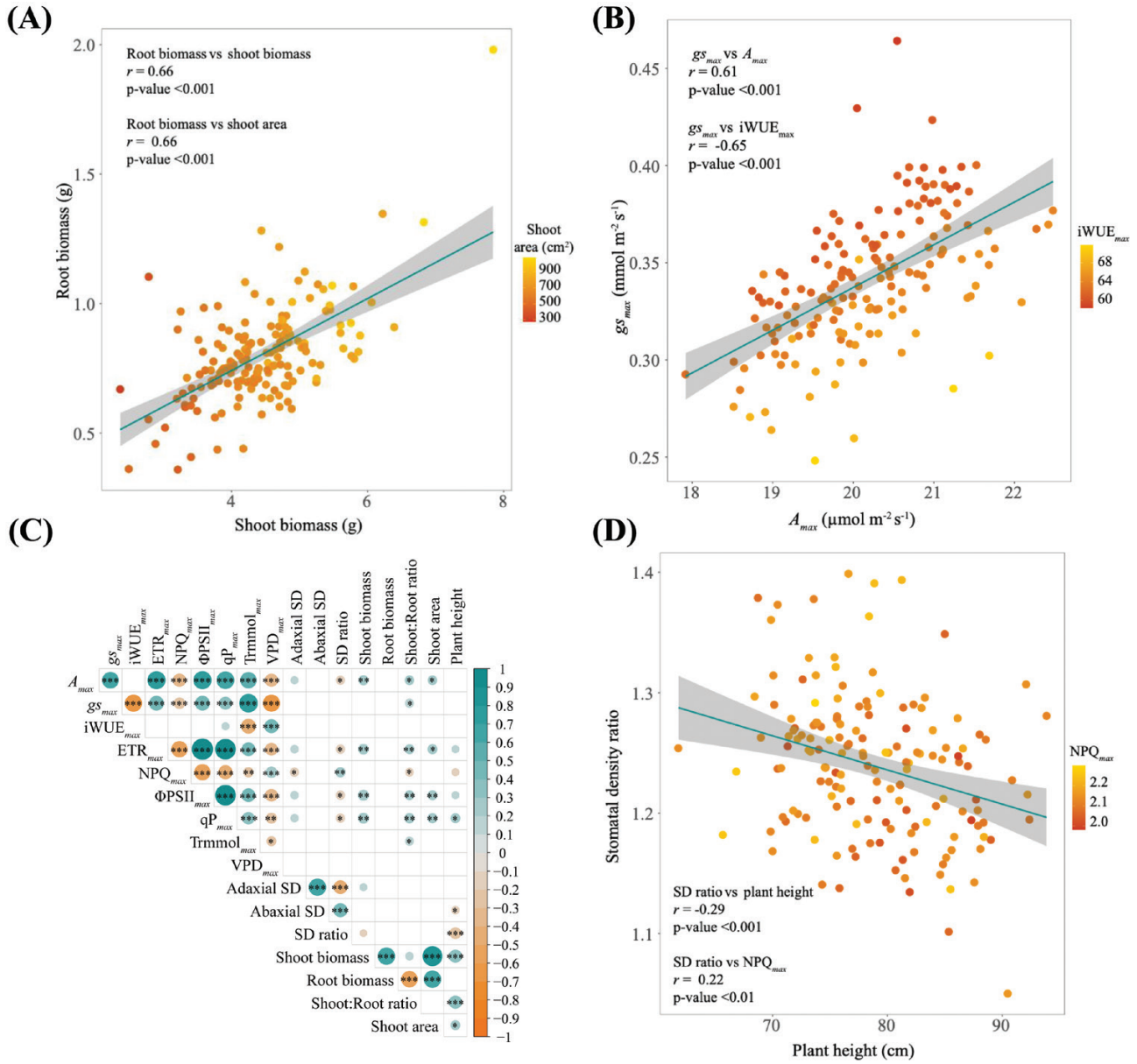


Fig. 3. *O. glaberrima* shows a range of interesting morphological and steady-state photosynthetic trait correlations. The colour gradient shows a second correlation against the *y*-axis variable. (A) Positive correlation between root and shoot biomass; the second correlation shows root biomass and shoot area. (B) The effect of g_{smax} on A_{max} and the second correlation of g_{smax} against $iWUE_{max}$. (C) Pearson correlation matrix showing associations between morphological and steady-state gas exchange traits, filtered to show trait associations at a $P<0.1005$ significance threshold. Correlations are scaled by colour, shown in the right-hand scale bar; asterisks indicate significance between traits (***$P<0.001$, **$P<0.01$, *$P<0.05$). (D) Negative correlation between SD ratio and plant height, while the SD ratio against NPQ_{max} shows a positive correlation.

Phenotypic analysis of dynamic photosynthesis

Dynamic responses are now recognized as important determinants of photosynthetic productivity. Responses of gas exchange, fluorescence, and photoprotection to light shifts were modelled and show significant variation in 29 traits (Table 1, column 'Sig'; Figs 5A–D, 6A, B; Supplementary Fig. S4A–F).

The well-documented divergence between the induction of g_s and A was observed, where a lag in g_s induction and reduction occurs relative to A (Figs 1, 5A–C). The mean upper limit estimates for A induction and reduction curves ($A_{i\ max}$ and $A_{r\ max}$) and g_s induction and reduction ($g_{sr\ max}$ and $g_{si\ max}$) curves were similar ($P<0.0001$, Supplementary Fig. S5A, B) to measured values. The estimated averages for the mean lower limits of the A induction ($A_{i\ min}$), g_s induction and reduction ($g_{si\ min}$ and $g_{sr\ min}$) curves are close to zero (Table 1).

The average time taken to reach 10% of the maximum induction curve was significantly less for CO_2 assimilation ($A_{i\ 10}$) than for g_s ($g_{si\ 10}$), whereas the time taken to reach 50% of the

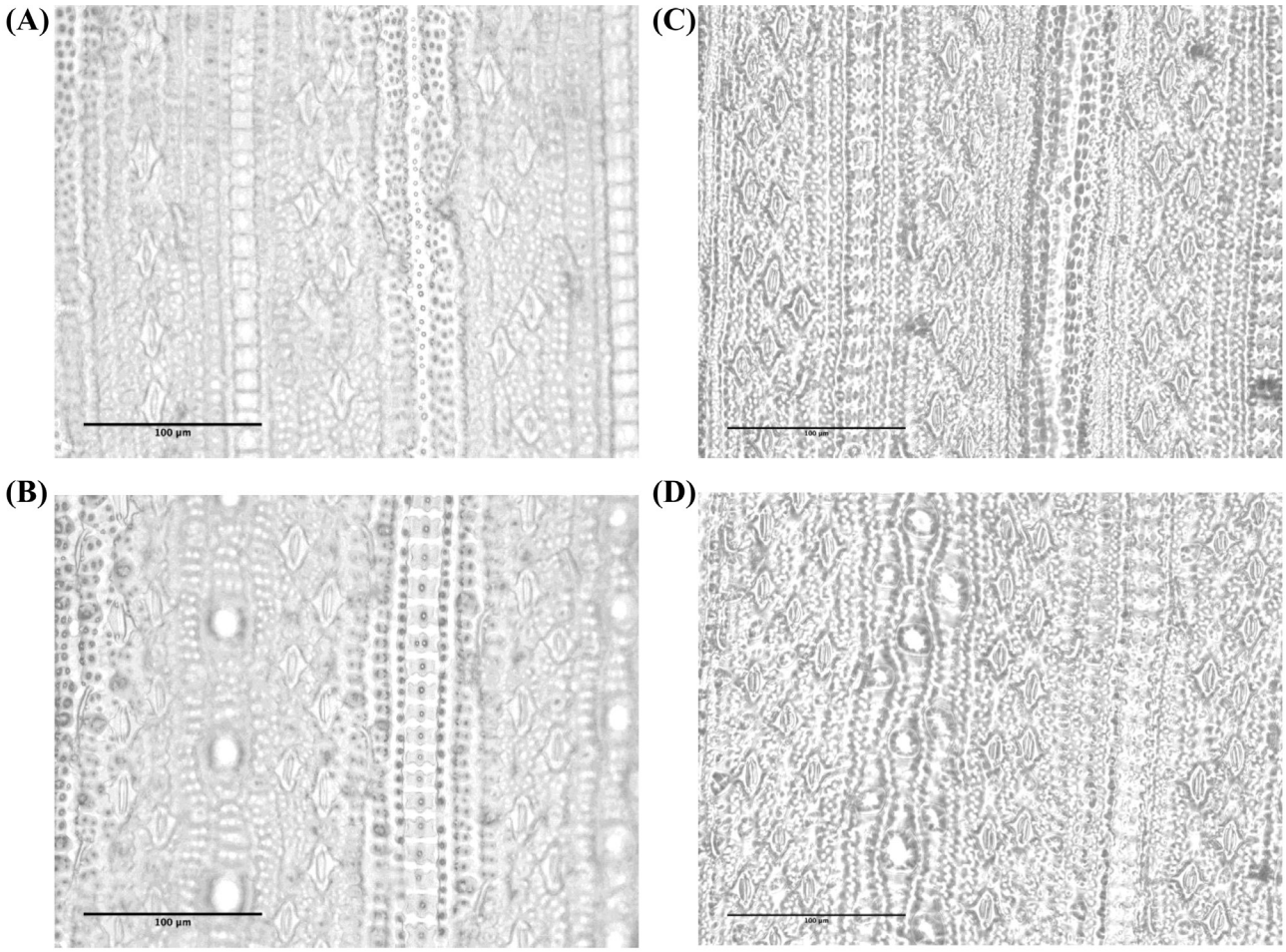


Fig. 4. Microscope images showing examples of the *O. glaberrima* accessions with highest (TOG_14116) and lowest (TOG_5464) recorded SD. These images demonstrate the extent of SD variation in the population and the qualitative correlation between high SD and small stomatal size, (A) TOG_5464; adaxial SD=260 mm^{-2}, (B) TOG_5464; abaxial SD=325 mm^{-2}, (C) TOG_14116; adaxial SD=345 mm^{-2}, (D) TOG_14116; abaxial SD=426 mm^{-2}.

induction curve for $A_{i\ 50}$ and $g_{si\ 50}$ did not significantly differ. However, the average time to reach induction to 90% of the maximum ($A_{i\ 90}$) was significantly longer than that of $g_{si\ 90}$.

Strong interactions between stomatal and CO_2 assimilation indicate co-dependence (Fig. 5E, F). Notably the steepness of the g_s induction slope ($g_{si\ slope}$) highly correlates with key induction traits $g_{si\ 90}$, $g_{si\ rate}$, g_{smax}, $A_{i\ slope}$, $A_{i\ rate}$, $_iWUE_{max}$, NPQ_{islope}, NPQ_{imax}, and NPQ_{i50}. $g_{si\ rate}$ was also strongly correlated to many dynamic induction traits; g_{smax}, $g_{si\ 10}$, $g_{si\ 50}$, $g_{si\ 90}$, A_{max}, $A_{i\ 90}$, and $A_{i\ rate}$.

The rate of photosynthetic induction in high light was associated with rates of decline in low light (Fig. 5). $g_{si\ slope}$ versus $g_{sr\ slope}$, $gs_{i\ rate}$ versus $gs_{r\ rate}$, and $A_{i\ rate}$ versus $A_{r\ rate}$ were significant, suggesting that accessions which exhibited rapid stomatal opening also close at a greater rate (Fig. 5C, D; Supplementary Fig. S5). Further, traits associated with rapid stomatal closure, $g_{sr\ slope}$, $g_{sr\ 10}$, $gs_{r\ 50}$, and $g_{sr\ 90}$ showed significant associations with enhanced $iWUE_{max}$.

Like steady-state traits, A and g_s dynamics were also linked to plant biomass and morphology in these data, further supporting the role of photosynthesis in determining growth. A greater $A_{i\ rate}$ was positively correlated with total plant and shoot biomass. $A_{r\ rate}$ showed positive associations with total plant biomass, shoot biomass, shoot:root ratio, and shoot area. $A_{r\ slope}$ had negative associations with shoot biomass, shoot:root ratio, and plant height, while a more rapid $A_{r\ 90}$ was correlated to a greater shoot biomass, shoot:root ratio, and shoot area. $g_{si\ rate}$ showed positive associations with total plant biomass, shoot biomass, root biomass, and shoot area.

Again there were fewer links with stomatal morphology; a significant negative association was identified between the SD ratio and $A_{r\ rate}$. Only upper leaf SD was also found to have positive relationships to $g_{si\ 50}$ and $A_{r\ rate}$, and a negative relationship to $g_{sr\ min}$.

Non-photochemical quenching dynamics

NPQ is of particular interest here because it showed multiple relationships with photosynthesis and biomass. The model estimation of the NPQ induction and relaxation curve upper

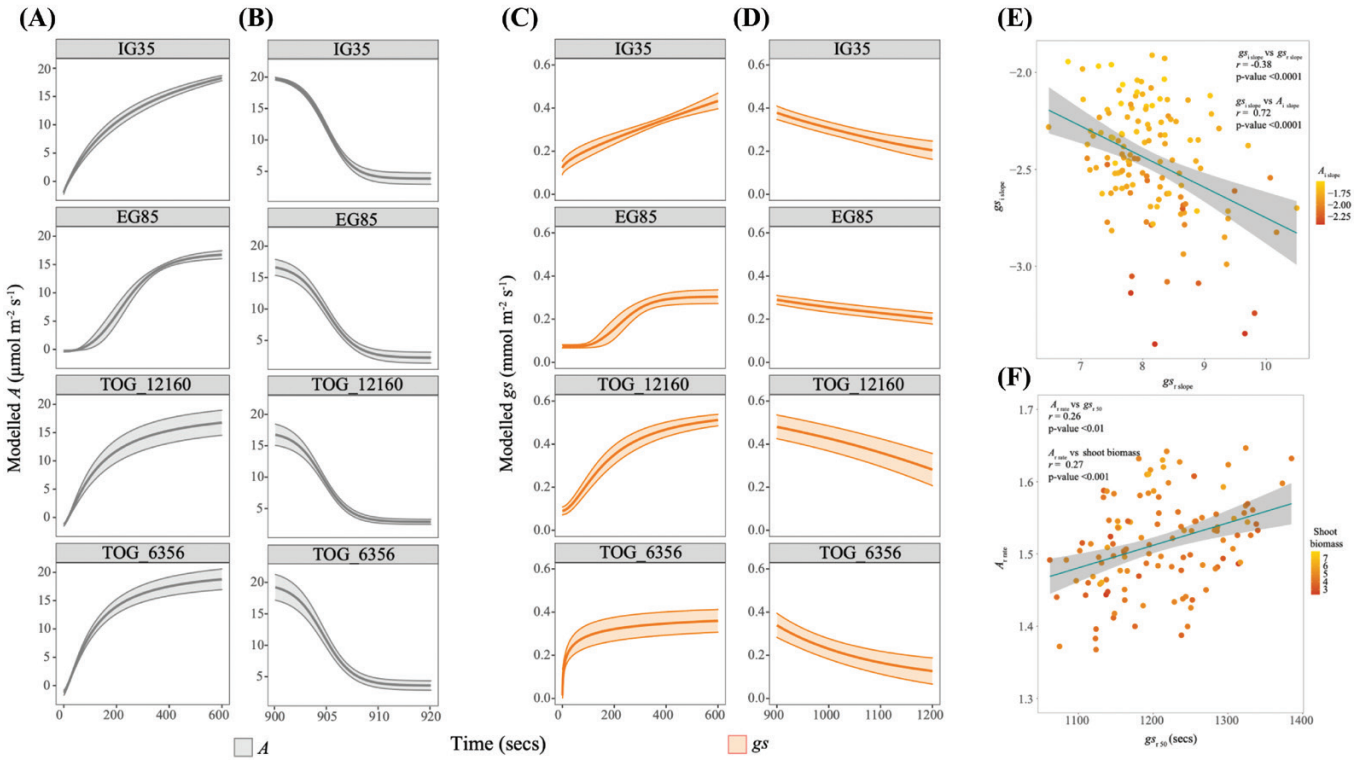


Fig. 5. Demonstrating the variation of (A, B) *A* and (C, D) g_s dynamic responses to light intensity changes within the *O. glaberrima* population using four example accessions. IG35 and TOG_6356 were used as example of a 'slow' and 'fast' responding accession, respectively, whereas EG85 and TOG_12160 are used to demonstrate the intermediate gradient of responses in the population. (E) $g_{si\ slope}$ was correlated with $g_{sr\ slope}$ and $A_{i\ slope}$; during induction, with a greater negative value indicating a steeper slope; this relationship is reversed for the decrease. (F) $A_{r\ rate}$ shows positive associations with $g_{sr\ 50}$ and shoot biomass.

limit ($NPQ_{i\ max}$ and $NPQ_{r\ max}$) was close to the measured value for NPQ_{max}, providing confidence in the method (Table 1; Supplementary Fig. S5C)

We observed limited significance between the kinetics of NPQ relaxation and kinetics of *A*. Importantly, there was a significant negative correlation between the *A* reduction curve lower limit ($A_{r\ min}$) achieved under 100 PPFD, and $NPQ_{r\ slope}$, $NPQ_{r\ 50}$, and $NPQ_{r\ 90}$, suggesting that *A* maintains a higher value under low light conditions when NPQ relaxes rapidly (Kromdijk *et al.*, 2016). Additionally $NPQ_{i\ slope}$, $NPQ_{i\ rate}$, and the time taken to induce 90% of the maximum ($NPQ_{i\ 90}$) strongly correlated with A_{max}.

Speed of induction was not closely related to NPQ capacity: only the time taken to reach 90% of the NPQ curve upper limit ($NPQ_{i\ 90}$) positively correlated to a greater NPQ_{max}. Like gas exchange traits, NPQ induction and relaxation traits were positively correlated ($NPQ_{i\ slope}$ versus $NPQ_{r\ slope}$ and $NPQ_{i\ rate}$ versus $NPQ_{r\ rate}$).

Interestingly, NPQ and g_s dynamic traits also showed numerous significant correlations. $g_{si\ slope}$ significantly correlated with $NPQ_{i\ slope}$, $NPQ_{i\ 10}$, $NPQ_{i\ 50}$, and $NPQ_{i\ 90}$. $g_{si\ rate}$ was positively related to $NPQ_{i\ slope}$ and $NPQ_{i\ 90}$. Accessions with steeper $g_{sr\ slope}$ were also found to have a greater $NPQ_{r\ rate}$ (Fig. 6D). These associations highlight a complex interdependent relationship between g_s, *A*, and NPQ and the recent link noted between underlying control of NPQ by PsbS and the dynamics of stomatal conductance and gas exchange (Fig. 6C, D) (Kromdjik *et al.*, 2016; Glowacka *et al.*, 2018).

Further NPQ relaxation traits were related to morphological and SD traits, indicating that photoprotection has a role in determining growth. $NPQ_{i\ slope}$ and $NPQ_{i\ 90}$ (Fig. 6E) showed negative correlations with shoot biomass and shoot area. $NPQ_{i\ rate}$ positively correlated to shoot biomass and shoot area. A more pronounced set of associations was observed during NPQ relaxation; shoot biomass and shoot area, respectively, showed negative correlations to $NPQ_{r\ slope}$, $NPQ_{r\ 50}$, and $NPQ_{r\ 90}$ (Fig. 6F), and positive correlations to $NPQ_{r\ 10}$ and $NPQ_{r\ rate}$. Root biomass showed a similar, but not as strong, association with $NPQ_{r\ slope}$, $NPQ_{r\ 10}$, $NPQ_{r\ 90}$, and $NPQ_{r\ rate}$.

Trait and ecological comparison between O. glaberrima *and* O. sativa

It is informative to compare the *O. glaberrima* trait variation with that of the elite Asian *O. sativa* cultivar, IR64 (File S3 at Zenodo; Table 1) even though caution should be observed using just one genotype. We highlight the slower induction rates of photosynthesis of IR64.

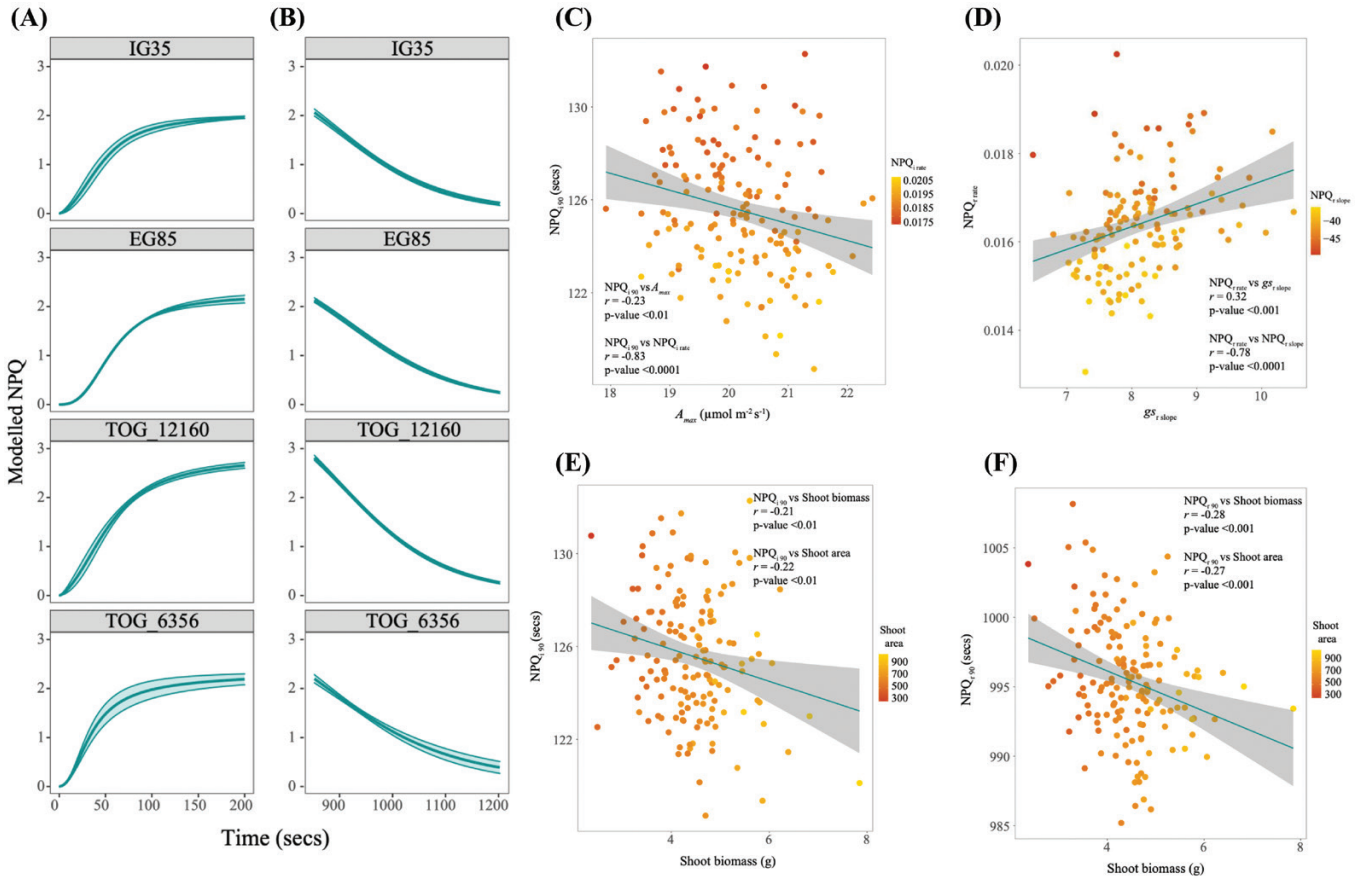


Fig. 6. Demonstrating the variation of NPQ (A) induction and (B) relaxation responses to light intensity changes within the *O. glaberrima* population using four example accessions, as explained in Fig. 5. (C) Negative correlations were identified with $NPQ_{i\,90}$ against A_{max} and $NPQ_{i\,rate}$. (D) $NPQ_{r\,rate}$ showed a positive relationship to $g_{sr\,slope}$, where a high value indicates a steeper slope and a negative correlation between $NPQ_{r\,rate}$ and $NPQ_{r\,slope}$, where for this model a more negative value suggests a steeper relaxation slope. $NPQ_{i\,90}$ (E) and $NPQ_{r\,90}$ (F) showed associations with both shoot biomass and shoot area.

IR64 had a slightly smaller shoot than *O. glaberrima* but a greater root biomass, reflected in the lower shoot:root ratio of IR64, suggesting a greater investment in roots. IR64 height was lower. IR64 displayed a greater SD on the abaxial leaf side than *O. glaberrima*, and IR64 had a lower SD ratio.

IR64 did not differ from *O. glaberrima* for A_{max} and NPQ_{max}. However, average ETR_{max} and $\phi PSII_{max}$ were higher in IR64. IR64 showed a slightly lower g_{smax} and greater $iWUE_{max}$ in comparison with *O. glaberrima*. The latter is likely to be a direct result of the higher levels of g_{smax} observed in *O. glaberrima*. Clear differences were found in dynamics of A, g_s, and NPQ between the two species.

During induction, IR64 was significantly slower than *O. glaberrima* for $g_{si\,10}$, $g_{si\,50}$, $A_{i\,10}$, $A_{i\,50}$, $A_{i\,90}$, $A_{i\,rate}$, $NPQ_{i\,10}$, and $NPQ_{i\,50}$ (Supplementary Fig. S4). This implies that IR64 had a longer g_s and NPQ lag phase. The initial rapidity of the g_s induction curve may facilitate the significantly faster A response observed in *O. glaberrima*, suggesting that *O. glaberrima* may be able to respond better to the onset of high light than IR64. During the decrease, IR64 and *O. glaberrima* did not significantly differ for $g_{sr\,10}$, $g_{sr\,50}$, $g_{sr\,90}$, $g_{sr\,rate}$, $A_{r\,10}$, $A_{r\,90}$, $A_{r\,rate}$, $NPQ_{r\,10}$, and $NPQ_{r\,rate}$. IR64 was found to have a faster reduction response for $A_{r\,50}$, $NPQ_{r\,50}$, and $NPQ_{r\,90}$ in comparison with *O. glaberrima* (Table 1 in comparison with File S3 at Zenodo).

During the multivariate analyses, we observed that *O. sativa* IR64 values cluster separately from *O. glaberrima* for both ecology and country of origin. This can be seen most clearly when plotting principal components (PCs) 1 and 3, where the two species cluster distinctly for the Asian country of origin and paddy field ecology (Fig. 8C).

Impact of country of origin and ecotype on O. glaberrima *trait adaptation*

An important aspect of *O. glaberrima*'s novelty is the independent evolution to *O. sativa* and adaptation to the variable African environment. We used PCA and H-clustering to explore natural trait variation and the adaptive effect of environmental climatic variables. Here we identify phenotypic trends which cluster according to country and environment, indicating adaptation and possibly variation in growth strategy.

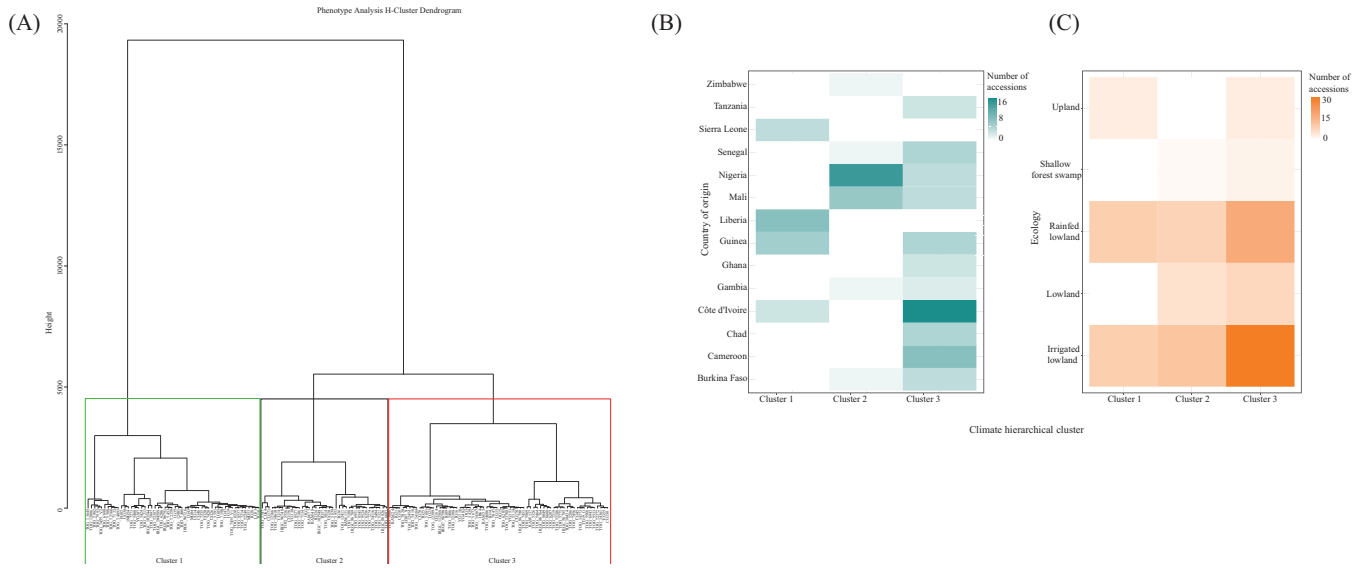


Fig. 7. Hierarchical clustering of 155 *O. glaberrima* accessions (A) for 64 phenotypic traits and the frequency of accessions for each country of origin (B) and ecological niche (C) in the clades (1–3) identified in the climate hierarchical clustering analysis.

The PCA and H-clustering were separated into two grouped analyses for phenotypic and climatic variables. For the PCA of phenotypic traits 12 PCs were selected as they explain 95% of the variance (Supplementary Fig. S6). The H-clustering analysis identified three clusters (Fig. 7A) with common sources of trait variation (File S7 at Zenodo). The accessions in cluster 1 are characterized by a slow g_s reduction time ($g_{sr\ 10/50/90}$), rapid A and NPQ induction time ($A_{i\ 50}$ and $NPQ_{i\ 90}$), steep A reduction curve ($A_{r\ slope}$), rapid A reduction time ($A_{r\ 50/90}$), high values for g_{smax}, $A_{r\ rate}$, $NPQ_{i\ rate}$, and shoot:root ratio, and low values for root biomass, VPD_{max}, and $iWUE_{max}$. Accessions present in cluster 2 demonstrate g_s reduction curves with a steep slope and rapid reduction times ($g_{sr\ 10/50/90}$), high trait values for NPQ_{max}, VPD_{max}, and $iWUE_{max}$, and low values for plant biomass, shoot biomass, shoot area, g_{smax}, ETR_{max}, and $\phi PSII_{max}$. Accessions in the largest group, cluster 3, show high trait values for total biomass, shoot biomass, shoot area, root biomass, and A_{max}, low levels of NPQ ($NPQ_{r\ min}$) under reduced light (100 PPFD), and rapid g_s reduction time ($g_{sr\ 50/90}$). Cluster 3 is the group where IR64 can be found, and it consists mostly of lowland type accessions.

Adaptation to different environments was explored during the multivariate analyses. In Fig. 8A and B, axes PC1 and 2 are shown overlaid with ecological niche and country of origin. *Oryza glaberrima* accessions cluster separately dependent upon their ecological origin, in particular upland or lowland (Fig. 8B). Accessions from lowland-type ecologies dominate, though it is still clear that upland and lowland show trait differences. Accessions also show a high degree of trait variation due to countries of origin that have contrasting climates (Fig. 8A). For example, distinct clustering can be seen between landlocked Burkina-Faso, which borders the Sahara, and coastal Gambia. A categorical analysis was performed to establish if the accessions that occupy each cluster of the H-clustering analysis share similar origins (Supplementary Fig. S7B, C). While there is no obvious relationship, a greater proportion of upland accessions occupy cluster 1, whereas a large proportion of lowland accessions are present in cluster 3 (Supplementary Fig. S7C).

The diversity of climates and elevations (Fig. 2A–C) are likely to have directly impacted trait adaptation and resilience. A PCA focused on climatic traits explored the relationship between climate and phenotype. The first four PCs explain 90% of trait variation in the population (File S8 at Zenodo). H-clustering identified three distinct clusters of accessions with common sources of variation in climatic variables (Files S9, S11 at Zenodo). A categorical analysis of ecological niche and country of origin for the accessions present in each cluster showed a clear distinction of climate-based clustering due to country of origin (Fig. 7C). Cluster 1 contains all accessions that originate from the neighbouring countries of Liberia and Sierra Leone. Cluster 2 contains all accessions from Zimbabwe and most accessions originating from Nigeria. Cluster 3, which contains the largest number of accessions, contains all individuals originating from Cameroon, Chad, Ghana, and Tanzania, and the majority of accessions from Côte d'Ivoire and Senegal.

With the extensive phenotypic and climatic variables reduced to a small number of components, we completed a correlation analysis between the phenotypic and climatic trait PCs to identify groups of climatic drivers on trait adaptation. A significant positive association was identified between trait PC1 and climatic PC4 (r= –0.20, P<0.05; Fig. 8D), suggesting that key traits contributing to phenotypic trait PC1, which includes photosynthetic traits and shoot biomass, have adapted in response to precipitation-related variables that contribute to climate PC4 loadings. Other significant associations were identified between phenotype PC5 and climate PC4 (r=0.25,

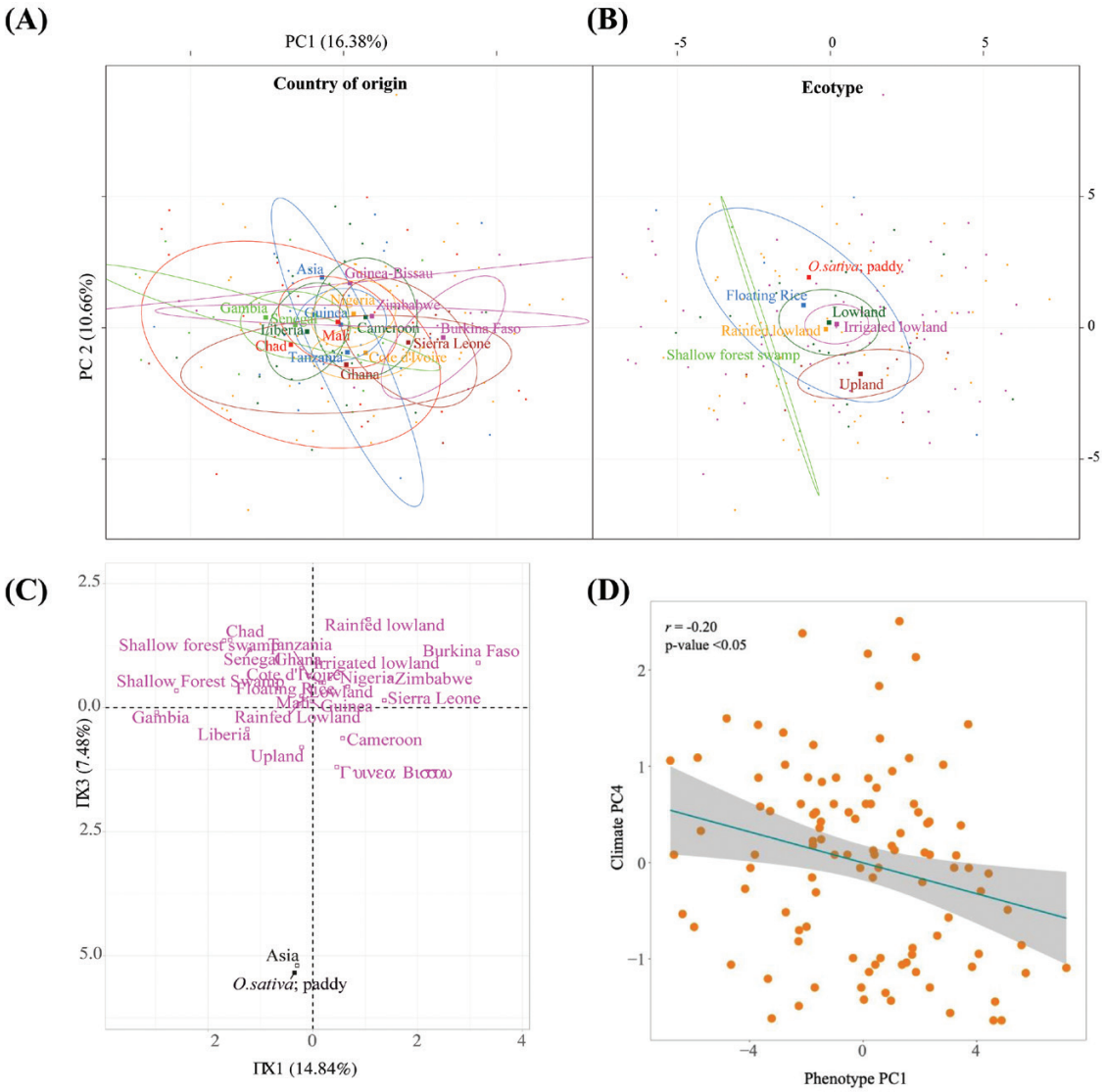


Fig. 8. Graphical PCA outputs; the phenotypic PCs 1 and 2 are overlaid with 95% confidence ellipses for the *O. glaberrima* accessions; (A) country of origin and (B) ecotype categorical variables. (C) PCs and 3 show the separate clustering of *O. sativa* (IR64), based on country of origin and ecotype categories, from *O. glaberrima*. (D) PC1 from the phenotypic traits data PCA was found to be a function of PC4 from the climatic data PCA analysis.

$P<0.05$), phenotype PC8 and climate PC2 ($r=0.24$, $P<0.05$), and phenotype PC11 and climate PC3 ($r=0.25$, $P<0.05$) File S12; Supplementary Fig. S8).

Discussion

Crop production in future climates has the challenge of increasing productivity whilst retaining resilience. To do so, optimizing interactions and trade-offs between carbon assimilation, photoprotection, and water loss will be essential. However, we do not yet have complete understanding of the genetic basis of the co-regulation of the interlinked processes and components (light harvesting, photoprotection, electron transport, carbon assimilation, and stomatal conductance) involved. Recent progress shows that crop productivity and WUE are only partly dependent upon 'steady-state' maximum values of A_{max} and g_{smax}. SD, stomatal conductance, and photoprotection dynamics have been identified as critical traits to optimize carbon assimilation and minimize abiotic stress (Kromdijk *et al.*, 2016; Caine *et al.*, 2019; Faralli *et al.*, 2019). However, elite gene pools may be genetically narrow and poorly adapted to challenging environmental conditions. Attention is increasingly focused upon underutilized crop species and wild relatives as a source of genetic diversity to improve resilience in commercial species (Draic *et al.*, 2011). Whilst the variation for photosynthesis induction has been partly characterized in *O. sativa*, this is not true of *O. glaberrima* (Acevedo-Siaca et al., 2020, 2021). The *O. glaberrima* association panel used here was developed as a resource for crop improvement, which may have diversity not

available in *O. sativa* (Agnoun *et al.*, 2012). For the first time, a comprehensive analysis of photosynthesis- and morphology-related traits has been completed in *O. glaberrima*. Our novel approach uses a large pool of accessions, with a large range of heritable natural variation to explore the natural variation and relationships in these traits. While we cannot here make a meaningful comparison between *O. glaberrima* and *O. sativa*, we observed key differences, with the former showing faster photosynthesis induction. This may be an indication of adaptation to drier soils and air generally, requiring faster opening and closure of stomata (Lawson and Vialet-Chabrand, 2019).

Here we have described extensive natural variation in *O. glaberrima* for steady-state, induction, and relaxation/reduction responses for A and g_s. This suggests underlying genetic diversity to these traits that could be identified and exploited. We identified indications of heritability (H^2) and underlying genetic variation (PGV) in many of these traits (Table 1). Trait heritability values were comparable with estimates of similar traits from maize (Choquette *et al.*, 2019), but they are marginally lower than those previously demonstrated in *O. sativa* (Qu *et al.*, 2017), though a strong genetic component is still indicated. A calculation of heritability using genomic data would provide a more accurate estimation (Zhu and Zhou, 2020). This would be useful when selecting traits for genetic introgression or characterization. The large number of accessions used here (155) permits a statistical comparison that was not possible in related studies on dynamic photosynthesis in *O. sativa* where fewer lines were analysed (Acevedo-Siaca *et al.*, 2020).

A global PCA and clustering analysis showed a distinction between clusters of high biomass (cluster 3), low biomass (cluster 2), and low root biomass (cluster 1). The fast g_s decrease, low g_{smax}, high NPQ_{max}, and high $iWUE_{max}$ of cluster 2 would suggest a conservative type geared toward water conservation, whilst the high total biomass of cluster 3 is consistent with a fast growth type displaying a rapid g_s decrease, low NPQ, and a higher A_{max}. The association of cluster 3 with wetter lowland environments is consistent with higher productivity. We therefore see a general consistency in these two clusters with photosynthetic, water use, and biomass production 'strategy'. It is also notable that steady-state A_{max} correlates well with biomass, suggesting that capacity for higher photosynthesis is still important. Increases in photosynthetic capacity are known to improve light responses in rice (Sun *et al.*, 2016).

Extensive natural variation identified in dynamic photosynthetic traits

In recent years, there has been a shift in photosynthesis-related research towards dynamic responses in place of steady-state values. It is now recognized that irradiance fluctuations in field conditions, and the ability of stomatal and photosynthetic responses to respond instantaneously, can substantially affect plant productivity (Taylor and Long, 2017). To enable greater productivity in dynamic environments such as a crop canopy, one would anticipate that all components of photosynthesis would respond rapidly to 'track' light closely. Each component has a different effect; thus, fast activation of the Calvin cycle and CO_2 assimilation during induction is beneficial, while rapid reduction of NPQ and fast stomatal closure at transition to low light enable the attainment of improved CO_2 efficiency and iWUE at low light.

It is clear that we see some independence of dynamic traits, but interesting associations appear which indicate a link with biomass. Recent research suggests that major yield gains can be made by enhancing photoprotection capacity and NPQ dynamic responses (Kromdijk *et al.*, 2016; Hubbart *et al.*, 2018). Rapid NPQ relaxation can remove the limitation on quantum yield of CO_2 assimilation, allowing a quicker recovery of photosynthetic efficiency upon A reduction (Kromdijk *et al.*, 2016; Murchie and Ruban, 2020). Our findings support this: NPQ relaxation dynamics were the only group found to have ubiquitous associations with increased shoot biomass and area. Notably, we also observed that values for A under low light were greater in those accessions that exhibited rapid NPQ relaxation and those that have lower NPQ capacity under low light ($NPQ_{r\ min}$). It is also hypothesized that faster induction of CO_2 assimilation may reduce the need for photoprotection during induction (McAusland and Murchie, 2020); however, we found no association between A induction traits and NPQ dynamic or steady-state values. We did find that faster NPQ induction is associated with greater photosynthetic capacity, shoot area, and biomass.

Whilst no associations were identified between NPQ and A reduction dynamics, we found strong positive correlations between the speed of g_s and NPQ dynamics. This may highlight the importance of the key NPQ protein, PSII subunit S (PsbS), on stomatal conductance, as shown by Głowacka *et al.* (2018), whereby PsbS overexpression, which increases both NPQ capacity and NPQ dynamic rate (Kromdijk *et al.*, 2016; Głowacka *et al.*, 2018; Hubbart *et al.*, 2018), also reduces the extent of stomatal opening in tobacco. This may be reflected here by the negative correlation between NPQ_{max} and g_{smax}, also that the g_s induction rate was lower when NPQ induction was faster. This highlights the need to further explore the associations between NPQ and g_s dynamics: these have not been elucidated although there is a general principle that limitations imposed by g_s or Rubisco activation state would result in a further reduction of electron transport and an enhanced NPQ. We suggest that in *O. glaberrima* NPQ may be a major player in both g_s and A reduction dynamics. Akin to the relationship between A and g_s, there is a trade-off in NPQ as it reduces photosynthetic quantum yields under low irradiance.

No association was identified between the water use-related traits, g_s and $iWUE_{max}$, and SD; this may be because the variation was less than that needed to produce changes in gas exchange traits (Caine *et al.*, 2019; Mohammed *et al.*, 2019). It is also possible that this highlights the importance of stomatal size

and morphology, rather than density, on these traits. Smaller stomata have been shown to have improved WUE, g_{smax}, and dynamics (Drake *et al.*, 2013; Dittberner *et al.*, 2018; Lawson and Vialet-Chabrand, 2019; Chatterjee *et al.*, 2020). However, the positive correlations we identified between SD ratio, NPQ_{max}, and the level of NPQ achieved under low light ($NPQ_{r\ min}$) is unusual. The significant negative association between SD ratio and $A_{r\ rate}$ has no direct interpretation but may indicate that the SD ratio is a trait worthy of further work. Upper leaf SD had positive relationships to $g_{si\ 50}$ and $A_{r\ rate}$, and negative to $g_{sr\ min}$, also indicating that distinction between the leaf surfaces may be important.

Understanding the interplay of photoprotective, stomatal, and assimilation dynamics should include detailed morphological characterization (Ohsumi *et al.*, 2007; Drake *et al.*, 2013; McAusland *et al.*, 2016), together with the associated mesophyll conductance (Campany *et al.*, 2016; Deans *et al.*, 2019). The proportion by which photosynthetic dynamics are limited by stomata or biochemistry seems to be species dependent (Tinoco-Ojanguren and Pearcy, 1993; Taylor and Long, 2017; De Souza *et al.*, 2020). *Oryza sativa* photosynthetic induction has been shown to be predominantly limited by biochemistry (Acevedo-Siaca et al., 2020, 2021), and the same assumption might be extended to *O. glaberrima* due to a similar genomic composition (Stein *et al.*, 2018); however, we conclude from our data that stomatal limitations may be more pronounced in *O. glaberrima*.

Accessions have adapted to variable ecological and environmental regimes in different countries

No comprehensive studies exist that tease apart the ecological and environmental variables that correlate with specific trait adaptation in *O. glaberrima*. This information is useful from an evolutionary perspective but may be essential in the selection of cultivars for abiotic stress tolerance varieties and trait-related genetic characterization.

Of note, we identified a significant association between the climate PC4 and phenotype PC1 (Fig. 8D; File S12 at Zenodo). This relationship suggests that key photosynthetic traits contributing to PC1 have adapted in response from climatic pressures associated with PC4, such as elevation and the combined effect of temperature and precipitation. However, these are broad observations for climatic–trait correlations across the African continent, lacking resolution that can be seen in studies on a discrete geographical area (Wolfe and Tonsor, 2014).

For the selection of abiotic stress tolerance-adapted cultivars, the H-clustering analyses would be of particular use, as this generated three distinct clades of *O. glaberrima* accessions stemming from similar climatic and phenotypic variables. Furthermore, the climatic H-clustering demonstrated clear grouping of accessions due to country of origin (Fig. 7B), suggesting that a higher resolution analysis of environmental effect on trait adaptation would be beneficial.

We identified adaptation based upon ecotype in the PCA (Fig. 8B), supporting a known distinction between *O. glaberrima* upland and lowland phenotypes (Ghesquière, 1997). However, there is no comprehensive description in the literature of the physiological differences that contribute to these ecotypes. Though due to the unequal representation of accessions from each ecological niche in this analysis, it is difficult to obtain a clear indication of the effect of ecotype on trait adaptation.

The environmental analysis completed here produces useful information of accessions displaying similar phenotypic qualities because of environmental adaptation. Equally, this also highlights the requirement for a dedicated study to truly elucidate the environmental and ecological trait adaptation of *O. glaberrima*, utilizing equally represented accessions from a range of ecotypes and assessing physiological adaptation to climatic variables at a range of spatial scales.

Conclusions

Here, we have demonstrated that *O. glaberrima* has broad, heritable natural variation in a range of important traits, which are likely to aid in the improvement of *O. sativa*. This is the first study to describe photosynthetic, photoprotection, and dynamic traits in *O. glaberrima*, the size of which is not matched in panels of *O. sativa* accessions. The phenotyping efforts compiled here will provide a basis for the identification of interesting traits for physiology research, aid in the selection of accessions for crop improvement efforts, and provide information for genetic characterization.

Supplementary data

The following supplementary data are available at *JXB* online.

Fig. S1. Original, un-fitted data for induction and reduction of CO_2 assimilation (A), stomatal conductance (gs) and NPQ vs time.

Fig. S2. Correlations between the best linear unbiased predictor (BLUP) values and the original mean.

Fig. S3. Correlation matrix of all phenotypic traits measured.

Fig. S4. Modelled curves for two extreme *O. glaberrima* accessions and *O. sativa* IR64, plotted on a log scale.

Fig. S5. Linear regression plots showing strong positive correlations between the actual measurement vs modelled estimate values.

Fig. S6. Plots showing the scree plot and trait loadings for the phenotypic data PCA analysis.

Fig. S7. H-clustering dendrogram of 105 *O. glaberrima* accessions and frequency plots generated from the H-clustering analysis of the phenotypic data from 155 *O. glaberrima* accessions.

Fig. S8. Correlation matrix for all phenotypic and climatic data, alongside their principal components.

Table S1. List of parameter abbreviations, definitions and units of measurement.

Table S2. List of *O. glaberrima* ID codes, country of origin and ecology.

Table S3. Estimated model outputs.

Acknowledgements

We would like to acknowledge the life and scientific rigour of Dr Pracha Treeingtong. Pracha sadly passed away upon the commencement of this manuscript. Our heartfelt condolences and thoughts are with Pracha's friends and family. We would like to thank Dr François Sabot and Dr Philippe Cubry from RICE, CIRAD, Montpelier for their hard work in the development of this *O. glaberrima* germplasm as a resource. We would also like to thank Dr Marie-Noelle Ndjiondjop from AfricaRice for providing us with the ecological information for each *O. glaberrima* accession and her ongoing dedication to its research. We would also like to thank the software developer, Matthew Hartley, who wrote the Python code 'DataProcessor' and helped to develop loops in R-Studio to manage the large volume of data produced in this study.

Author contributions

EHM, SM, and RS: conceptualization; SBC with input from PT: investigation, data curation, formal analysis; SBC with input from JF, HS, and PT: methodology; SBC with input from EHM: writing—original draft; RS, SM, and EHM: funding acquisition; SBC, EHM, SM with input from JF, RS: writing—review and editing; SM, EHM, and RS: project administration.

Conflict of interest

The authors declare no conflicts of interest.

Funding

SC is in receipt of a BBSRC PhD scholarship as part of the Nottingham Doctoral Training Partnership. EM receives funding from the Biotechnology and Biological Sciences Research Council [grant no. BB/R004633/1].

Data availability

Data supporting the findings of this study are available within the paper and within its supplementary data published online. The data Files (S1–S12) are deposited at the Zenodo repository: https://doi.org/10.5281/zenodo.5555930; Murchie, 2021. Any other data are available from the corresponding author upon request.

References

Acevedo-Siaca LG, Coe R, Quick WP, Long SP. 2021. Variation between rice accessions in photosynthetic induction in flag leaves and underlying mechanisms. Journal of Experimental Botany **72**, 1282–1294.

Acevedo-Siaca LG, Coe R, Wang Y, Kromdijk J, Quick WP, Long SP. 2020. Variation in photosynthetic induction between rice accessions and its potential for improving productivity. New Phytologist **227**, 1097–1108.

Agnoun Y, Biaou SSH, Sié M, Vodouhè RS, Ahanchédé A. 2012. The African rice *Oryza glaberrima* Steud: knowledge distribution and prospects. International Journal of Biology **4**, 158.

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.

Barratt GE, Sparkes DL, Mcausland L, Murchie EH. 2021. Anisohydric sugar beet rapidly responds to light to optimize leaf water use efficiency utilizing numerous small stomata. AoB Plants **13**, doi: 10.1093/aobpla/plaa067

Bimpong IK, Serraj R, Chin JH, *et al.* 2011. Identification of QTLs for drought-related traits in alien introgression lines derived from crosses of rice (*Oryza sativa* cv. IR64) × *O. glaberrima* under lowland moisture stress. Journal of Plant Biology **54**, 237–250.

Black R, Allen L, Bhutta Z. 2008. Climate change, food insecurity and hunger: key messages for UNFCCC negotiators. http://www.careclimatechange.org/files/reports/IASC_CC_FS.pdf

Blum A. 2009. Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. Field Crops Research **112**, 119–123.

Bocco R, Lorieux M, Seck PA, Futakuchi K, Manneh B, Baimey H, Ndjiondjop MN. 2012. Agro-morphological characterization of a population of introgression lines derived from crosses between IR 64 (*Oryza sativa indica*) and TOG 5681 (*Oryza glaberrima*) for drought tolerance. Plant Science **183**, 65–76.

Burgess AJ, Retkute R, Preston SP, Jensen OE, Pound MP, Pridmore TP, Murchie EH. 2016. The 4-dimensional plant: effects of wind-induced canopy movement on light fluctuations and photosynthesis. Frontiers in Plant Science **7**, 1392.

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.

Campany CE, Tjoelker MG, von Caemmerer S, Duursma RA. 2016. Coupled response of stomatal and mesophyll conductance to light enhances photosynthesis of shade leaves under sunflecks. Plant, Cell & Environment **39**, 2762–2773.

Challinor AJ, Watson J, Lobell DB, Howden SM, Smith DR, Chhetri N. 2014. A meta-analysis of crop yield under climate change and adaptation. Nature Climate Change **4**, 287–291.

Chatterjee J, Thakur V, Nepomuceno R, *et al.* 2020. Natural diversity in stomatal features of cultivated and wild *Oryza* species. Rice **13**, 58.

Choquette NE, Ogut F, Wertin TM, Montes CM, Sorgini CA, Morse AM, Brown PJ, Leakey ADB, McIntyre LM, Ainsworth EA. 2019. Uncovering hidden genetic variation in photosynthesis of field-grown maize under ozone pollution. Global Change Biology **25**, 4327–4338.

Cubry P, Pidon H, Ta KN, *et al.* 2020. Genome wide association study pinpoints key agronomic QTLs in African rice *Oryza glaberrima*. Rice **13**, 66.

Deans RM, Brodribb TJ, Busch FA, Farquhar GD. 2019. Plant water-use strategy mediates stomatal effects on the light induction of photosynthesis. New Phytologist **222**, 382–395.

De Souza AP, Wang Y, Orr DJ, Carmo-Silva E, Long SP. 2020. Photosynthesis across African cassava germplasm is limited by Rubisco and mesophyll conductance at steady state, but by stomatal conductance in fluctuating light. New Phytologist **225**, 2498–2512.

Dittberner H, Korte A, Mettler-Altmann T, Weber APM, Monroe G, de Meaux J. 2018. Natural variation in stomata size contributes to the local adaptation of water-use efficiency in *Arabidopsis thaliana*. Molecular Ecology **27**, 4052–4065.

Draic P, Flood J, Harbinson J, Aarts MGM. 2011. Natural genetic variation in plant photosynthesis. Trends in Plant Science **16**, 327–335.

Drake PL, Froend RH, Franks PJ. 2013. Smaller, faster stomata: scaling of stomatal size, rate of response, and stomatal conductance. Journal of Experimental Botany **64**, 495–505.

Faralli M, Matthews J, Lawson T. 2019. Exploiting natural variation and genetic manipulation of stomatal conductance for crop improvement. Current Opinion in Plant Biology **49**, 1–7.

Ghesquière A, Séquier J, Second G, Lorieux M. 1997. First steps towards a rational use of African rice, *Oryza glaberrima*, in rice breeding through a 'contig line' concept. Euphytica **96**, 31–39.

Głowacka K, Kromdijk J, Kucera K, Xie J, Cavanagh AP, Leonelli L, Leakey ADB, Ort DR, Niyogi KK, Long SP. 2018. Photosystem II Subunit S overexpression increases the efficiency of water use in a field-grown crop. Nature Communications **9**, 868.

Gu J, Yin X, Stomph T-J, Struik PC. 2014. Can exploiting natural genetic variation in leaf photosynthesis contribute to increasing rice productivity? A simulation analysis. Plant, Cell & Environment **37**, 22–34.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology **25**, 1965–1978.

Hubbart S, Smillie IRA, Heatley M, Swarup R, Foo CC, Zhao L, Murchie EH. 2018. Enhanced thylakoid photoprotection can increase yield and canopy radiation use efficiency in rice. Communications Biology **1**, 22.

Kromdijk J, Głowacka K, Leonelli L, Gabilly ST, Iwai M, Niyogi KK, Long SP. 2016. Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. Science **354**, 857–861.

Lawson T, Blatt MR. 2014. Stomatal size, speed, and responsiveness impact on photosynthesis and water use efficiency. Plant Physiology **164**, 1556–1570.

Lawson T, Vialet-Chabrand S. 2019. Speedy stomata, photosynthesis and plant water use efficiency. New Phytologist **221**, 93–98.

Lawson T, von Caemmerer S, Baroli I. 2010. Photosynthesis and stomatal behaviour. In: Lüttge U, Beyschlag W, Büdel B, Francis D, eds. Photosynthesis and stomatal behaviour. Progress in Botany 72. Berlin Heidelberg: Springer, 265–304.

Lê S, Josse J, Husson F. 2008. FactoMineR: an R package for multivariate analysis. Journal of Statistical Software **25**, 1–18.

Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL. 2014. Microsoft coco: common objects in context. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, eds. Computer Vision—ECCV 2014. Lecture Notes in Computer Science, vol 8693. Cham: Springer, 740–755.

Linares OF. 2002. African rice (*Oryza glaberrima*): history and future potential. Proceedings of the National Academy of Sciences, USA **99**, 16360–16365.

McAusland L, Murchie EH. 2020. Start me up; harnessing natural variation in photosynthetic induction to improve crop yields. New Phytologist **227**, 989–991.

McAusland L, Vialet-Chabrand S, Davey P, Baker NR, Brendel O, Lawson T. 2016. Effects of kinetics of light-induced stomatal responses on photosynthesis and water-use efficiency. New Phytologist **211**, 1209–1220.

Merk HL, Yarnes SC, Van Deynze A, *et al*. 2012. Trait diversity and potential for selection indices based on variation among regionally adapted processing tomato germplasm. Journal of the American Society for Horticultural Science **137**, 427–437.

Mohammed U, Caine RS, Atkinson JA, Harrison EL, Wells D, Chater CC, Gray JE, Swarup R, Murchie EH. 2019. Author Correction: rice plants overexpressing OsEPF1 show reduced stomatal density and increased root cortical aerenchyma formation. Scientific Reports **9**, 14827.

Murchie EH. 2021. Data from: Natural variation in photosynthetic traits in a diverse population of *Oryza glaberrima*. [Dataset]. Zenodo. Doi: 10.5281/zenodo.5555931

Murchie EH, Kefauver S, Araus JL, Muller O, Rascher U, Flood PJ, Lawson T. 2018. Measuring the dynamic photosynthome. Annals of Botany **122**, 207–220.

Murchie EH, Lawson T. 2013. Chlorophyll fluorescence analysis: a guide to good practice and understanding some new applications. Journal of Experimental Botany **64**, 3983–3998.

Murchie EH, Ruban AV. 2020. Dynamic non-photochemical quenching in plants: from molecular mechanism to productivity. The Plant Journal **101**, 885–896.

Ohsumi A, Kanemura T, Homma K, Horie T, Shiraiwa T. 2007. Genotypic variation of stomatal conductance in relation to stomatal density and length in rice (*Oryza sativa* L.). Plant Production Science **10**, 322–328.

Qu M, Zheng G, Hamdani S, Essemine J, Song Q, Wang H, Chu C, Sirault X, Zhu XG. 2017. Leaf photosynthetic parameters related to biomass accumulation in a global rice diversity survey. Plant Physiology **175**, 248–258.

Ray DK, West PC, Clark M, Gerber JS, Prishchepov AV, Chatterjee S. 2019. Climate change has likely already affected global food production. PLoS One **14**, e0217148.

Ren S, He K, Girshick R, Sun J. 2015. Faster r-cnn: towards real-time object detection with region proposal networks. Advances in Neural Information Processing Systems **39**, 1137–1149.

Ritz C, Baty F, Streibig JC, Gerhard D. 2015. Dose–response analysis using R. PLoS One **10**, e0146021.

Robinson GK. 1991. That BLUP is a good thing: the estimation of random effects. Statistical Science **6**, 15–32.

Sarla N, Swamy BPM. 2005. *Oryza glaberrima*: a source for the improvement of *Oryza sativa*. Current Science **89**, 955–963.

Slattery RA, Walker BJ, Weber APM, Ort DR. 2018. The impacts of fluctuating light on crop performance. Plant Physiology **176**, 990–1003.

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.

Sun J, Ye M, Peng S, Li Y. 2016. Nitrogen can improve the rapid response of photosynthesis to changing irradiance in rice (*Oryza sativa* L) plants. Scientific Reports **6**, 30315.

Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. 2016. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, NV, 2818–2826.

Taylor SH, Long SP. 2017. Slow induction of photosynthesis on shade to sun transitions in wheat may cost at least 21% of productivity. Philosophical Transactions of the Royal Society B: Biological Sciences **372**, 20160543.

Tinoco-Ojanguren C, Pearcy RW. 1993. Stomatal dynamics and its importance to carbon gain in two rainforest Piper species - I. VPD effects on the transient stomatal response to lightflecks. Oecologia **94**, 388–394.

Wambugu PW, Ndjiondjop MN, Henry R. 2019. Advances in molecular genetics and genomics of African rice (*Oryza glaberrima* Steud). Plants **8**, 376.

Wolfe MD, Tonsor SJ. 2014. Adaptation to spring heat and drought in northeastern Spanish *Arabidopsis thaliana*. New Phytologist **201**, 323–334.

York LM. 2018. Phenotyping crop root crowns: general guidance and specific protocols for maize, wheat, and soybean. Methods in Molecular Biology **1761**, 23–32.

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, USA **114**, 9326–9331.

Zhu H, Zhou X. 2020. Statistical methods for SNP heritability estimation and partition: a review. Computational and Structural Biotechnology Journal **18**, 1557–1568.

4.3 CHAPTER THREE SUPPLEMENTARY MATERIAL

Due to the large volume of supplementary material included in this chapter the material is not included here to reduce the thesis file size. It can be accessed at the Zenodo repository: https://doi.org/10.5281/zenodo.5555930: Murchie., 2021.

Supplementary material legends are listed below.

4.3.2 Supplementary Figures

Supplementary Figure S3.1: Original, un-fitted data for *A*, *gs* and NPQ vs time. Gas exchange data was logged every 10 s. (a) Mean values for *A* (mmol m^{-2} s^{-1}); (b) individual replicates for *A*; (c) Mean values for *gs* (mmol m^{-2} s^{-1}); (d) individual replicates for *gs*; **(e)** Mean values for NPQ; (f) individual replicates for NPQ.

Supplementary Figure S3.2: Correlations between the best linear unbiased predictor (BLUP) values and the original mean. The BLUP value is generated from the linear mixed effect model that was used to account for undesirable variation due to spatial and temporal effects on the phenotype. The BLUP is then used to calculate the adjusted mean value. Here we grouped the plots into trait types: (a) steady state gas exchange and chlorophyll fluorescence; (b) dynamic CO_2 assimilation; (c) dynamic stomatal conductance; (d) dynamic NPQ; (e) plant morphology.

Supplementary Figure S3.3: Correlation matrix of all phenotypic traits measured. *R* value is indicated by colour, shown in the right-hand scale bar. Stars show the significance between traits; p<0.001***, p<0.01**, p<0.05*.

Supplementary Figure S3.4: Modelled curves for two extreme *O. glaberrima* accessions and *O. sativa* IR64, plotted on a log scale, for example *A*, *gs* and NPQ induction and relaxation dynamic traits; **(**a) $A_{\text{i rate}}$, (b) $A_{\text{r slope}}$, (c) $gs_{\text{i rate}}$, (d) $gs_{\text{r slope}}$, (e) $NPQ_{\text{i slope}}$ and (f) $NPQ_{\text{r slope}}$. The lines on all plots from the y and x-axes to the curve show the time to reach 50% of the curve maximum.

Supplementary Figure S3.5: Linear regression plots showing strong positive correlations between the actual measurement vs modelled estimate values for (a) A_{max} vs $A_{\text{i max}}$, (b) gs_{max} vs $gs_{\text{i max}}$ and (c) NPQ_{max} vs $NPQ_{\text{i max}}$.

Supplementary Figure S3.6: Plots showing the (a) scree plot and (b) trait loadings for the phenotypic data PCA analysis.

Supplementary Figure S3.7: (a) H-clustering dendrogram of 105 *O. glaberrima* accessions analysed for similarities based on climatic traits, showing 3 distinct clusters. Frequency plots generated from the H-clustering analysis of the phenotypic data from 155 *O. glaberrima* accessions, showing the frequency of accessions in each cluster for (b) ecological niche and (c) country of origin.

Supplementary Figure S3.8: Correlation matrix for all phenotypic and climatic data, alongside with their principal components. Correlation strength is indicated by colour, shown in the right-hand scale bar. Stars show the significance between traits; $p<0.001$***, $p<0.01$**, $p<0.05$*.

4.3.2 Supplementary Tables

Supplementary Table S3.1: List of parameter abbreviations, definitions and units of measurement.

Supplementary Table S3.2**:** List of *O. glaberrima* ID codes, country of origin and ecology.

Supplementary Table S3.3a: Estimated LL.4 model outputs on carbon assimilation (A) IRGA induction data, showing the 4 replicates for accession IRGC_96726.

Supplementary Table S3.3b: Estimated LL.4 model outputs on stomatal conductance (gs) IRGA induction data, showing the 4 replicates for accession MG04.

Supplementary Table S3.3c: Estimated LL.3 model outputs on non-photochemical quenching (NPQ) IRGA induction data, showing the 4 replicates for accession TOG_12188.

Supplementary Table S3.3d: Estimated W2.4 model outputs on non-photochemical quenching (NPQ) IRGA relaxation data, showing the 3 replicates for accession EG55. EG55 was one of a small number of accession where only 3 replicates were measured.

Supplementary Table S3.3e: Estimated LL.4 model outputs on stomatal conductance (gs) IRGA relaxation data, showing the 4 replicates for accession TOG_5326.

Supplementary Table S3.3f: Estimated LL.4 model outputs on carbon assimilation (A) IRGA relaxation data, showing the 4 replicates for accession UG26.

4.3.2 Supplementary Data Files

Supplementary Data File S3.1: Methodology of stomata detection and counting through machine learning.

Supplementary Data File S3.2: Descriptive statistics of all data from *Oryza glaberrima* accessions.

Supplementary Data File S3.3: Descriptive statistics of all data from *Oryza sativa* accession IR64.

Supplementary Data File S3.4: Box and distribution plots of linear mixed effects model adjusted means for *Oryza glaberrima* and *Oryza sativa*

Supplementary Data File S3.5: The Python code for extracting a defined percentage of a curve.

Supplementary Data File S3.6: This shows the Principal Component Analysis of key traits.

Supplementary Data File S3.7: This shows the Hierarchical Cluster analysis of key traits.

Supplementary Data File S3.8: This shows the Principal Component Analysis, climate data.

Supplementary Data File S3.9: This shows the Hierarchical Cluster analysis , climate data .

Supplementary Data File S3.10: Correlation table of phenotypic traits form *Oryza sativa* and *Oryza glaberrima*

Supplementary Data File S3.11: List of accessions in each hierarchical cluster clade.

Supplementary Data File S3.12: Correlation table of Climate and Phenotypic data, 105 accessions.

CHAPTER FOUR

Genome Wide Association Study (GWAS) on stomatal, steady-state and dynamic photosynthetic traits in a diverse population of African rice (Oryza glaberrima).

4.1 Chapter four introductory statement

This chapter consists of a draft research manuscript, using the phenotypic data obtained in Chapter Three to identify trait genomic intervals and candidate genes for those traits. We intend to progress and submit this manuscript for publication during 2022. The draft manuscript shows the completed GWAS outputs and candidate gene lists for morphological, photosynthetic dynamic and steady-state traits. In addition to the work shown here, we would like to complete gene *O. sativa* expert enrichment lists and compare these to the *O. glaberrima* GWAS outputs, in line with the work by Cubry et al., 2020.

At the start of this study the sequencing of the *O. glaberrima* population and downstream bioinformatic work was still in progress. At the beginning of the PhD (2017) we had access to 150 re-sequenced, but not yet cleaned accessions. There are a now a total of 163 sequenced *O. glaberrima* accessions, with no intentions for further sequencing. The GWAS results described here are the product of three cycles of association analysis and continuous improvement over the course of this study, with these stages described below:

1.) Circa. 2018: no cleaned SNP file had been produced by the Research Institute for Development (IRD) at the University of Montpellier, but they allowed me access to the raw SNP files through the Gigwa database (Sempéré et al., 2016). A bioinformatician, Dr. Niraj Shah, in the Advanced Data Analysis Centre at Uni. Of Nottingham was commissioned to clean the raw SNP files. However, when this was used to complete an initial association analysis using FarmCPU and GAPIT in R-studio, it was clear that the data was being affected by confounding factors. On reflection, the file was cleaned with *O. bathii* accessions also included and the missing allele frequency was not calculated based on the specific subset of accessions under use. This would have affected the ability to accurately calculate and adjust for the effects of population structure and kinship.

2.) Shortly after identifying these issues, the Rice Group shared a cleaned and imputed SNP file, alongside an *O. glaberrima* specific GWAS pipeline. After successfully

running this on my traits up until this point, it seemed that there were environmental interactions, likely from measuring the accessions being grown in non-randomised batches over time, which were effecting the GWAS results.

3.) A linear mixed effects model was used to account for temporal and spatial effects in the raw phenotypic data and from this Best Linear Unbiased Predictors (BLUPs) and adjusted means were generated that were then used in the reported analyses (Chapters 3 and 4). The final GWAS output shown here uses adjusted means, the cleaned and imputed SNP file and GWAS pipeline from the RICE group for a robust analysis.

Genome Wide Association Study (GWAS) on stomatal, steady-state and dynamic photosynthetic traits in a diverse population of *Oryza glaberrima*.

Sophie B. Cowling[1], Philippe Cubry[2], Pracha Treeintong[1], John Ferguson[1], Ranjan Swarup[1], Erik H. Murchie[1], Sean Mayes[1*]

[1] Division of Plant and Crop Science, School of Biosciences, University of Nottingham, Sutton Bonington Campus, LE12 5RD, UK.
[2] DIADE, University of Montpellier, IRD, Montpellier, France.
* Corresponding author: Sean.Mayes@nottingham.ac.uk

Paper author contributions:

SBC collected phenotype data. Performed phenotypic data cleaning, phenotypic and GWAS analysis, performed GWAS. Wrote and developed the manuscript with advice from EHM and SM.
PC generated the African Rice GWAS pipeline and ran the GWAS results files through the SNP2Gene analysis, producing candidate gene lists from KnetMiner.
JF produced R scripts for the linear mixed effects model, BLUPS and heritability analysis.
PT completed glasshouse experimental design. PT and SBC collected phenotypic data.
RS, EHM and SM supervised project and manuscript development.

Keywords: (5-8)
GWAS, *O. glaberrima*, photosynthesis, natural variation, rice, stomata.

ABSTRACT

While climate change will mean a generally unstable climate, the main consequence is global warming. For this reason, breeding crops for future climates means improving water use efficiency and heat stress tolerance, while balancing the trade-off for productivity. African rice (*Oryza glaberrima*) exhibits many abiotic stress tolerance traits and may offer useful genetic variation beyond that found in the highly selected *O. sativa*. We measured 52 phenotypic traits relating to climate resilience, including photosynthesis steady-state and dynamics, water use efficiency and stomatal density. Here we describe the results of a genome wide association study (GWAS) across these diverse traits in a population of 150 re-sequenced *O. glaberrima* accessions. We identified candidate genes including those known for mitigating drought and heat stress, alongside regulators of key phytohormones. Of note, we identified several candidate genes for cytochrome subcomplexes across a range of photosynthesis related traits. This study provides a resource for the future breeding selection through application of linked markers for trait loci and genetic validation for traits of interest. It also highlights the importance *O. glaberrima* as a novel genetic resource in rice breeding for climate resilience.

4.2 INTRODUCTION

The African rice species, *Oryza glaberrima*, underwent independent evolution and domestication to its Asian cousin, *O. sativa*. While the two species share orthologous genes that are associated with domestication (Wang et al., 2014), they show substantial genetic differentiation as they have not shared a common ancestor for approximately 0.6 – 0.9 million years (Purugganan, 2014; Huang et al., 2015). *O. glaberrima* was domesticated in West Africa (Cubry et al., 2018) ~3 thousand years ago, from the wild Sahelian species *O. bathii*, and some 7 thousand years after *O. sativa*. Consequently, *O. glaberrima* retained many traits that facilitate abiotic and biotic stress resilience to the African environment, including drought, pest, salinity and soil toxicity resistance (Sarla & Swamy, 2005). As a semi-aquatic plant, *O. sativa* is highly sensitive to drought events and the changing climate. Conservative predictions indicate up to 37% *O. sativa* yield losses due to temperature increases by the end of the century, but the reality may be far more serious (Bocco *et al.*, 2012; Zhao *et al.*, 2017). The abiotic stress tolerance and ability to thrive under low-input conditions make *O. glaberrima* an ideal candidate for gene identification and translation to *O. sativa* and other crop species.

O. glaberrima has been recognised as a potential source of genetic diversity for over two decades (Ghesquière et al., 1997; Sarla & Swamy, 2005). But considering the extent of potentially desirable variation observed in *O. glaberrima*, there is limited literature describing the phenotypic and underlying genetic variation that can be taken forward for translation. Allele mining has successfully identified Rust Yellow Mottle Virus (RYMV) resistance genes (Thiémélé et al., 2010; Pidon et al., 2020), association studies have been undertaken on salt tolerance (Meyer et al., 2016) and key agronomic traits (flowering time, panicle branching and RYMV resistance; Cubry et al., 2020).
The recent re-sequencing of diverse *O. glaberrima* accessions provides a new resource to explore the genetic basis of functional traits, in what remains a relatively untapped resource (Cubry et al., 2018, 2020). Genome wide association studies (GWAS) have proven a powerful tool when exploring complex traits in *O. sativa*, such as elucidating the genetic basis of stomatal traits (Chen et al., 2020). GWAS may also prove an effective method for identifying the genetic basis of climate resilient traits in *O. glaberrima*.

Water use efficiency and the tolerance of high light levels are important components for the development of abiotic stress resilient crops, but this also needs to be balanced with

photosynthetic efficiency for productivity (Lawson & Blatt, 2014; Kissoudis et al., 2016; Hubbart et al., 2018). Stomata are key players when developing resilient and productive crop varieties for future climates (Xu et al., 2016; Faralli et al., 2019; Buckley et al., 2020). These small pores on the leaf surface regulate CO_2 assimilation (*A*) and water lost via transpiration, both are essential in water use efficiency (WUE) and productivity in water limited environments (Lawson et al., 2010). Physical stomatal traits, such as density and morphology, are known to influence stomatal conductance (*gs*) (Drake et al., 2013). For example, reducing stomatal density in rice has been shown to improve water conservation under drought and high temperature conditions while maintaining yields (Caine et al., 2019; Mohammed et al., 2019). Recent developments in stomatal research have shown that along with stomatal density, altering stomatal dynamics can conserve water and enhance productivity through the close coupling of *gs* to dynamic photosynthetic responses to light (Lawson & Blatt, 2014; Lawson & Vialet-Chabrand, 2019; Acevedo-Siaca et al., 2020). It has been long established that stomatal traits are important in carbon assimilation, but photoprotection mechanisms are also important for preventing yield losses under high light conditions (Taylor and Long, 2017; Murchie et al., 2018; Acevedo-Siaca et al., 2020). Non-photochemical quenching (NPQ) is a photoprotective mechanism that dissipates excess energy as heat. The rapid relaxation of NPQ has been shown to improve photosynthetic efficiency and yield gains under fluctuating light and high light conditions (Kromdijk et al., 2016; Hubbart et al., 2018). For the first time, these complex stomatal and dynamic photosynthesis related traits have recently been characterised in *O. glaberrima* (Cowling et al., 2021). Elucidating the genetic basis of this suite of traits in African rice can contribute towards the development of crops that are resilient to the projected change in climate.

As a notably resilient species, we hypothesised that *O. glaberrima* will prove an interesting source of candidate genes and quantitative trait loci (QTLs) in useful photosynthetic and stomatal traits. We used a diverse panel of 150 whole-genome re-sequenced *O. glaberrima* accessions (Cubry et al., 2018, 2020) to characterise 52 phenotypic traits. Using four complementary GWAS methods we were able to successfully identify genes and regions for many traits, including stomatal density and numerous dynamic traits for stomatal conductance, photosynthesis and NPQ. We hope this study highlights the importance of *O. glaberrima* as a potential source of novel genetic variation and provides useful material for developing future rice resilience in a changing climate.

4.3 MATERIALS AND METHODS

4.3.1 Plant material and measurements conditions

The description for plant material and growth conditions can be found in chapter three materials and methods, section 3.3.1.

4.3.2 Acquisition of phenotypic data

The acquisition of phenotypic data, including infrared gas analyser (IRGA), stomatal density and morphological measurements can be found in chapter three, materials and methods, sections 3.3.2 to 3.3.4.

4.3.3 Phenotypic data analysis

All data analysis was completed in R-Studio (v. 4.0.1). We used a linear mixed-effects model (lme4 package, v.1.1-26) to account for the variance due to temporal and spatial effects of measuring the accessions in batches. The linear mixed-effects model calculated best linear unbiased predictions (BLUPs), and from this generated adjusted mean for each trait. The model took into account the effects of: accession, sowing date, measurement date, nested location within the glasshouse and, for gas-exchange traits, LiCor machine. All results reported here use the adjusted means data generated from the mixed effects model. Broad sense heritability (H^2) was calculated using this mixed effect model as the variance attributed to *O. glaberrima* accession divided by the sum of the variation due to genotype over the total variance.

Trait distribution was calculated using the Shapiro-Wilk test, using a more stringent $\alpha = 0.01$, as the Shapiro-Wilk test is biased by outliers and tends to report false negatives in sample sizes exceeding 50. Traits that showed a non-normal distribution were transformed using the Box-Cox transformation method in the package Forecast (v. 8.14). Ten traits were resistant to transformation, due to heavy skew, and were checked for outliers using the Tukey's method. The association models use parametric methods, which assume the input of datasets with a Gaussian distribution. To meet the test assumptions, data points that exceeded the 1.5 * interquartile range were removed and trait distribution was checked again using a Shapiro Wilk test. Three remaining traits were then transformed using the Box-Cox method. Traits with negatives values, such as $gs_{\text{i slope}}$, that cannot be transformed using the Box-Cox method, were log transformed. A complete list of traits, H^2 values, descriptive statistics, transformation method and outlier removal can be found in Table 4.1.

The estimation of dynamic photosynthetic traits for *A*, NPQ and *gs* was completed using the 'drc' (v. 3.0) package. This package analyses dose-response curves, in this case light response, through a suite of model fitting methods (Ritz et al., 2015). The 'drc' (v. 3.0) package was used extract parameters for the induction and relaxation responses, including the curve slope ($_{i/r\ slope}$), lower limit ($_{i/r\ min}$) and the time taken to reach a defined percentage of the dependant variable, in this case 10 ($_{i/r\ 10}$), 50 ($_{i/r\ 50}$) and 90 ($_{i/r\ 90}$) %. *A* and *gs* induction and relaxation responses were modelled using a LL.4 (log-logistic 4-parameter). While a LL.3 (log-logistic 3-parameter) model was used for NPQ induction and a W2.4 (4-parameter Weibull2) model for NPQ relaxation. The comparison of eight different models, followed by Akaike's information criterion analysis was used to select the best model fit to the response curves.

Principal component analysis (PCA) and population structure analysis was completed using the function 'snmf()' (sparse non-negative matrix factorisation(sNMF)) in the R-package LEA (v. 3.0.0). This produces results akin to the programme STRUCTURE. Ancestral groups (K) were estimated to be between 1:10 groups and run for 5 repetitions. The estimated K was used to assist some association analysis methods to account for the confounding effects of population structure. Dendrogram analysis was completed by computing genetic distance using Nei's distance (Takezaki & Nei, 1996), and establishing population clusters (k = 9 as indicated by Bayes Information Criterion (BIC)) using hierarchical clustering. This analysis used the R packages adegenet (v. 2.1.3) and NAM (v. 1.7.3), for the multivariate analysis of genetic markers and nested association mapping respectively. The dendextend (v. 1.14) and circlize (v.0.4.1) packages were used for generating dendrogram plots. Pearson correlation coefficients were calculated using the Corrplot package (v. 0.84).

Table 4.1: List of traits, variation and data distribution in the *O. glaberrima* association panel**.** The skew and Shapiro-Wilk p-value are given in columns two and three respectively. Traits that were found to deviate significantly ($p<0.01$) from the normal distribution, per the Shapiro-Wilk normality test, were transformed. Whether a trait was transformed, and the method used, is detailed in the 'Trans.method' column, where a blank cell means no transformation occurs, 'BC' refers to the Box-Cox method and 'log' is $\log_{10}$. In traits resistant to transformation, outliers were identified and removed using the Tukey method. 'Outliers removed' details how many outliers were removed. Trait distributed was then re-tested using the Shapiro Wilk normality test after transformation or outlier removal, the result can be found in 'Post trans p-value' column.

Trait	Skew	p-value	Outliers removed	Trans. method	Post trans p-value
$A_{i\ 10}$	2.08	<0.00	9	BC	0.09
$A_{i\ 50}$	1.55	<0.00	8	BC	0.08
$A_{i\ 90}$	0.59	0.001		BC	0.03
$A_{i\ min}$	0.42	0.054			
$A_{i\ rate}$	-0.01	0.678			
$A_{i\ slope}$	-1.27	<0.00	9		0.22
$A_{r\ 10}$	-0.48	0.059			
$A_{r\ 50}$	1.67	<0.00	8	BC	0.30
$A_{r\ 90}$	0.82	<0.00	7	BC	0.38
$A_{r\ min}$	0.28	0.109			
$A_{r\ rate}$	-2.41	<0.00	3	BC	0.19
$A_{r\ slope}$	0.64	0.001	7		0.37
Abaxial SD	-0.25	0.050			
Adaxial SD	-0.33	0.168			
A_{max}	0.13	0.190			
ETR_{max}	0.06	0.740			
$gs_{i\ 10}$	1.00	<0.00		BC	0.98
$gs_{i\ 50}$	0.86	0.001		BC	0.89
$gs_{i\ 90}$	1.03	<0.00	6	BC	0.02
$gs_{i\ min}$	1.37	<0.00		log	0.01
$gs_{i\ rate}$	0.65	0.002		BC	0.35
$gs_{i\ slope}$	-0.71	0.003	3		0.21
$gs_{r\ 10}$	1.20	<0.00	6	BC	0.01
$gs_{r\ 50}$	0.21	0.035			
$gs_{r\ 90}$	0.41	0.005		BC	0.05
$gs_{r\ min}$	-0.43	0.013			
$gs_{r\ rate}$	3.91	<0.00	10	BC	0.13
$gs_{r\ slope}$	0.81	0.001		BC	0.37
gs_{max}	0.33	0.276			
$iWUE_{max}$	0.73	<0.00		BC	0.01
NPQ_{max}	0.17	0.706			
$NPQ_{i\ 10}$	0.08	0.989			

$NPQ_{i\ 50}$	0.08	0.744		
$NPQ_{i\ 90}$	0.13	0.577		
$NPQ_{i\ rate}$	0.22	0.443		
$NPQ_{i\ slope}$	-0.62	0.001	log	0.07
$NPQ_{r\ 10}$	-0.05	0.868		
$NPQ_{r\ 50}$	0.33	0.207		
$NPQ_{r\ 90}$	0.27	0.414		
$NPQ_{r\ min}$	0.04	0.897		
$NPQ_{r\ rate}$	0.37	0.047		
$NPQ_{r\ slope}$	-0.33	0.158		
$\phi PSII_{max}$	0.06	0.740		
Plant height	0.15	0.299		
qP_{max}	0.01	0.822		
Ratio SD	0.17	0.114		
Root biomass	1.78	<0.00	BC	0.01
Shoot area	0.47	0.004	BC	0.02
Shoot biomass	0.58	0.007	BC	0.20
Total biomass	0.82	<0.00	BC	0.11
$Trmmol_{max}$	0.07	0.596		
VPD_{max}	-0.07	0.264		

4.3.4 Genome re-sequencing

The genome re-sequencing, relevant bioinformatics and genomic description was completed by the RICE team at DIADE, IRD, University of Montpellier, a detailed description of these methods can be found in Cubry et al. (2018). The genotypic data in this study uses single nucleotide polymorphism (SNP) markers from 150 of 163 *O. glaberimma* re-sequenced accessions. The prepared libraries were sequenced using Illumina (USA) HiSeq2000 or HiSeq4000 machines at 2x101 or 2x151 bp reads. The *Oryza sativa japonica* cv. Nipponbare reference genome (MSU Rice Genome Annotation project/IRGSP 1.0, release 7.0; Kayahara et al., 2013)) was used to map and call *O. glaberrima* SNPs. SNPs with >5% missing data were filtered out and the remainder of the missing SNP data was imputed using the 'impute()' function in the LEA package in R-studio. After cleaning and imputation, 892539 SNP

markers were used to complete the GWAS. Linkage disequilibrium was computed as 150kb by Cubry et al. (2020), using the PopLDdecay software.

4.3.5 Association Analysis

The GWAS was completed in R-Studio (v. 4.0.1), using a pipeline consisting of four association analysis methods; 1.) Latent Factor Mixed Model (LFMM; Frichot et al, 2013) , 2.) Genome Association and Prediction Integrated Tool (GAPIT; Lipka et al., 2012), 3.) Fixed and random model Circulating Probability Unification (FarmCPU; Liu et al., 2016) and 4.) Efficient Mixed Model Association (EMMA; (Kang et al., 2008). These four methods comprise of two different classes of association models, LFMM and Mixed Linear Model (MLM). The MLM method is used in the GAPIT, FarmCU and EMMA methods but each differs in how they account for confounding factors. The pipeline described here was developed by and described in Cubry et al. (2018), the script for this analysis can be found on the 'Africrop/gwas_african_rice' GitHub repository. At the beginning of each trait analysis, SNPs with a MAF <5% are filtered out. The results of each GWAS analysis and model fit were graphically represented by Manhattan and QQ-plots respectively, using the R-package qqman (v.0.1.4). Traits and regions were considered reliable and retained for further exploration when statistically significant regions were identified in two or more GWAS methods. Candidate genes were identified and considered within a 50kb window (+/- 25kb) of the statistically significant SNP, using a threshold of $p<10^{-5}$. Both the threshold value and size of the window surrounding each associated SNP were used in the methodology successfully used by Cubry et al. (2020) to identify flowering time genes in this *O. glaberrima* panel.

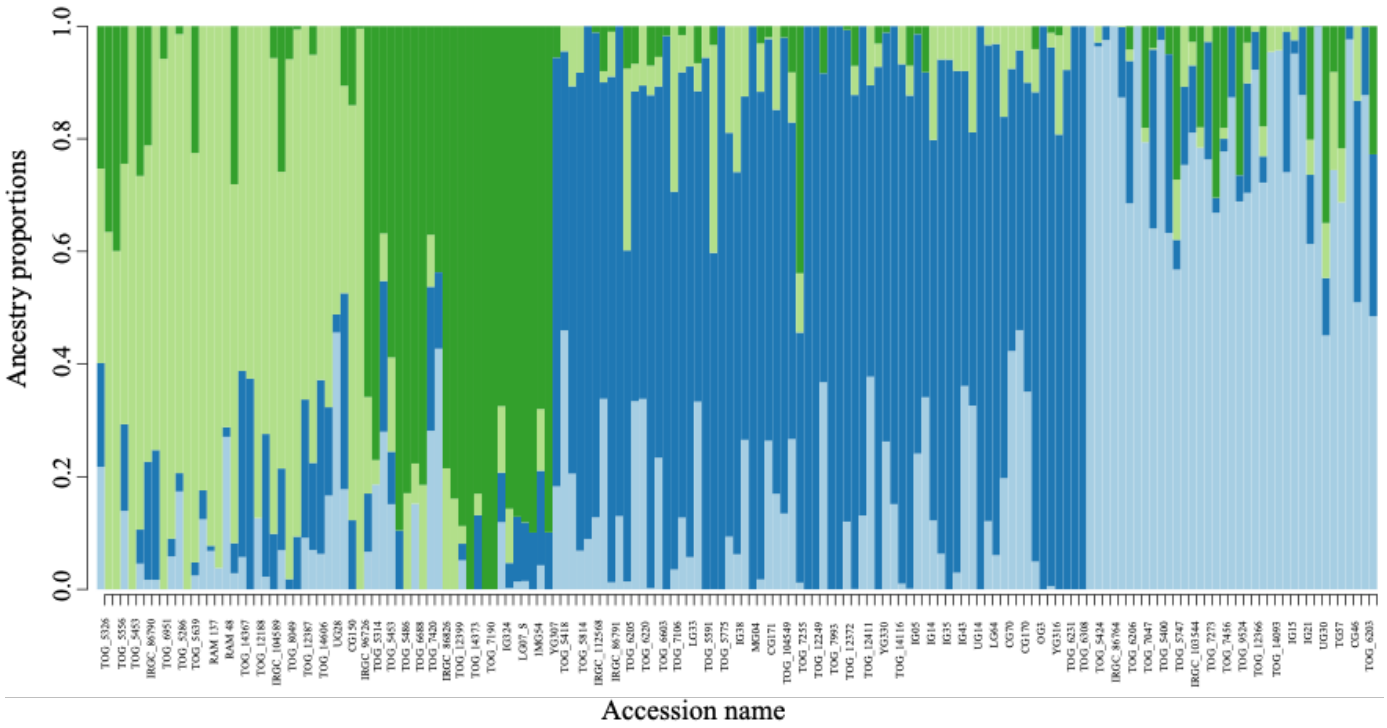
4.4 RESULTS

4.4.1 Associations between population structure, agro-ecology and trait diversity

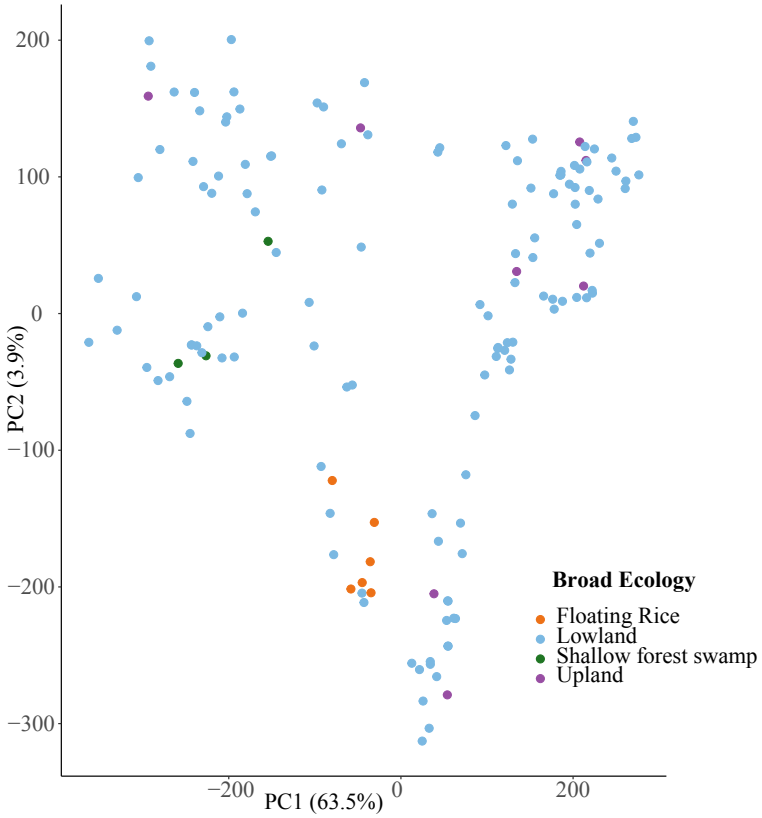
The genetic structure within *O. glaberrima* was found to be rather diffuse, with a substantial amount of admixture between groups. All analyses used to explore population structure in the *O. glaberrima* association panel indicated four main sub-populations (Fig. 4.1a-c). The sNMF method estimated four ancestral groups (K), selected from the cross-entropy criterion values (Supp. Fig. S4.1a). The cross-entropy criterion values showed an initial decrease from K = 1- 4, a slight knee can be seen at K=4 and the values plotted against K = 5:10 do

continue to decrease but only very slightly. The slow decay shown in the cross-entropy criterion values is also reflected in the kinship plot (Supp. Fig S4.1b). Therefore, K = 4 was selected as the best estimation for the number of ancestry groups. A principal components analysis (PCA) also shows four dispersed clusters of accessions, where the first two PCs account for 69.4% of the total genetic variance. H-clustering based dendrogram and kinship matrix (Supp. Fig S4.1b) also show four major clades in the *O. glaberrima* population.

(a)



(b)


200
100
0
−100
−200
−300
PC2 (3.9%)
−200
0
200
PC1 (63.5%)
Broad Ecology
Floating Rice
Lowland
Shallow forest swamp
Upland

(c)

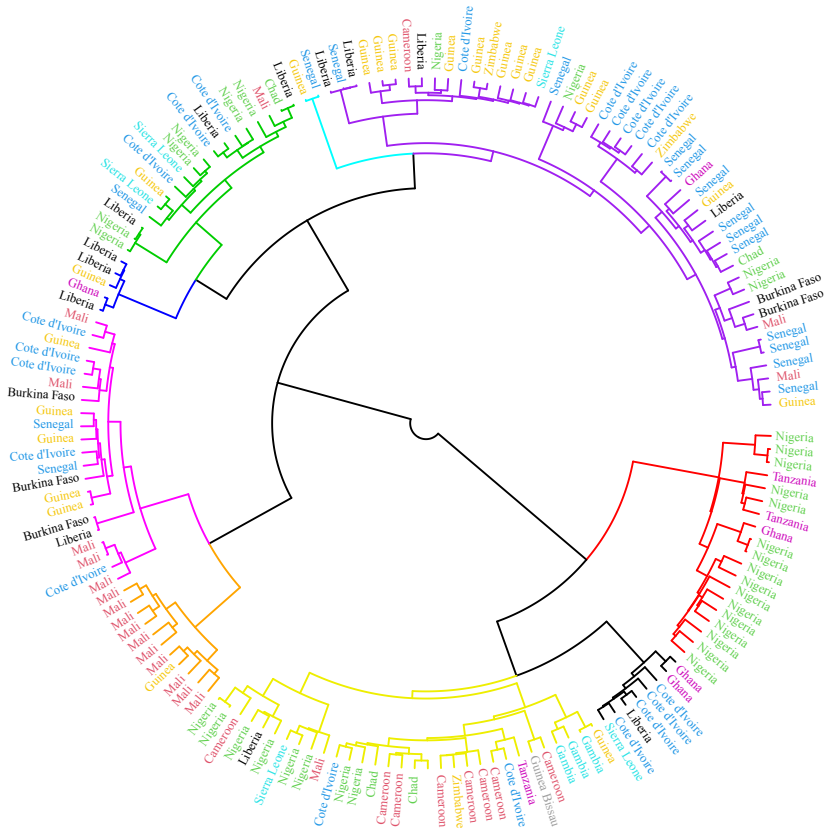


Figure 4.1: Evaluating *O. glaberrima* population structure and the effect of agroecological niche and country of origin on genetic similarity, using 892539 SNP markers. (a) Ancestry matrix bar plot, computed using the sNMF method, showing the estimated the population structure when K = 4. Each bar corresponds to an accession and the coloured proportion indicates the estimated ancestry coefficient (q). (b) Principal Component Analysis (PCA), showing a scatter plot of PCs 1 and 2. Individual accessions are coloured by the agroecological niche for their collection location. (c) Dendrogram computed from hierarchical clustering analysis using Nei's distance of SNP data, each terminal branch indicates an *O. glaberrima* accessions, which is labelled and coloured by country of origin.

The independent adaptation of *O. glaberrima* to variable African environments, not encountered by *O. sativa* in Asia, is an important source of the natural variation for potential breeding traits into *O. sativa*. We used the population analyses shown here to explore the potential association between the agro-ecological environment and genetic structure within the *O. glaberrima* sub-populations. Individuals shown in the PCA and dendrogram analyses were overlaid with their agro-ecological niche and country of origin respectively (Fig. 4.1, Supp. Figs S4.1, S4.2 and S4.3), based on the geo-coordinates from where the accessions were collected. The PCA shows (Fig. 4.1b) that accessions from shallow forest swamp and floating rice appear to cluster together. This can also be seen in Supp. Fig. S4.3c dendrogram, where all the floating rice and shallow forest swamp individuals cluster together in their respective clades. There is no obvious genetic clustering of accessions from other ecologies, either in the dendrogram or PCA out puts (Fig. 4.1b-c and Supp. Fig. S4.2a-d, S4.3c). All floating rice accessions, shown in the orange sub-clade, originate from Mali, while most accessions from forest swamp conditions originate from Nigeria (Fig. 4.1c). This suggests a distinct genetic similarity of accessions from these two countries, while the remaining accessions show no apparent clustering based upon country of origin.

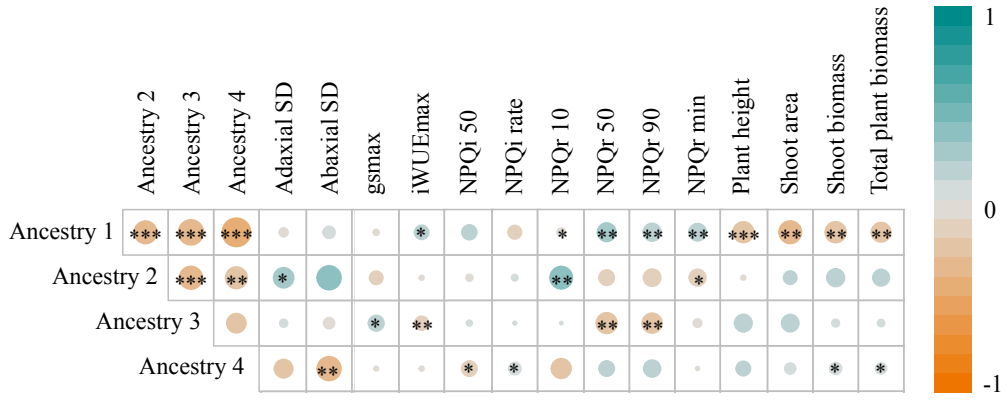


Figure 4.2: *O. glaberrima* ancestry groups show clear associations with specific phenotypic traits. The Pearson correlation matrix shows associations between the four ancestry groups, generated from the sNMF analysis, and phenotypic traits. Correlation values are shown by the coloured scale bar and asterisks indicate the significance between associations (∗∗∗p<0.001, ∗∗p<0.01, ∗p<0.05). Traits with significance results only are shown here, see Supp. Table S4.1 for the correlation matrix containing all traits in this study.

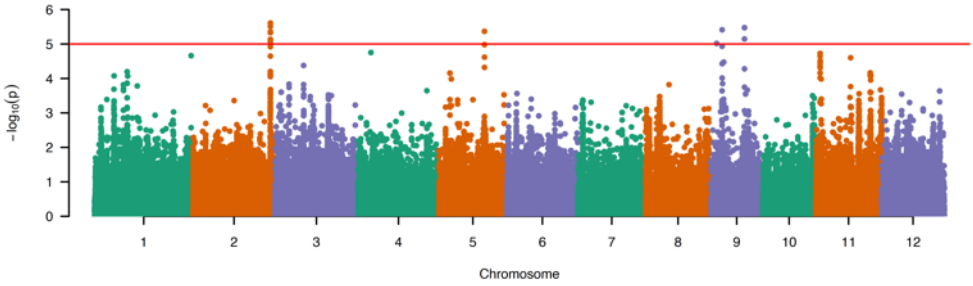
The *O. glaberrima* ancestry groups identified in the sNMF population structure analysis showed clear associations with groups of phenotypic traits (Fig. 4.2; Supp. Fig. S4.4 and Supp. Table S4.1). Ancestry group 1 showed significant associations with NPQ relaxation dynamics and biomass (p<0.05 – 0.001). While ancestry group 2 showed associations with enhanced adaxial stomatal density (p<0.05), a slower initial NPQ relaxation time ($NPQ_{r\ 10}$; p<0.01) and lower levels of NPQ achieved under low light ($NPQ_{r\ min}$; p<0.05). Ancestry group 3 showed correlations to enhanced stomatal conductance (p<0.05), lower iWUE and longer time taken for NPQ relaxation (p<0.01). Finally, individuals belonging most strongly to ancestry group 4 showed a higher abaxial stomatal density (p<0.01) and indications of an association with faster NPQ induction rate and shoot biomass (p<0.05).

4.4.2 Genome Wide Association Study

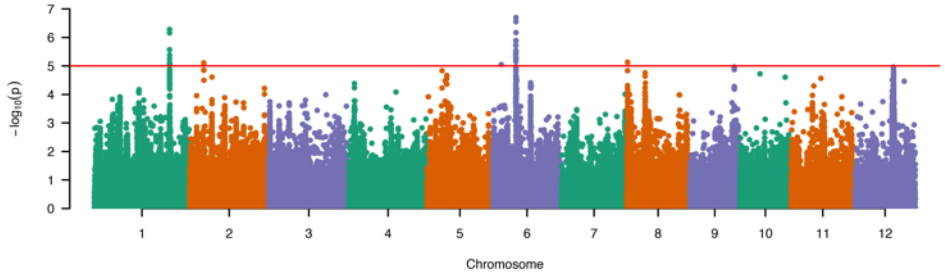
We used four GWAS methods (LFMM, GAPIT, FarmCPU and EMMA) to identify statistical associations between 52 phenotypic traits and 892, 539 single nucleotide polymorphism (SNP) genetic markers across 150 *O. glaberrima* individuals. For 44 out of 52 traits a total of 5992 candidate genes were identified in a 50kb window around each SNP (threshold = $p<10^{-5}$), across the four methods used here. With the results consolidated to remove duplicates between methods, a total gene list of 754 was created. No significant associations were identified for A_{max}, root biomass and six photosynthetic dynamic traits; $A_{r\ min}$, $gs_{i\ 50}$, $gs_{i\ rate}$, $gs_{r\ 50}$, $gs_{r\ min}$ and $NPQ_{r\ min}$. Significant associations were identified in 13 traits but only through a single GWAS method, these will not be reported here but are still valid and are presented in supplementary data. The remaining 30 traits showed significant genetic associations in two or more GWAS methods and will be reported here. Supplementary Table S4.2 details a list of traits and which GWAS methods produced significant associations. The EMMA method produced the largest number of significant associations, for 34 out of 52 traits, and of the 13 traits where only one method detected SNP/trait associations, EMMA was responsible for 71% of these. The other methods gave; LFMM (29/52 traits), FarmCPU (26/52 traits) and

lastly GAPIT showed the most conservative results (9/52 traits). The GWAS results are separated and discussed in respect to their physiological groups of morphology, steady state and then dynamic photosynthesis related traits. A full list of SNP/trait associations and Manhattan plots, based on a threshold of $p<10^{-5}$, can be found in Supplementary Table S4.3 and Supplementary Figures S4.5.

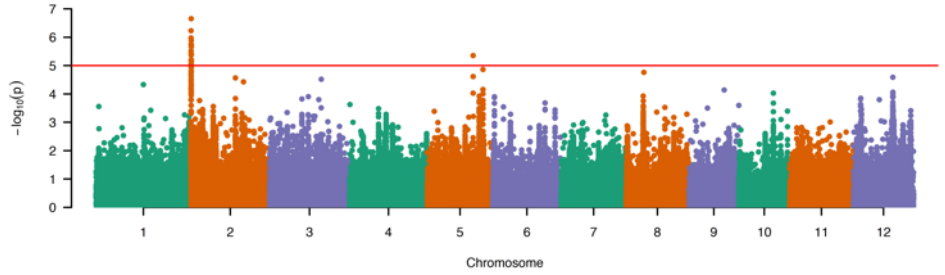
(a) Abaxial stomatal density



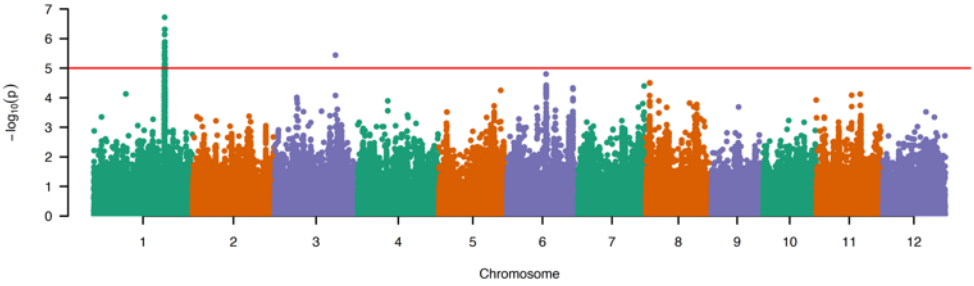
(b) $iWUE_{max}$



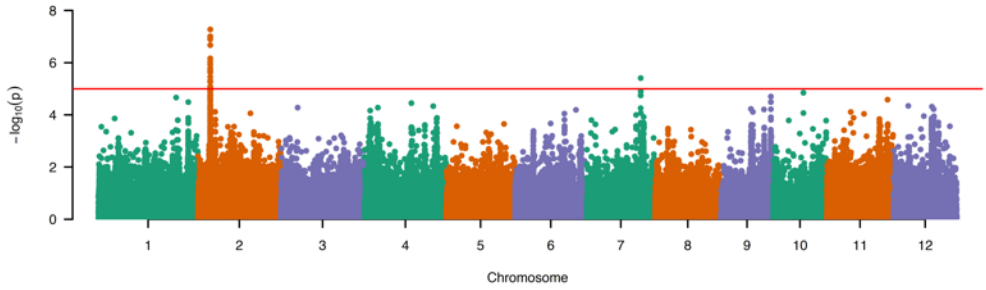
(c) NPQ_{max}



(d) $A_{i\ slope}$



(e) $gs_{i\ 90}$



(f) $NPQ_{r\ 10}$

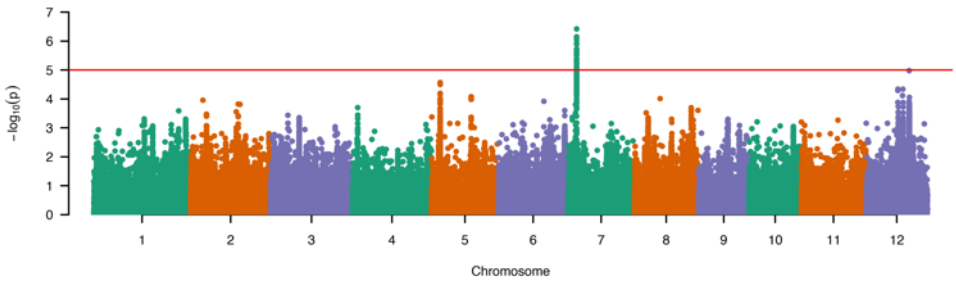


Figure 4.3: Manhattan plots of GWAS results with notably SNP/trait associations (threshold $p<10^{-5}$) across morphology, steady-state and then dynamic photosynthesis related traits, which include: (a) Abaxial SD, identified in two different GWAS methods (LFMM and EMMA). (b) $iWUE_{max}$, which showed significant SNPs across three methods (LFMM, FarmCPU and EMMA). (c) NPQ_{max} showed significant associations across all methods (LFMM, GAPIT, FarmCPU and EMMA). (d) $A_{i\ slope}$ showed significant SNP markers were identified across all GWAS methods. (e) $gs_{i\ 90}$ SNP/trait associations were identified across all GWAS methods. (f) Finally, significant associations were identified in three methods (LFMM, FarmCPU and EMMA) for $NPQ_{r\ 10}$. To view all Manhattan plots, please see Fig. S4.5.

4.4.3 Morphological traits

The GWAS of morphological traits in the *O. glaberrima* population identified 4 traits, abaxial stomatal density (SD), SD ratio, shoot area and plant height, with significant associations to SNPs. Abaxial SD was associated with 7 SNPs and a total of 31 candidate genes (Fig. 4.3a). SD ratio showed 3 SNP/trait associations and a total of 16 candidate genes, including a potential candidate for WRKY18, WRKY refers to a conserved amino acid domain and the large WRKY transcription factor gene family. Abaxial and SD ratio do not share significant SNPs. 7 SNPs and 22 candidate genes, including the *O. sativa* f-box/kelch repeat gene (*Os*FBK21), were associated with shoot area across all GWAS methods. 7 SNPs and 20 genes were identified for plant height across three methods, including the *Arabidopsis*

thaliana gene CSLA5 (cellulose synthase-like a5), a variant in the Cellulose Synthase-Like gene family.

4.4.4 GWAS of steady-state photosynthetic traits

SNP/trait associations were detected for the key photosynthetic traits; gs_{max}, $iWUE_{max}$, NPQ_{max}, qP_{max}, $trmmol_{max}$ and VPD_{max}. These traits showed some of the most significant SNP/trait associations in this study. 3 SNPs and 19 candidate loci on chromosomes 1, 4 and 12 were associated with gs_{max}, across 3 GWAS methods. The gs_{max} candidate list included the *O. sativa* defined gene *Os*SAUR1, an Auxin responsive SAUR (small auxin up RNAs) gene family member (Jain et al., 2006). $iWUE_{max}$ was associated with one of the largest numbers of significantly associated 39 SNPS, with a candidate gene list 66 genes in regions on chromosomes 1, 2, 6 and 8 (Fig. 4.3b). Two methods identified SNP/trait associations for qP_{max}, with 5 SNPs and 13 genes on chromosomes 5 and 10.

For NPQ_{max} we identified 24 SNPs and 43 genes in regions on chromosomes 2 and 5, with significant SNPs across all GWAS methods. Of note, the candidate gene for ferredoxin NADP+ reductase chloroplast precursor and the *O. sativa* MADS-box transcription factor, *Os*MADS60, were identified for NPQ_{max}. Significant SNP associations were detected for $trmmol_{max}$ and VPD_{max} across all GWAS methods and VPD_{max} included large regions of significant SNP associations on chromosomes 4 and 7.

4.4.5 GWAS of dynamic photosynthesis-related traits

To our knowledge, this study completes the first GWAS analysis on dynamic photosynthesis-related traits. Significant SNPs were identified across 5 traits ($A_{i\ 10}$, $A_{i\ 90}$, $A_{i\ min}$, $A_{i\ rate}$ and $A_{i\ slope}$), which measured the induction curve phase of carbon assimilation (A). The greatest number of SNPs (38) and candidate genes (40) were associated with the steepness of $A_{i\ slope}$, with a region of note on chromosome 1 (Fig. 4.5a), which included the *O. sativa* genes *Os*MAN2 (endo-β-mannanase 2) and *Os*IAA6 (indole-3 acetic acid 6). The latter being a member of the rice Auxin/IAA gene family.

Significant SNPs were associated with 4 carbon assimilation traits ($A_{r\ 10}$, $A_{r\ 90}$, $A_{r\ slope}$ and $A_{r\ rate}$). To echo the results for carbon assimilation induction, the strongest associations with the steepness of the A relaxation slope, with 26 potential candidate genes. $A_{r\ rate}$ also displayed a strong signal on chromosome 12, with a putative YABBY domain containing gene.

Some of the most significant SNP/trait associations were with the dynamic stomatal conductance traits $gs_{i\ 90}$ and $gs_{i\ min}$. $gs_{i\ 90}$ associated with 42 significant SNPs and 24 candidate genes in regions on chromosomes 2 and 7 (Fig. 4.5b). Several of the candidate genes were found to be heat shock proteins. $gs_{i\ min}$ showed a region of 8 SNPs on chromosome 4 and $gs_{i\ slope}$ was associated with 16 SNPs, and 34 candidate genes, on chromosomes 3, 4 and 7. For the relaxation dynamics of stomatal conductance 3 traits ($gs_{r\ 10}$, $gs_{r\ 90}$ and $gs_{r\ rate}$) significant SNP/trait associations were detected with the FarmCPU and EMMA methods.

For non-photochemical quenching (NPQ) induction dynamics traits, $NPQ_{i\ slope}$ was associated to 1 SNP on chromosome 12, the same location was also identified when analysing $NPQ_{i\ 10}$ and includes a ca^{2+}/calmodulin-dependant protein kinase (CDPK) candidate gene. $NPQ_{i\ 10}$ also includes 23 other candidate genes, including two auxin-induced SAUR (small auxin up RNA) genes and two DUF1517 (domains of unknown function protein family) genes. We found 37 candidate genes associated with the $NPQ_{i\ rate}$. The only NPQ relaxation phenotype where significant SNPs were detected in two or more methods was $NPQ_{r\ 10}$, where 28 SNPs were identified in regions on chromosomes 6 and 7.

4.5 DISCUSSION

Under the current climate crisis and burgeoning human population, we are currently facing one of the most challenging periods of human history (Shukla et al., 2019. The ability to produce both climate resilient and productive crops will directly contribute to food security and the sustainable management of our environment (Ray *et al.*, 2019; Franco *et al*., 2020). But this is no easy task, it requires the trade-off between the two conflicting physiological processes, water use efficiency and photosynthesis (Lawson and Blatt, 2014). We have very little genetic understanding of these processes and even less of dynamic photosynthesis related traits, which are now known to be important for carbon assimilation and abiotic stress tolerance (Kromdijk et al., 2016; Taylor & Long, 2017; Caine et al., 2019). *O. glaberrima* is renowned for its resilience to challenging environments but the genetic basis of these traits is relatively uncharacterised.

Using 150 *O. glaberrima* re-sequenced accessions we have been able to identify genetic regions and candidate genes for 30 diverse traits across two-to-four GWAS methods. We

acknowledge the study uses a lower number of accessions, in comparison to most association analyses, which may limit the statistical power when identifying SNP/trait associations and the ability to detect rare or low frequency variants. For this reason, we used only reported results that were significant in two or more GWAS methods for added stringency and used a liberal threshold ($p<10^{-5}$). We appreciate this threshold is relaxed compared to the classical Bonferroni, but it was considered a necessary compromise and was used in a previous GWAS analysis in *O. glaberrima* by Cubry et al. (2020). A threshold based on a false discovery rate performed poorly, likely due to a lack of statistical power, and a Bonferroni threshold was considered too stringent under these circumstances. However, there is merit in exploring the traits where significant SNPs were identified in only one method if the trait is of particular interest. There is also substantial natural variation in the traits measured here, which will contribute to the successful detection of associated loci. Outlier data points in the raw dataset were removed to meet the parametric test assumptions of the association models used here. However, it is worth noting that these apparent outlying data points may be a true representation of the genetic variation in the population. Consequently, their removal may reduce the ability of the model to detect important low frequency variants and we recommend repeating the analysis with these outlying data points included. Furthermore, it is apparent that in some instances the GWAS methods used are overfitting. The reasons that are not apparent, as all model assumptions are met and we have accounted for confounding factors. On results that indicate a poor model fit (Fig. S4.5) we recommend considering re-running the analysis with all data points included, as this may affect the model fit, and reassessing confounding factors to ensure they were accurately accounted for.

Across these traits we identified 754 candidate genes, which may seem relatively few for the breadth of traits analysed here. However, these are all complex polygenic traits and are likely to exhibit multiple low effect sizes. When further exploring a particular suite of traits, we recommend ranking significant associations based on effect size and the threshold stringency could be reduced to $p<10^{-4}$ to further increase the pool of candidate loci, though this will be a trade-off with increased likelihood of false positive SNPs. The candidate gene lists generated here should be used with caution. The gene lists were generated based on *O. sativa*, due to the close synteny between the two species and because *O. glaberrima* lacks a fully annotated genome or genetic characterisation. It is possible there are novel trait-related regions in *O. glaberrima* that are not present in *O. sativa*. The candidate gene lists were generated from a 50kb window surrounding the significant SNP, we accept these are small compared to the calculated LD (150kb). Akin to the significant threshold ($p<10^{-5}$), the 50kb genomic window

was deemed an appropriate compromise between capturing the most likely candidates, without producing an overly large candidate list. However, these windows can be broadened in future, when focusing on a particular trait of interest. Considering the extent of LD (150kb) within this population (Cubry et al., 2020), and even the reduced genomic window, we lack the ability to accurately pinpoint responsible loci. Therefore, the gene lists may contain spurious results and functional studies are the only methods to truly establish if candidates are indeed trait related. This could also be improved in future by scaling the window based on local LD and estimated recombination rates.

4.5.1 *O. glaberrima* shows distinct physiology associations between ancestral groups

The abiotic resilience exhibited in *O. glaberrima* is attributed to the independent evolution to *O. sativa* and adaptation to a more variable African set of environments. Despite this, there are no comprehensive studies that tease apart the phenotypic variation and genetic structure within the population.

The *O. glaberrima* association panel used here demonstrates a diffuse population structure with relatively high amounts of admixture between ancestral groups, as seen in the ancestral matrix and PCA plots (Fig. 4.1a-b). This is beneficial for reducing confounding factors and successfully identifying significant SNP/trait associations. The weak population structure may be surprising in a species that evolved across Africa, with contrasting climates and elevation. However, historically *O. glaberrima* has been widely traded between African tribes and is still ceremonially important (Linares, 2002). This may account for what we see here and the lack of more distinct clustering of accessions that you would expect from separate agro-ecological origins, such as upland versus lowland accessions. The dendrogram showed distinct clustering of individuals that were collected from floating rice and shallow forest swamp ecologies. This suggests a group of accessions that may be physiologically distinct from other accessions and adapted for a specific environment. As far as we are aware, there is no description of the physiological traits associated with these agro-ecological niches and the environment itself. In future, it would be beneficial to utilise a range of accessions that are equally represented from ecological groups or countries of origin, as this study was unequally weighted.

The pairwise correlation between ancestry groups and phenotypic traits indicates that despite admixture within the *O. glaberrima* population used here, there is still physiological distinction between ancestral groups. Likely based on adaptation to adaptive pressures of

their specific African environment. Ancestry group 1 showed significant correlations to slower NPQ relaxation time and reduced biomass, further supporting the recent findings in this area (Kromdijk et al., 2016; Hubbart et al., 2018). While the variation displayed in this group is not desirable for crop improvement, it does provide a group of individuals that could be used to elucidate the genetic basis in the interplay between NPQ relaxation dynamics, biomass and stomatal traits. Other ancestry groups also demonstrate correlations related to the interplay of NPQ dynamics and stomatal traits (density and conductance), which may reflect the effect of the NPQ protein Photosystem II Subunit (PsBs) on stomatal opening and density, as observed by Głowacka et al. (2018). These are traits with a complex interplay, but the correlations shown here support that NPQ may be a major player relating to photosynthesis related traits with a solid genetic basis that can be harnessed.
The analysis performed in this area highlights useful information on the grouping of individuals based on environmental adaptation and demonstrates the need for a dedicated study in this area. We grouped accessions based on broad categories, such as country or ecological origin. Going forward, it would be interesting to also integrate how *O. glaberrima* ancestry groups are related based on bio-climatic variables at relatively fine resolution. While complex, the information gleaned from teasing apart these relationships may be essential when selecting groups of individuals with a genetic adaptation to abiotic stress tolerance for crop improvement.

4.5.2 Important genetic candidates for climate resilient traits

Recent research has shown that harnessing the combined effects of optimal stomatal density, capacity and rapid dynamic responses of photosynthesis related traits allows plants to respond instantaneously to fluctuating field conditions and stressful environmental stimuli, improving water use and productivity (Drake et al., 2013; Kromdijk et al., 2016; Taylor & Long, 2017; Vialet-Chabrand et al., 2017). Here, we have begun the journey in characterising the genetic basis of these important traits in a species known for its abiotic stress tolerance.

Among the morphological traits measured here, stomatal density candidate genes hold the most interest when selecting characteristics for climate resilience. For reasons unknown, Stomatal density (SD) ratio has been shown to have associations with photoprotection and relaxation dynamics in *O. glaberrima* (Cowling et al., 2021). Here we identified 16 candidate genes for SD ratio (Table S4.3). This list includes a potential candidate for WRKY18, a

transcription factor associated with abiotic stress response (Chen et al., 2010). In banana (*Musa acuminata*) *Musa*WRKY18 has been shown to be exhibit high expression in stomatal guard cells when under stress in banana (Tak et al., 2021). It is also thought to be important in abscisic acid (ABA) signalling, where Chen et al. (2010) identified reduced ABA sensitivity in *wrky18* knockout mutants. The ABA phytohormone plays an important role in mitigating drought stress and a regulatory role in stomatal movement (Fujita et al., 2006; Verslues & Zhu, 2007), thus this may be a good candidate to take forward for functional validation. Another key stomatal trait, maximal stomatal conductance (gs_{max}), showed significant associations to gibberellin (GA) and auxin related candidate genes in the same region on chromosome 1. The GA receptor gene GID1b identified here acts as a positive regulator in GA signalling (Ueguchi-Tanaka et al., 2005; Griffiths et al., 2006). While GA is important in plant growth, where *gid1* causes dwarfing, it has also been shown to be a regulator in stomatal development (González et al., 2017), root to shoot communication and the reduction of stomatal conductance during drought stress (Gaion & Carvalho, 2021). In other peaks, an interesting CAF1 (ccr4-associated factor 1) candidate was identified on chromosome 4. Increased CAF1 expression is induced by phytohormones under stressful environmental conditions and has been shown through expression studies to be of importance in biotic and abiotic stress tolerance (Wang et al., 2021).

Several highly significant SNPS were associated with water use efficiency ($iWUE_{max}$) on chromosomes 1 and 6. We identified several highly significant SNPs on chromosomes 1 and 6 for water use efficiency ($iWUE_{max}$). A candidate gene for a thylakoid lumenal protein was identified on chromosome 1. These proteins are essential for thylakoid function and restoration, which helps to prevent desiccation and facilitate rapid recovery after drought stress (Georgieva et al., 2010; Jarvit et al., 2013). A thylakoid lumenal protein was also identified for the photoprotection dynamic traits $NPQ_{i\ 10}$, $NPQ_{r\ 50}$ and $NPQ_{r\ 90}$. Thylakoid lumenal proteins have been shown to be important for the acclimation to fluctuating light conditions and photoprotection (Liu & Last, 2017).

Maximal non-photochemical quenching (NPQ_{max}) showed strong SNP/trait associations to a region on chromosome 2. This region included a candidate gene for ferredoxin NADP+ reductase chloroplast precursor, a precursor to ferredoxin-NADP+ oxidoreductase protein (FNR) that catalyses NADP+ reduction during the final steps of the Calvin cycle. The electron transfer process also generates a proton gradient across the thylakoid membrane, this drives ATP synthesis but also induces NPQ (Murchie & Ruban, 2020). *Arabidopsis thaliana fnr* mutants show higher levels of NPQ, demonstrating the ferredoxin NADP+ reductase

chloroplast precursor is an important gene when developing crop cultivars with appropriate levels of photo protection (Lintala et al., 2012; Bednarczyk et al., 2020). We feel this is a good candidate to explore further.

Due to the volume of dynamic traits measured here, we will discuss one example from stomatal conductance, carbon assimilation and non-photochemical quenching suites of traits. $gs_{i\ 90}$ was associated with significant region on chromosomes 2 (Table S4.3). Two DNAJ heat shock proteins, also known as Hsp or Hsf (heat shock proteins or transcription factors), were identified as candidates. These are chaperone proteins that interact with other heat shock proteins, mutants have been shown to be sensitive to ABA-mediated in stomatal closure, while over expression lines demonstrated improved drought and heat tolerance (Wang et al., 2015).

A region on chromosome 1 had the strongest association with $A_{i\ slope}$, and a number of candidate genes (Table S4.3). A candidate closest to a significant SNP was an ATP-dependant protease, many homologues of these proteases play essential roles in regulating the chloroplast stroma, by proxy they may also be important in chloroplast function for photosynthesis (Sjögren et al., 2006). A putative protein for NADP-ME (NADP-malic enzyme), an essential decarboxylase in C_4 species, the protein is non-photosynthetic in C_3 species but is associated with abiotic stress-resilience (Chen et al., 2019). A member of the auxin/IAA gene family, *Os*IAA6, that has been shown to be important in drought tolerance and tillering (Jung et al., 2015). *Os*MAN2, part of the endo-beta-mannanase gene family, thought to be important in cell wall synthesis (Yuan et al., 2007; Schröder et al., 2009). None of the 40 candidate genes for $A_{i\ slope}$ (Supp. Table S4.3) showed a clear relationship to the physical rapidity of carbon assimilation induction. However, many genes listed are biochemically characterised, but their functions are largely unknown.

The NPQ induction dynamic traits $NPQ_{i\ 10}$, $NPQ_{i\ 90}$ and $NPQ_{i\ slope}$ showed candidate genes in the DUF (domains of unknown function) and CDPK (calcium dependant protein kinase) gene families. Genes in the DUF gene families are thought to be involved with mitigating heat stress with recent research suggesting that *DUF* proteins are linked to the photosynthetic pathway and its regulation (Zhou et al., 2020; Nabi et al., 2021). CDPK proteins are a diverse superfamily, which are thought to be key mediators in a range of stimuli, including abiotic stress (Cheng et al., 2002). Calcium (Ca^{+}) and reactive oxygen species (ROS) are signalling molecules during photosynthesis when plants experience abiotic stress. ROS activate *CDPK* proteins, beginning a signalling cascade to induce abiotic stress mechanisms (Mohanta et al., 2018). NPQ is thought to limit ROS generation, so it could a be reasonable to suggest high

levels of ROS contribute to the induction of NPQ (Murchie & Ruban, 2020). As mentioned for $iWUE_{max}$ we also identified a putative thylakoid lumenal protein for $NPQ_{i\ 10}$, $NPQ_{r\ 50}$ and $NPQ_{r\ 90}$. Thylakoid lumenal proteins are required for abiotic stress tolerance, photosynthetic acclimation to fluctuating light and potentially important in photoprotection. NPQ regulation occurs in the lumen and this is the site where delta pH is sensed via proton accumulation, which causes an NPQ induction. The activity of the well-known NPQ regulatory PsbS (photosystem II subunit S) protein is modulated by the protons in the lumen, therefore it seems logical that other NPQ regulatory proteins could be found in the lumen. For example, a study in *Arabidopsis thaliana* showed that photoprotection impaired mutants grew normally when the thylakoid lumenal protein *mph2* was knocked out (Liu & Last, 2017).
Finally, we also identified cytochrome candidate genes in multiple dynamic and steady state photosynthesis related traits, including $iWUE_{max}$, $gs_{i\ slope}$, $gs_{r\ 10}$, $gs_{r\ rate}$, $NPQ_{i\ rate}$, $NPQ_{i\ 90}$ and stomatal density ratio. The cytochrome complex generates the proton gradient across the thylakoid membranes in chloroplasts and is essential in the electron transport chain step of photosynthesis (Hippler et al., 2021). The interruption of the electron transport chain can increase NPQ and it's been suggested that an increased turnover of the cytochrome complex increases photoprotection (Yoshida et al, 2006; Li et al., 2018). While the relevance of cytochrome proteins to photosynthesis and photoprotection is obvious, the associations to stomata related traits (*gs*, iWUE and SD) seen here is unclear. However, the consistency of cytochrome candidates across a range of photosynthesis related traits makes it clear this result should be pursued further.

4.5.3 Conclusion

This study reports the results of an extensive GWAS on 52 phenotypic traits related to climate resilience in *O. glaberrima*. This is the largest GWAS conducted in *O. glaberrima* and the only study to explore the genomic basis of the combined induction and relaxation dynamics for carbon assimilation, photo protection and stomatal conductance in any species. Our analysis identified candidate genes that have shown to be associated with the mitigation of drought and heat stress, including regulatory genes for key phytohormones, cytochrome and heat shock proteins. The efforts detailed here will provide a basis for the future filtering of key candidate genes, through trait-specific expert enrichment lists, and functional studies to validate gene-trait associations. Here, we provide useful information that can contribute to

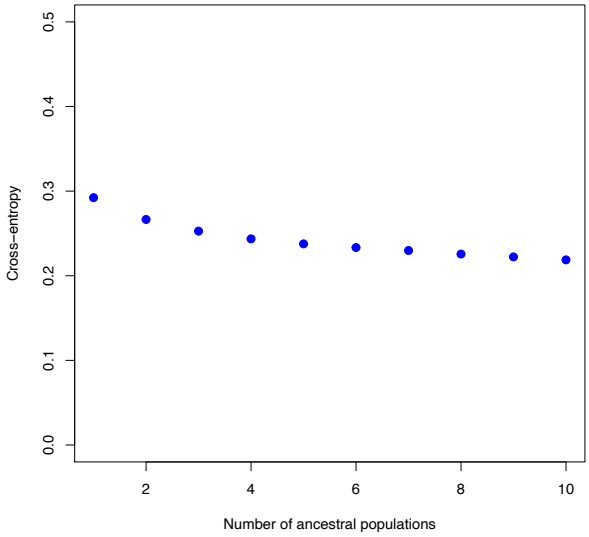
the selection of *O. glaberrima* accessions for crop breeding efforts and the long-term development of climate resilient crop varieties.

4.6 CHAPTER FOUR SUPPLEMENTARY MATERIAL

4.6.1 Supplementary Figures

Figure S4.1: (a) Cross entropy criterion values for SNP data (892539 markers) in a population of 155 *O. glaberrima* accessions, generated from the sNMF ancestry analysis, plotted against the estimated ancestry number (K). (b) The kinship matrix and heatmap were generated using the VanRaden genomic relationship method within the GAPIT association analysis method. The colour variation indicates the relatedness between individuals (darker indicates higher relatedness) in the population and the length of the branches is proportional to the genetic distance.

(a)



(b)

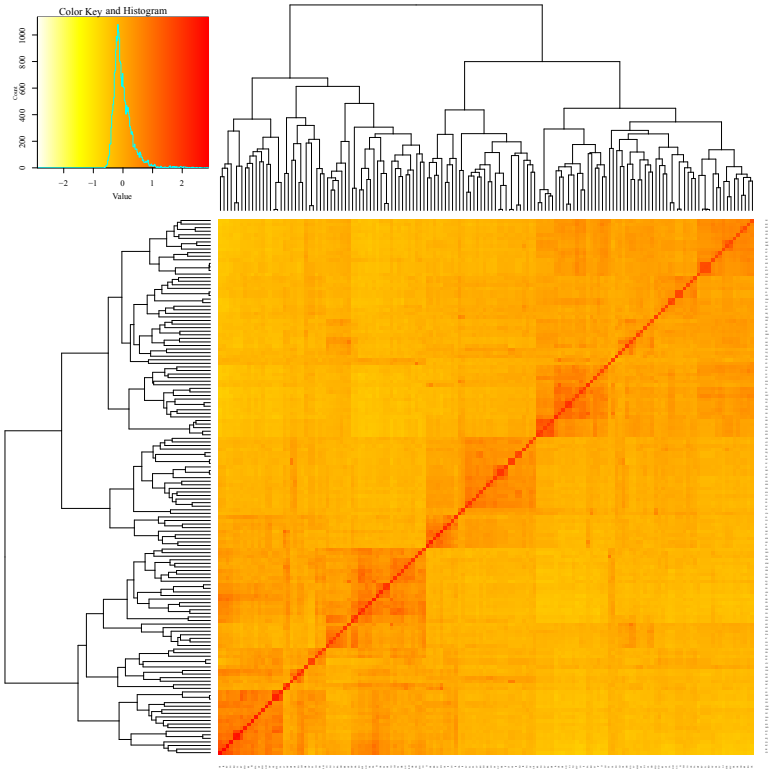


Figure S4.2: PCA output of SNP data, each individual point denotes a *O. glaberrima* accession, overlaid with colour relating to the accessions origin, shown in the legend in the bottom right of each plot, for (a) broad ecology category, (b) descriptive ecology category, (c) country and (d) African region.

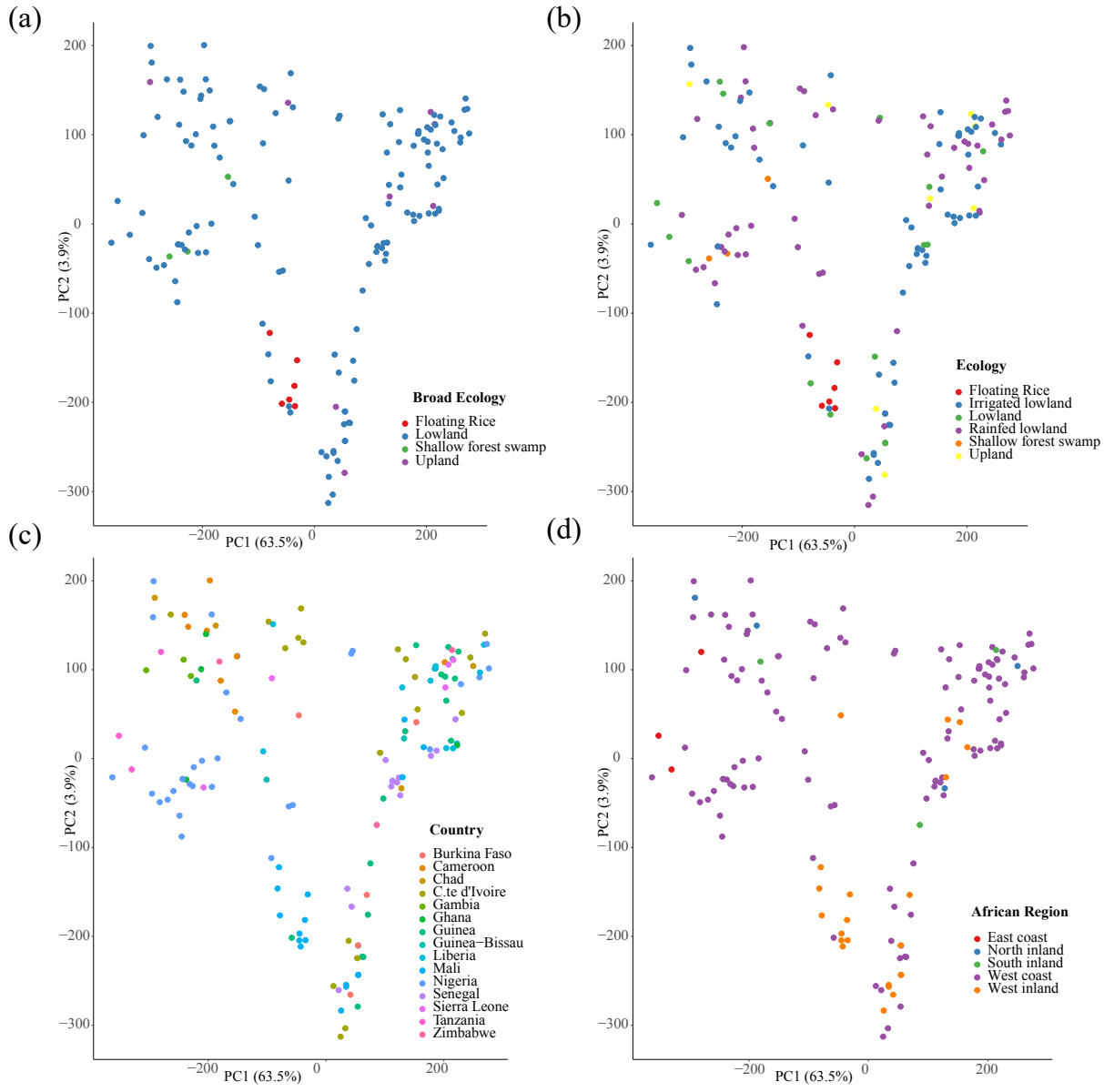
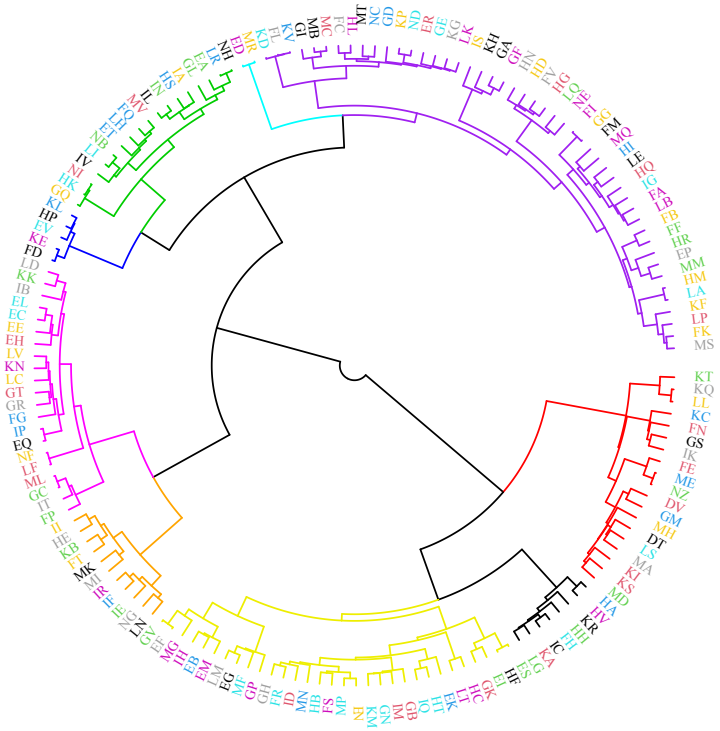
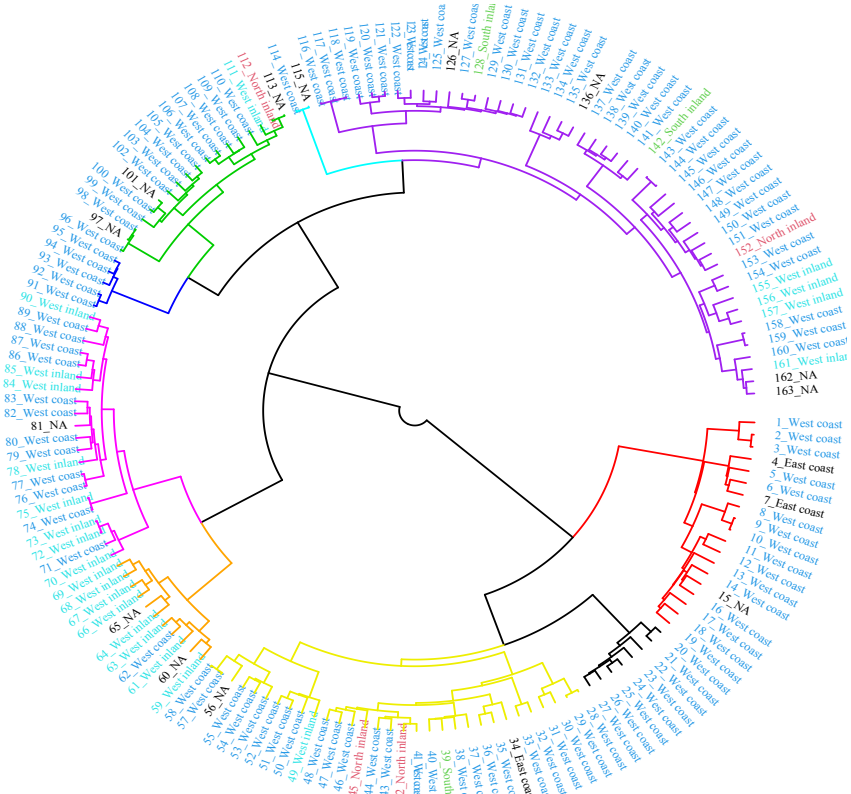


Figure S4.3: Dendrogram generated from a cluster analysis, computed using Nei's distance and the tree cut using k-means (k=9). The analysis was completed using SNP data in 155 *O. glaberrima* accessions. Each individual arm denotes a *O. glaberrima* accession, overlaid with colour relating to the (a) accession sequencing codes , (b) African region (blue = west coast, red = north inland, green = south inland, teal = west inland, black = east coast), and (c) agro-ecological niche (red = lowland, blue = upland, green = shallow forest swamp, black = floating rice, NK = not known).

(a)



(b)



(c)

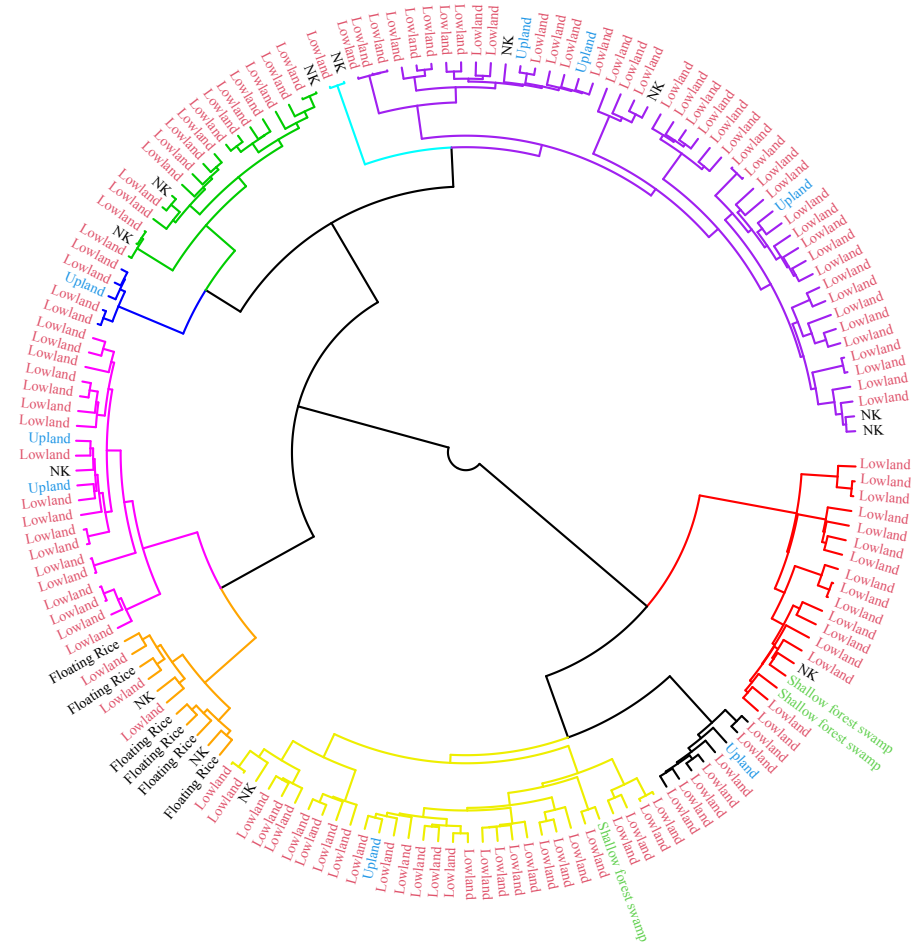


Figure S4.4: Pearson correlation matrix between four *O. glaberrima* ancestry groups and 52 phenotypic traits, the correlation coefficient values can be found in Table S.4.1. Correlations are displayed by the colour, indicated in the right-hand scale bar, filtered to only show associations p<0.1005. Asterisks indicate the significance between associations (∗∗∗*P*<0.001, ∗∗*P*<0.01, ∗*P*<0.05).

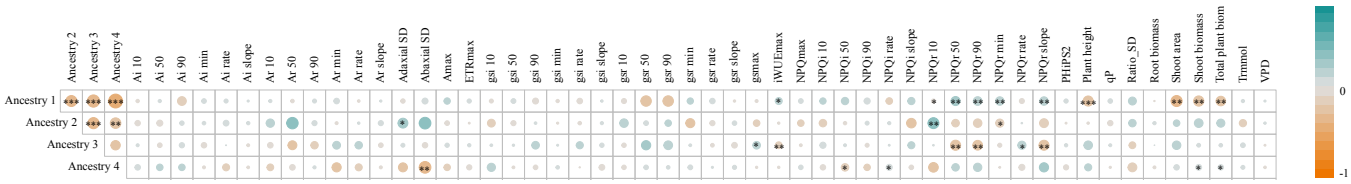


Figure S4.5: * Due to the size of this figure, it can be found the end of the thesis, after Chapter 5*. Complete list of GWAS Manhattan and diagnostic plots for all traits analysed and algorithms used. Each sub-figure shows a Manhattan plot for each of the four models used in the association analysis and the corresponding diagnostic qq-plot, to indicate the goodness of fit of the model. The red line on the Manhattan line indicates the significance threshold for this analysis, which is set at p<0.00001.

4.6.2 Supplementary Tables

Table S4.1: The table shows the correlation coefficients from a Pearson correlation test between *O. glaberrima* ancestry groups and phenotypic traits. 'A' refers to the ancestry group. Asterisks indicate the significance between associations (**** p<0.0001, ∗∗∗p<0.001, ∗∗p<0.01, ∗p<0.05).

A1	A1	A2	A3	A4
A2	-0.40****	1		
A3	-0.38****	-0.36****	1	
A4	-0.43****	-0.25**	-0.16	1
A_i_10	-0.03	-0.06	0.01	0.11
A_i_50	-0.02	0.03	-0.07	0.07
A_i_90	-0.15	0.11	0.05	0.01

A_i_min	-0.03	0	-0.08	0.14
A_i_rate	-0.03	-0.05	-0.01	0.1
A_i_slope	-0.01	0.09	0.04	-0.14
A_r_ED10	0.01	0.1	-0.02	-0.1
A_r_50	-0.02	0.16	-0.11	-0.04
A_r_90	-0.05	0.09	-0.08	0.05
A_r_Min	-0.01	-0.04	0.09	-0.04
A_r_rate	-0.06	-0.01	0.09	-0.01
A_r_Slope	0	0	-0.04	0.04
AdaxialSD	-0.02	0.16*	-0.06	-0.1
AbaxialSD	0.06	0.16	-0.04	-0.22**
Amax	-0.03	0	0.02	0.03
ETR	-0.06	0.05	0	0.02
gs_i_10	-0.09	-0.03	-0.02	0.16
gs_i_50	-0.09	0.02	0.01	0.08
gs_i_90	-0.1	0.08	0.04	-0.01
gs_i_min	-0.03	0.06	0.05	-0.09
gs_i_rate	-0.06	-0.08	0.14	0.03
gs_i_slope	-0.01	0.12	-0.02	-0.1
gs_r_10	0.06	0.1	-0.07	-0.1
gs_r_50	-0.11	0	0.14	-0.01
gs_r_90	-0.1	-0.01	0.14	0.01
gs_r_min	0	-0.05	0.01	0.04
gs_r_rate	0.11	0.03	-0.09	-0.07
gs_r_slope	0.01	-0.02	-0.04	0.06
gsmax	-0.12	-0.05	0.20*	0
WUE	0.19*	-0.04	-0.21**	0.04
NPQ	0.02	-0.1	0.03	0.05
NPQ_i_10	0.04	-0.11	0.13	-0.06
NPQ_i_50	0.07	-0.08	0.15	-0.16*
NPQ_i_90	0.12	-0.04	0.04	-0.15
NPQ_i_rate	-0.12	0.04	-0.05	0.16*
NPQ_i_slope	0.04	-0.13	0.09	0

NPQ_r_10	-0.16*	0.22**	0.07	-0.12
NPQ_r_50	0.24**	-0.08	-0.23**	0.03
NPQ_r_90	0.25**	-0.12	-0.23**	0.05
NPQ_r_min	0.23**	-0.19*	-0.01	-0.08
NPQ_r_rate	-0.15	-0.02	0.20*	0.01
NPQ_r_slope	0.23**	-0.1	-0.26**	0.09
PHiPS2	-0.06	0.05	0	0.02
Plant_height	-0.30***	0.08	0.13	0.16
qP	-0.06	0.02	-0.02	0.08
Ratio_SD	0.1	-0.01	0.03	-0.16
RootBiomass	-0.07	0.06	0.01	0.01
Shoot_area	-0.22**	0.07	0.05	0.15
ShootBiomass	-0.25**	0.12	-0.01	0.20*
Total_plant_biomass	-0.23**	0.11	-0.01	0.18*
Trmmol	-0.07	-0.13	0.13	0.09
VPD	0.04	-0.02	-0.08	0.05

Table S4.2: Table detailing a list of traits analysed against each GWAS test (LFMM, FarmCPU, GAPIT and EMMA). 'Yes' indicates significant trait-related SNPs were detected using a particular GWAS method, 'NA' means a non-significant result. The number of significant GWAS results is colour coded for ease.

No. Sig Tests

1

2

3

4

Trait	**LFMM**	**FarmCPU**	**GAPIT**	**EMMA**
$A_{i\ 10}$	Yes	NA	NA	Yes
$A_{i\ 50}$	NA	NA	NA	Yes
$A_{i\ 90}$	Yes	NA	NA	Yes

$A_{i\ min}$	Yes	Yes	NA	Yes
$A_{i\ rate}$	Yes	Yes	Yes	Yes
$A_{i\ slope}$	Yes	Yes	Yes	Yes
$A_{r\ 10}$	Yes	Yes	NA	NA
$A_{r\ 50}$	Yes	NA	NA	NA
$A_{r\ 90}$	Yes	Yes	NA	NA
$A_{r\ min}$	NA	NA	NA	NA
$A_{r\ rate}$	Yes	Yes	NA	Yes
$A_{r\ slope}$	Yes	Yes	NA	NA
Abaxial SD	Yes	NA	NA	Yes
Adaxial SD	NA	NA	NA	Yes
A_{max}	NA	NA	NA	NA
ETR_{max}	NA	NA	NA	Yes
$gs_{i\ 10}$	Yes	NA	NA	NA
$gs_{i\ 50}$	NA	NA	NA	NA
$gs_{i\ 90}$	Yes	Yes	Yes	Yes
$gs_{i\ min}$	Yes	Yes	Yes	NA
$gs_{i\ rate}$	NA	NA	NA	NA
$gs_{i\ slope}$	Yes	Yes	NA	Yes
$gs_{r\ 10}$	NA	Yes	NA	Yes
$gs_{r\ 50}$	NA	NA	NA	NA
$gs_{r\ 90}$	NA	Yes	NA	Yes
$gs_{r\ min}$	NA	NA	NA	NA
$gs_{r\ rate}$	Yes	Yes	NA	NA
$gs_{r\ slope}$	Yes	NA	NA	NA
gs_{max}	Yes	Yes	NA	Yes
$iWUE_{ma}$	Yes	Yes	NA	Yes
NPQ_{max}	Yes	Yes	Yes	Yes
$NPQ_{i\ 10}$	Yes	NA	NA	Yes
$NPQ_{i\ 50}$	NA	Yes	NA	NA
$NPQ_{i\ 90}$	Yes	Yes	NA	Yes
$NPQ_{i\ rate}$	Yes	Yes	NA	Yes
$NPQ_{i\ slope}$	NA	Yes	NA	Yes

$NPQ_{r\ 10}$	Yes	Yes	NA	Yes
$NPQ_{r\ 50}$	NA	NA	NA	Yes
$NPQ_{r\ 90}$	NA	NA	NA	Yes
$NPQ_{r\ min}$	NA	NA	NA	NA
$NPQ_{r\ rate}$	NA	NA	NA	Yes
$NPQ_{r\ slope}$	NA	NA	NA	Yes
$\Phi PSII_{max}$	NA	NA	NA	Yes
Plant height	NA	Yes	Yes	Yes
qP_{max}	Yes	NA	NA	Yes
Ratio SD	Yes	Yes	NA	NA
Root biomass	NA	NA	NA	NA
Shoot area	Yes	Yes	Yes	Yes
Shoot biomass	NA	NA	NA	Yes
Total biomass	NA	NA	NA	Yes
$Trmmol_{max}$	Yes	Yes	Yes	Yes
VPD_{ma}	Yes	Yes	Yes	Yes

Table S4.3: * Due to the size of this table, it can be found the end of the thesis, after Chapter 5*. Gene list for 30 phenotypic traits from the GWAS analysis detailed in chapter 4. The traits listed here were identified in 2 or more GWAS methods, SNPs were filtered based on a $p<10^{-5}$ threshold.

CHAPTER FIVE

Thesis Discussion

5.1 Introduction

Recent research has shown that crop yields are already decreasing under rising global temperatures and will continue to do so under the projected future climate conditions (Zhao et al., 2017; Ray et al., 2019). African rice (*O. glaberrima*) is comparatively more tolerant to many abiotic stresses than Asian rice (*O. sativa*), including drought and heat stress, because of extensive adaptation to variable African environments (Albar et al., 2003; Linares, 2002; Orjuela et al., 2014). While *O. glaberrima* is not suitable for commercial rice production due to harvesting difficulties and low yields (Linares, 2002) it does provide a novel source of genetic variation.

The potential *O. glaberrima* represents as a source for crop improvement has been recognised for over two decades (Ghesquière et al., 1997; Sarla & Swamy, 2005) but there is limited literature describing the phenotypic and genetic variation present. The recent genomic re-sequencing of 163 accessions has provided a new genomic resource to fully explore the potential *O. glaberrima* has to offer (Cubry et al., 2018). Using 155 of these re-sequenced accessions we characterised the phenotypic variation in 58 morphology, steady-state and dynamic photosynthesis related traits. Using this phenotypic data, a genome wide association study (GWAS) was completed using four different algorithms. Forty-four traits found to have significant SNP/trait associations were identified and over 700 associated candidate genes, many known for their role in drought, heat and abiotic stress tolerance. Multi-variate analyses were used to tease apart the ecological and environmental factors that contribute to trait and genetic adaption in *O. glaberrima*. This information will be essential when selecting accessions for specific traits of interest and/or adapted to specific environmental conditions. This thesis describes the analysis of complex traits and their underlying genetic basis and provides a solid base to further elucidate the functional validity of these findings.

5.2 Key Findings

This study aimed to explore the heritable variation in the *O. glaberrima* germplasm, to contribute towards the potential improvement of rice in future climate. The findings from this research can be described in three key sections, which are summarised below.

5.2.1 Characterising heritable phenotypic variation in *O. glaberrima*

Chapter three describes extensive natural variation in traits associated with increased photosynthetic efficiency, photoprotection and water use efficiency.

- Chapter two describes the development of a novel automated Machine Learning software that accurately identifies and counts *Oryza spp.* stomata from micrograph images. This allowed the characterisation of the variation in stomatal density across 155 *O. glaberrima* accessions.
- As *O. glaberrima* is proposed as a source of novel genetic variation for *O. sativa* crop improvement, it is useful to compare the phenotypic variation between the two species. In comparison to *O. glaberrima*, IR64 showed a slightly lower gs_{max} and higher iWUE*max*, but exhibited slower stomatal conductance, carbon assimilation and non-photochemical quenching (NPQ) induction dynamics. The two species also clustered separately in the principal component (PCA) and hierarchical (H)-clustering analyses, based on phenotypic differences. These findings suggested *O. glaberrima* has a range of unique phenotypic variation, compared to *O. sativa*, with an indication of adaptation to drier conditions that require rapid stomatal movements and responses to high light. However, these findings should be interpreted with caution as only a single *O. sativa* genotype was used.
- Ubiquitous associations were found between NPQ relaxation dynamics, during the transition from high – low light, and increased biomass. These findings support research that rapid NPQ relaxation can enhance productivity via the faster recovery of photosynthetic efficiency (Kromdijk *et al.*, 2016; Hubbart *et al.*, 2018; Murchie and Ruban, 2020). This was further supported by the finding that accessions with rapid NPQ relaxation, and a low NPQ capacity, exhibited higher carbon assimilation under low light. The ability to rapidly respond to fluctuating conditions is important at all levels of the canopy. The flag leaf is estimated to provide 50% of assimilates and is closely linked to rice yields (Nakano et al., 1995). It is essential the flag leaf can maximise photosynthetic efficiency, but the lower canopy is still of great importance, as it provides the remainder of assimilates. Therefore, it is important that at all levels the plant canopy can rapidly respond to changes in light that result from canopy architecture, cloud cover and wind (Burgess et al., 2016), to maximise photosynthetic processes and productivity.

- Unexpected but consistent correlations were identified between NPQ and stomatal conductance dynamic, and steady state traits. This may indicate the importance of the key NPQ protein, Photosystem II Subunit (PsbS) on stomatal conductance (Głowacka et al*.,* 2018). Other associations were identified between stomatal density ratio, NPQ_{max} and the level of NPQ achieved under low light ($NPQ_{r\ min}$), for unknown reasons.
- Heritable (H^2) variation and underlying genetic variation (PGV) was identified in many traits. Heritability values were comparable to similar traits in maize (Choquette *et al*., 2019) but lower than those described in *O. sativa* (Qu *et al*., 2017). A genetic component was indicated in most traits and this was supported through later the GWAS findings (chapter four).
- An automated Machine Learning software was developed that accurately identifies and counts *Oryza spp.* stomata from micrograph images.

5.2.2 Identifying the genetic basis of complex traits

Chapter four reports the genome wide association study completed on the phenotypic traits described in chapter three. The following are highlights of the most promising trait gene associations identified from the research:

- A GWAS analysis using four methods identified significant (threshold $p<10^{-5}$) SNP trait associations for 44 out of 52 traits. Thirty of these were identified using two or more methods.
- Stomatal candidate genes were identified that have phytohormone regulatory roles associated with stomatal movement and conductance. The results for abaxial to adaxial stomatal density ratio included a candidate for *WRKY18*, which is highly expressed in stomatal guard cells (Tak et al., 2021). *WRKY18* is also thought to be important in ABA signalling, a phytohormone that regulates stomatal movement (Fujita et al., 2006; Verslues & Zhu, 2007; Chen et al., 2010). gs_{max} showed significant associations to the gibberellin (GA) receptor, GID1B, which acts as a positive regulator in GA signalling (Ueguchi-Tanaka et al., 2005; Griffiths et al., 2006) and reduces stomatal conductance during drought stress (Gaion & Carvalho, 2021).

- iWUE$_{max}$ and NPQ dynamic traits included thylakoid lumenal proteins as candidate genes. These proteins restore and maintain thylakoid function, and have been shown to be important for photoprotection and abiotic stress tolerance responses (Georgieva et al., 2010; Jarvit et al., 2013; Liu & Last, 2017).
- NPQ$_{max}$ showed strong SNP/trait association to a region that includes a candidate gene for a ferredoxin NADP+ reductase chloroplast, a precursor protein to the ferredoxin-NADP+ oxidoreductase protein (FNR). FNR catalyses NADP+ reduction during the final steps of the Calvin cycle. The *FNR* gene, and therefore it's precursor, are important in NPQ induction (Lintala et al., 2012; Bednarczyk et al., 2020).
- The NPQ induction dynamic traits NPQ$_{i\ 10}$, NPQ$_{i\ 90}$ and NPQ$_{i\ slope}$ all included candidate genes in the *CDPK* (calcium dependant protein kinase) gene family. CDPK proteins are triggered by high levels of Ca^{2+} ions, caused by abiotic stresses, triggering a cascade that activates abiotic stress resilience mechanisms. This causes the production of reactive oxygen species (ROS) during abiotic stress (Kumar et al., 2018) and NPQ limits ROS generation. It is possible that CDPK proteins may contribute to the process of NPQ induction (Mohanta et al., 2018; Murchie & Ruban, 2020).
- Cytochrome candidate genes were implicated in multiple traits, including iWUE$_{max}$, $gs_{i\ slope}$, $gs_{r\ 10}$, $gs_{r\ rate}$, NPQ$_{i\ rate}$, NPQ$_{i\ 90}$ and stomatal density ratio. Cytochrome proteins are essential during the electron transport chain stage of photosynthesis and genetic variation in these genes may affect a range of complex photosynthetic traits.

5.2.3 Ecological and environmental impacts on phenotypic and genetic adaptation

The scientific literature describes that *O. glaberrim*a harbours exceptional abiotic stress resilience, in comparison to *O. sativa* (Ghesquière et al., 1997; Sarla & Swamy, 2005). This could be attributed adaptation to less intense agricultural inputs, domestic selection and the evolution to different selection pressures than *O. sativa*. In this context, it seems important to understand the association between environmental pressures and suites of traits and genotypes. This can inform the selection of useful accessions for future research. This area of work was explored in chapters three and four, the findings are summarised below:

- A PCA and clustering analysis identified three distinct clusters of accessions with adaptations to different environments and growth strategies. Cluster 2 contained individuals with physiologies geared towards water conservation, with high levels of water use efficiency, low gs_{max}, fast relaxation times for stomatal conductance and carbon assimilation. Whereas cluster 3 individuals were consistent with productive lowland phenotypes, with accessions demonstrating low levels of NPQ under low light, high trait values for A_{max}, shoot area, shoot and root biomass. Traits attributed to individuals in cluster 1 suggest an intermediate type, with accessions showing rapid induction and relaxation rates for A and NPQ but high gs_{max} and slow gs relaxation dynamics, which is reflected in a low water use efficiency.
- A PCA of trait data, overlaid with agro-ecological and country origins, showed a distinct clustering between upland and lowland accessions. However, this was not fully supported in the genetic dendrogram analysis in chapter four but there was distinct clustering of accessions that were collected from floating rice and shallow forest swamp ecologies. The multivariate analyses in chapter three demonstrated a degree of clustering of individuals based on country of origin. This was supported to some extent in chapter four, where individuals from Mali and Nigeria form distinct clades in the dendrogram.
- In chapter three the effect of environmental variables on trait adaptation was explored. We identified a significant correlation between the PCs 1 and 4, generated from the PCA for phenotypic and climatic variables respectively. Suggesting that traits contributing to PC1 adapted in response the selective drivers of elevation, and the combined effort of temperature and precipitation. Further to this, a hierarchical-clustering analysis identified three clusters, showing accessions with common sources of climatic variation.
- A pairwise correlation between ancestry groups and phenotypic traits indicated physical similarities between groups. Ancestry group 1 showed significant correlations to NPQ relaxation dynamics and biomass. While other ancestry groups showed correlations between NPQ and stomatal traits. These findings echo those identified in chapter three, where correlations were identified between these traits.

5.3 Research Limitations

All research projects have their limitations, such as constraints in methodology or experimental design, that can have implications on the research outcomes. This section identifies the key limitations in the study, how they were addressed to minimise any potential impact and what improvements could have been made.

5.3.1 Stomatal counting method

- When developing the automated stomatal counting method, detailed in chapter two, only a relatively small number of training images were used. However, we were still satisfied with the accuracy of the resource for the purposes of this study (r=0.94 between human and automated counts, see chapter two for more details).
- Finally, when testing the accuracy of the software against human manual counts, it would have been useful to quantify human error. The automated counting may be more accurate than the human manual counts, but this was not explored and could affect the ability to accurately quantify the accuracy of the method.

5.3.2 Glasshouse experimental design

- Limitations were identified in the experimental plan, designed for the phenotyping efforts described in chapter three. This were largely an effect of limitations for tropical glasshouse space, where direct soil planting was possible.

  The experimental plan grouped accessions into rows of five replicates, where twelve accessions were grown and measured together. The lack of random design and measurement of accessions across the seasonal transition from summer to autumn introduces opportunities for spatial and temporal variation. The systematic design was chosen to minimise the possibility of genotype misidentification and damaging plants by moving large IRGA machines.
- To reduce the variation due to temporal and spatial effects, a linear mixed-effects model (lme) was used to account for the random effects that may have affected the population, including sowing and measurement date, location in glasshouse and IRGA machine. The lme model coefficients quantify the amount of variation attributed to each effect included in the model. It was found that sowing date was consistently the effect with the highest source of variation, while the position in the glasshouse often had negligible variation attributed to it. We used BLUPs here, as

they are commonly used to account for the random effects when measuring plants in fluctuating conditions (Robinson, 1991; Merk et al., 2012; Zendonadi et al., 2021). However, there are alternative methods that could also be used to account for this, such as applying a spatial smoother using a generalised additive model, as described in Campos et al. (2021).

- *O. sativa* IR64 was used as a check-genotype during this phenotyping experiment but was not introduced until the fifth rotation of accessions. Therefore, it could not be used to normalise the entire dataset against. Ideally, two *O. sativa* genotypes (IR64 and Nipponbare, the latter being a hardier cultivar), should have been included from the beginning of the phenotyping efforts. It would have added consistency and assisted with normalisation for environmental effects.
- In hindsight, a randomised design, proceeded by a power experiment constructed with the assistance of a biological statistician, would have preferable. Alternatively, the phenotyping could have been completed in controlled growth cabinets to minimise for environmental effects, but from experience the species does not grow well in these compared to a glasshouse.
- A duplicate experiment the following year would have been desirable, to ensure trait measurements could be replicated. However, we have confidence in the phenotypic measurements. The thesis of Treeintong (2019) details the selection of extreme traits and accessions from the phenotyping experiment in chapter three, the subsequent replicated measurements are supported in his experiment.

5.3.3 The estimation of heritability and GWAS statistical power

- Heritability (H^2) values can be useful when choosing traits of interest to pursue. The values calculated here may vary dependant on the method used to calculate broad sense heritability. Our values were generated by using the lme model coefficients, whereby the variation attributed to genotype was divided by the total variation attributed to all effects included in the model. However, there are numerous methods that can be employed to calculate the phenotype portion of heritability, which can substantially alter the H^2 value. For example, when the method used in Cubry et al. (2020) was recently trialled, it increased all H^2 estimates for traits measured here. While it may be worth dedicating time to identifying the most accurate method of

heritability estimation, as they can be a useful guide, H^2 values are not essential when identifying the genetic basis in traits of interest.

- It is recognised that the number of accessions used in the association analysis is relatively low for a GWAS, where the norm is >300 genotypes, and will limit the statistical power when detecting significant SNPs. To maximise the likelihood of detecting significant associations we accounted for confounding factors, used a relatively low significance threshold and multiple GWAS methods that are known for optimising statistical power.
- It is also noted that for some traits and methods, the GWAS models were likely to be overfitting. This may be due to unresolved confounding factors and will need to be further examination before the manuscript is submitted for publication.

5.4 Recommendations for Future Research

This study has demonstrated the importance of *O. glaberrima* as an interesting and potentially useful source of heritable variation. However, it is just the beginning of a long path to truly elucidating the interplay between photosynthetic traits, genetic basis of the traits discussed here and their translation to crop improvement. In this section the immediate recommendations for future research will be discussed, based on the research goals of physiology, genetics and population adaptation.

5.4.1 Physiology research

The characterisation of *O. Glaberrima* traits and their interactions has raised many interesting avenues that could be pursued in future:

- An obvious first step would be to complete drought experiments on the population used here. This would be used to establish if accessions that display the desirable water conservation strategies detailed in chapter three, are in fact more resilient to water limited conditions. Furthermore, it would be interesting to include a range of *O. sativa* cultivars, to establish the degree of drought resilience between the two species.
- Consistent associations were identified between NPQ relaxation dynamics and biomass. Future research should be completed to confirm if rapid NPQ relaxation

not only equates to greater biomass accumulation in *O. glaberrima*, but also greater yield.

- Many correlations were identified between NPQ and stomatal traits, it would be useful to explore the basis of these relationships. In particular, the impact of the key NPQ protein, Photosystem II Subunit (PsbS), on stomatal conductance and density, using knock out mutants and expression studies in *O. sativa*.
- Future research could include a dedicated stomatal characterisation in *O. glaberrima*. A wide variation in stomatal morphology was observed, such a prominence in the epidermis and size, but was not quantified due to time constraints. Stomatal morphology may have a greater impact on associated traits, such as water use efficiency and dynamics, than simply stomatal density.
- In this study we were not able to establish if photosynthetic dynamics are mostly limited by stomata or biochemistry. *O. sativa* photosynthetic induction has been shown to be predominantly limited by Rubisco content (Acevedo-Siaca et al., 2021). Therefore, it would be interesting to explore the Rubisco content across the accessions in the *O. glaberrima* resource used here and where the limitation in photosynthesis dynamics lies.
- This study does include the measurement of root biomass but does not progress into characterising the natural variation of root density, architecture, or anatomy (such as aerenchyma formation) in *O. glaberrima*. It would be interesting to associate these traits at a whole plant level and describe root-shoot associations.

5.4.2 Genetic characterisation

The GWAS detailed here was the first step towards generating a formal list of trait candidate gene lists and elucidating gene-trait associations in *O. glaberrima*.

- Where possible, we recommend the candidate gene lists generated here are further informed using expert gene lists generated from *O. sativa*. This is a list of genes, obtained from the literature, that are known to be associated with a trait of interest. The candidate genes can then be checked for enrichment against the expert gene list. Although, this will not be possible for all traits of interest, such as photosynthesis dynamics, but only for those that have a well characterised genetic basis in *O. sativa*. As the candidate gene lists were also generated using the SNP2gene tool, it would be

useful to dedicate time to manually searching for candidate genes in specific traits of interest.

- Key candidate genes for traits of interest, such as *GID1B* for stomatal conductance, can be selected and explore using functional studies to validate causal genes.
- Long read re-sequencing efforts are currently underway in *O. glaberrima* by IRD, University of Montpellier. Once this are complete, it is recommended to re-run the GWAS in case of regions that may be missed.
- Considering the Illumina re-sequencing efforts were completed at a higher enough depth, the generation of a full *O. glaberrima* chromosome-level reference genome should be developed.
- Finally, the driving research rationale in this thesis was the translation of useful natural variation in *O. glaberrima* to improve the resilience of *O. sativa* in a changing climate. Therefore, much consideration should be given as to how the research reported here can contribute towards the testing, selection and introgression of useful traits between the species. The translation of useful genomic regions between the two species could take place using traditional breeding methods, where the identification and generation of flanking genetic markers and chromosome substitution lines could be used to guide introgression. However, the S_1 locus incompatibility between the two species is still a challenge, but good progress is being made by the development of interspecific bridge lines that are homozygous for the S_1^s allele (Wambugu et al., 2019). Parallel efforts should also be dedicated to the identification of causal genes and mutations in traits of interest, that can be implemented in *O. sativa* through genome editing tools. While there are also challenges here due to the controversy of genetically modified crops, these are becoming increased accepted as essential to future food security (Gao, 2021).

5.4.3 Exploration of adaptative pressures

- The research completed in chapters three and four shows interesting groups of individuals with specific phenotypic and genetic environmental adaptation traits. This highlights the need for a dedicated study to fully elucidate the environmental and agro-ecological trait adaptation of *O. glaberrima*.
- AfricaRice has a germplasm of over 2000 *O. glaberrima* accessions, using this it would be possible to obtain equally represented accessions from a range of ecologies

and assessing physiological adaptation to climatic variables at a range of spatial scales to tease apart these complex relationships.

5.4.4 Developing the stomata detector software

- The Stomata Detector software is a resource that is ready to use for the scientific community and does not require specialist skills. However, to be truly useful the software would require re-training with a greater number of training images and broad range of taxa. A larger training set, ideally across key crop species (rice, maize, wheat, soy), would improve the accuracy and prove a more useful resource for the scientific community. It would also be of particular use to the plant science community if the software could also measure stomatal morphology, such as pore size, guard cell length and width.
- Currently, there is no reliable resource that can automate the counting of stomata from micrographs. Fetter et al. (2019) released a promising web portal resource after the development of Stomata Detector, but from experience it was never functional and is currently exhibiting an error message. Jayakody et al. (2021) have developed a method that can reliably identify and bound stomatal guard cells across taxa. However, like many other stomatal resources, accessing and using the material requires a degree of computational knowledge, such as the use of GitHub and coding.

5.5 Conclusion

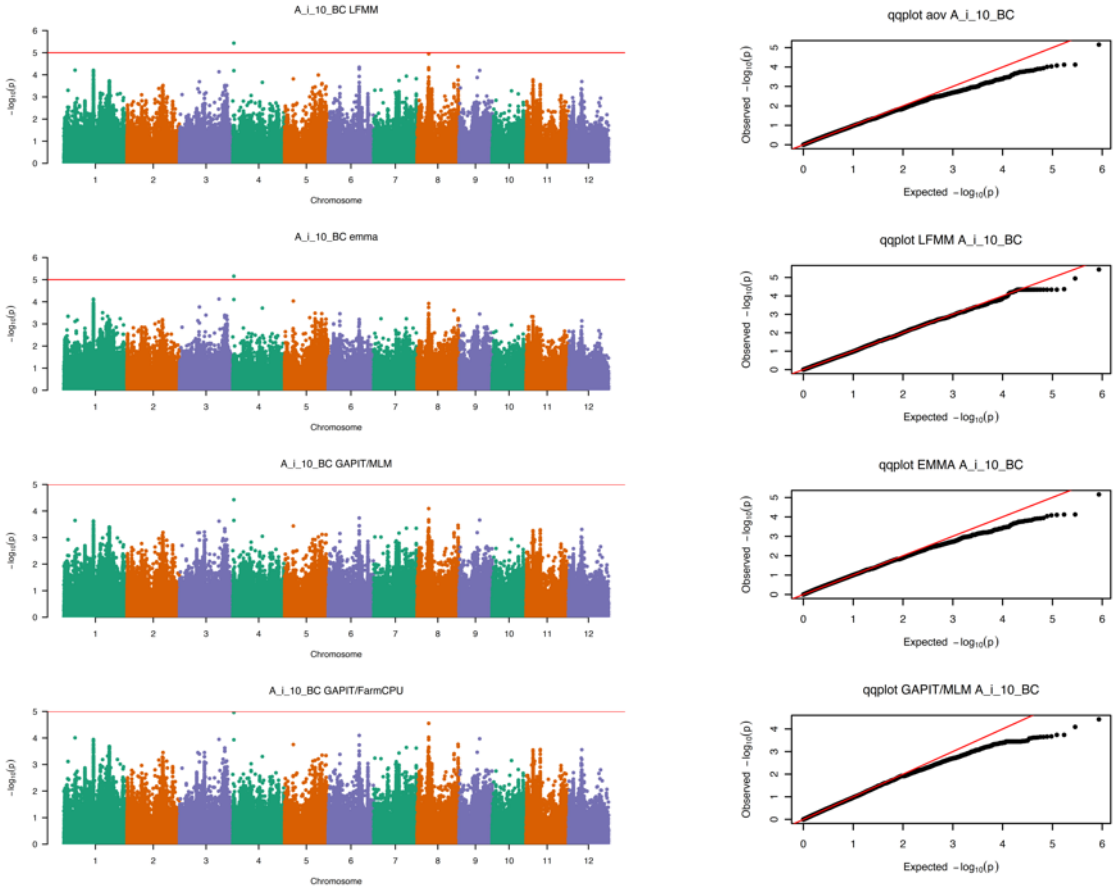
This thesis documents the results of extensive phenotypic characterisation, and the subsequent association study, in the reported abiotic stress tolerant species *O. glaberrima.* Promising heritable variation was identified in a range of important traits relating to climate resilience and productivity. The GWAS highlighted significant gene-trait associations and candidate genes known with previous evidence from other species of mitigating drought and heat stress, including regulatory genes for key phytohormones, cytochrome and heat shock proteins. Multivariate and genetic analyses completed here show the clustering of accessions with similar physical qualities, indicative of environmental adaptation strategies.
This novel data can be used to identify interesting traits for physiology research, the functional validation of candidate genes and assist in the selection of accessions for crop improvement efforts. Finally, I hope the contents of this thesis can be used to bridge the gap

between trait characterisation and the translation for crop improvement in future changing climates.

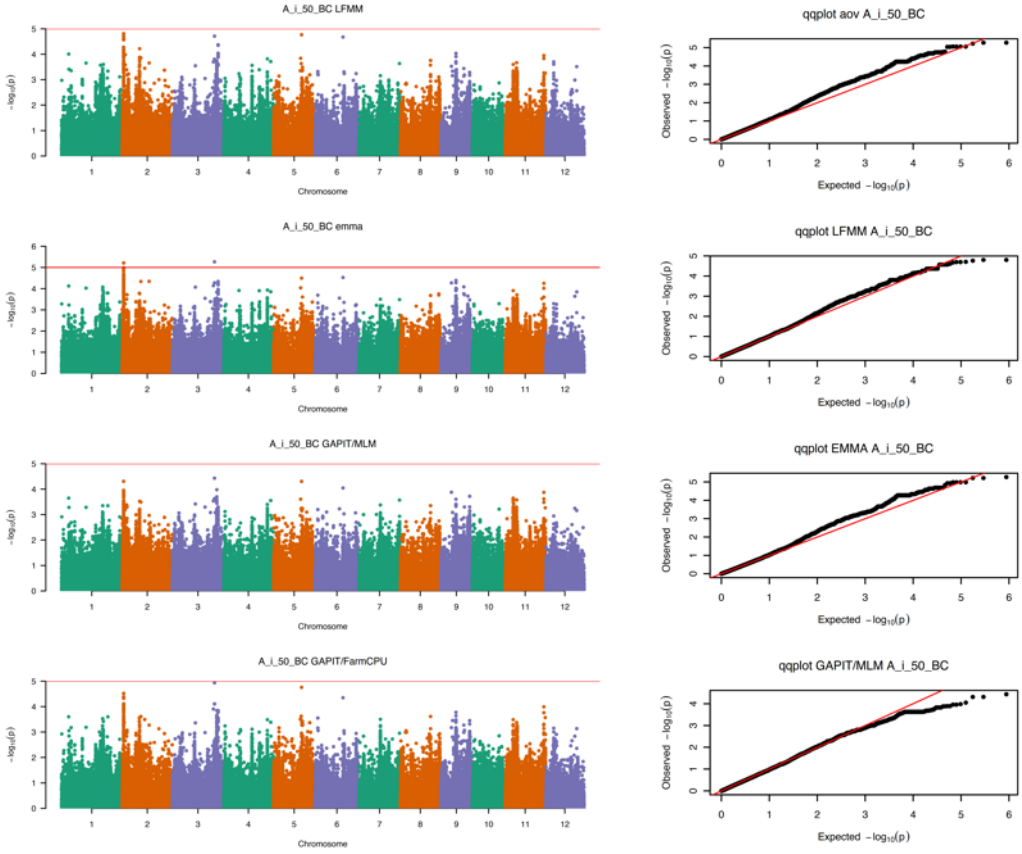
CHAPTER FOUR SUPPLEMENTARY MATERIAL FIGURE S4.5 AND TABLE S4.3

Figure S4.5: Complete list of GWAS Manhattan and diagnostic plots for all traits analysed and algorithms used. Each sub-figure shows a Manhattan plot for each of the four models used (from top to bottom, LFMM, EMMA, GAPIT and FarmCPU) in the association analysis and the corresponding diagnostic qq-plot, to indicate the goodness of fit of the model. The red line on the Manhattan line indicates the significance threshold for this analysis, which is set at $p<0.00001$.

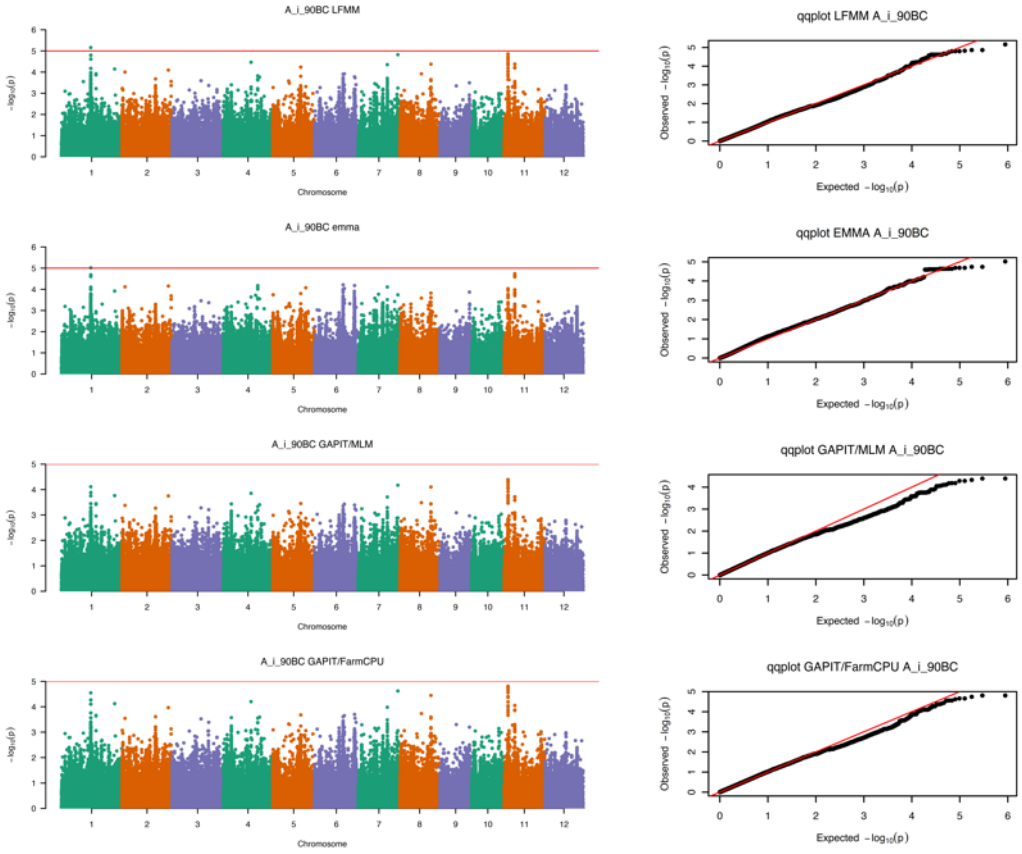
$A_{i\ 10}$



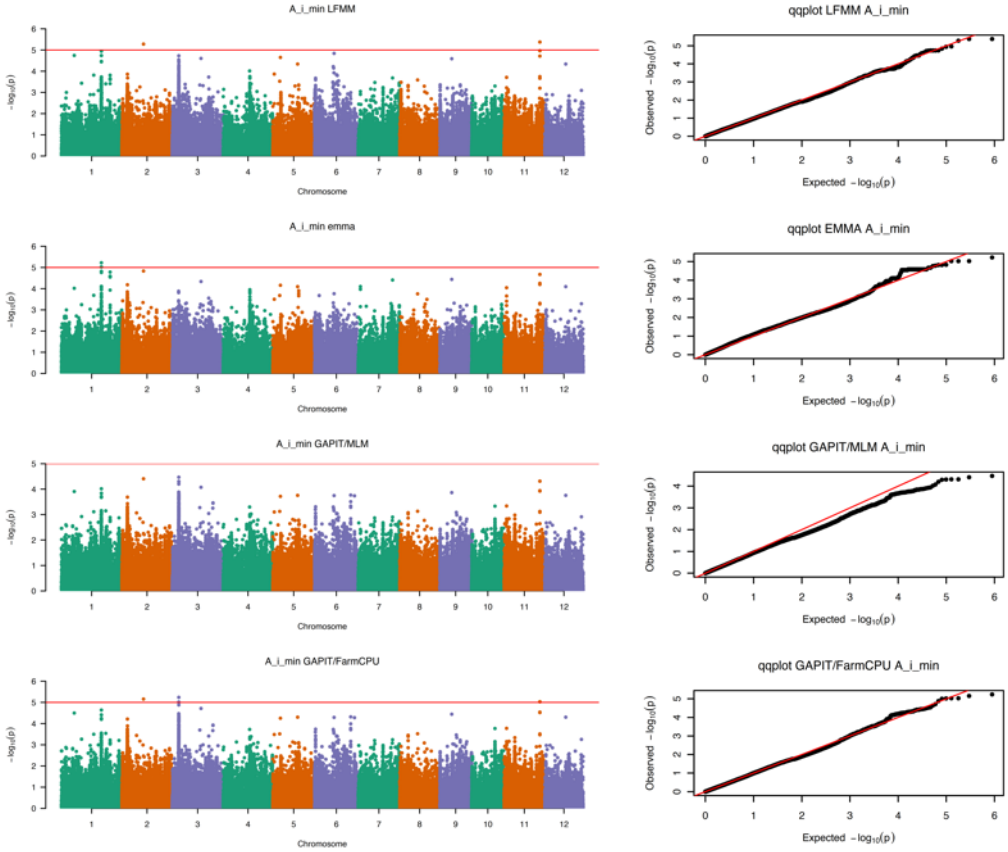
$A_{i\,50}$


A_i_50_BC LFMM
qqplot aov A_i_50_BC
A_i_50_BC emma
qqplot LFMM A_i_50_BC
A_i_50_BC GAPIT/MLM
qqplot EMMA A_i_50_BC
A_i_50_BC GAPIT/FarmCPU
qqplot GAPIT/MLM A_i_50_BC
Chromosome
Observed −log10(p)
Expected −log10(p)

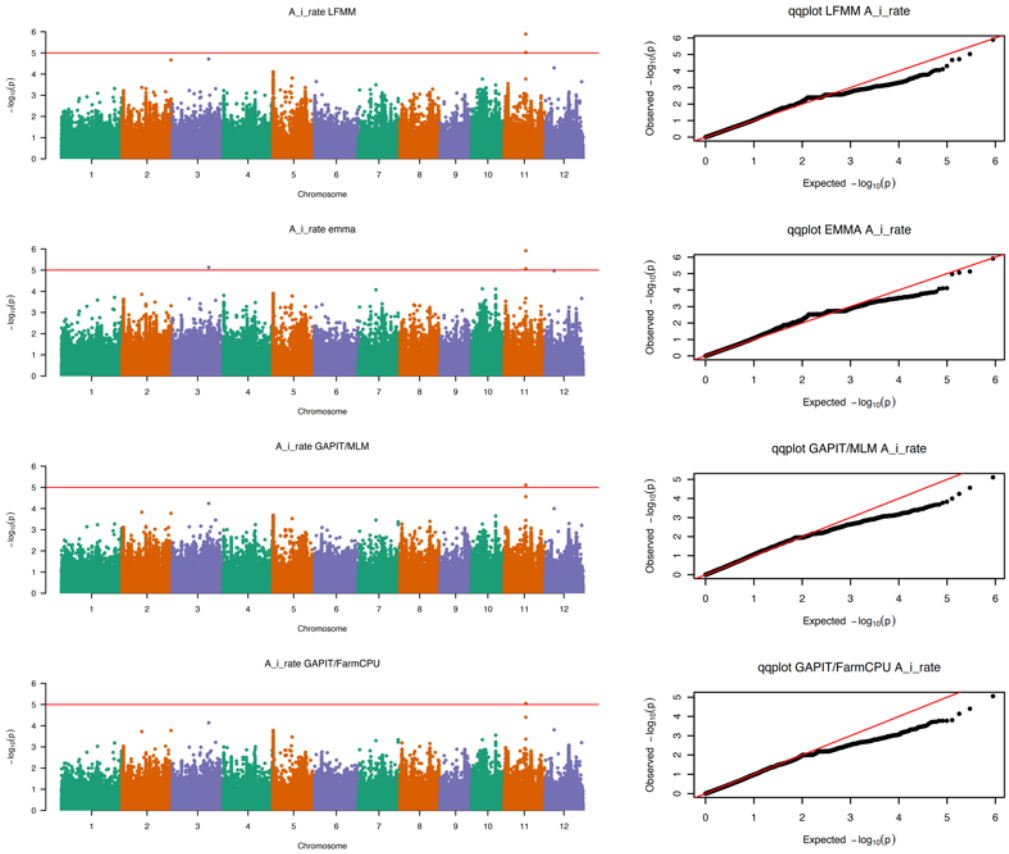
$A_{i\,90}$


A_i_90BC LFMM
qqplot LFMM A_i_90BC
A_i_90BC emma
qqplot EMMA A_i_90BC
A_i_90BC GAPIT/MLM
qqplot GAPIT/MLM A_i_90BC
A_i_90BC GAPIT/FarmCPU
qqplot GAPIT/FarmCPU A_i_90BC
Chromosome
Observed −log10(p)
Expected −log10(p)

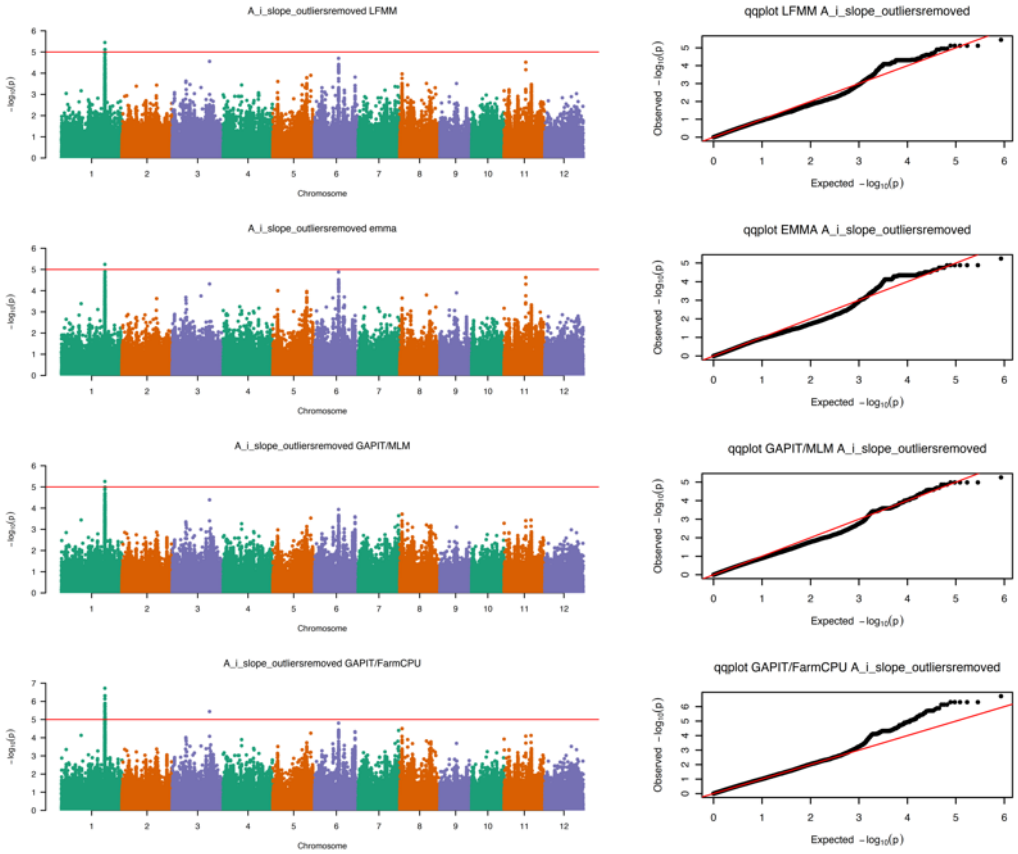
$A_{\text{i min}}$



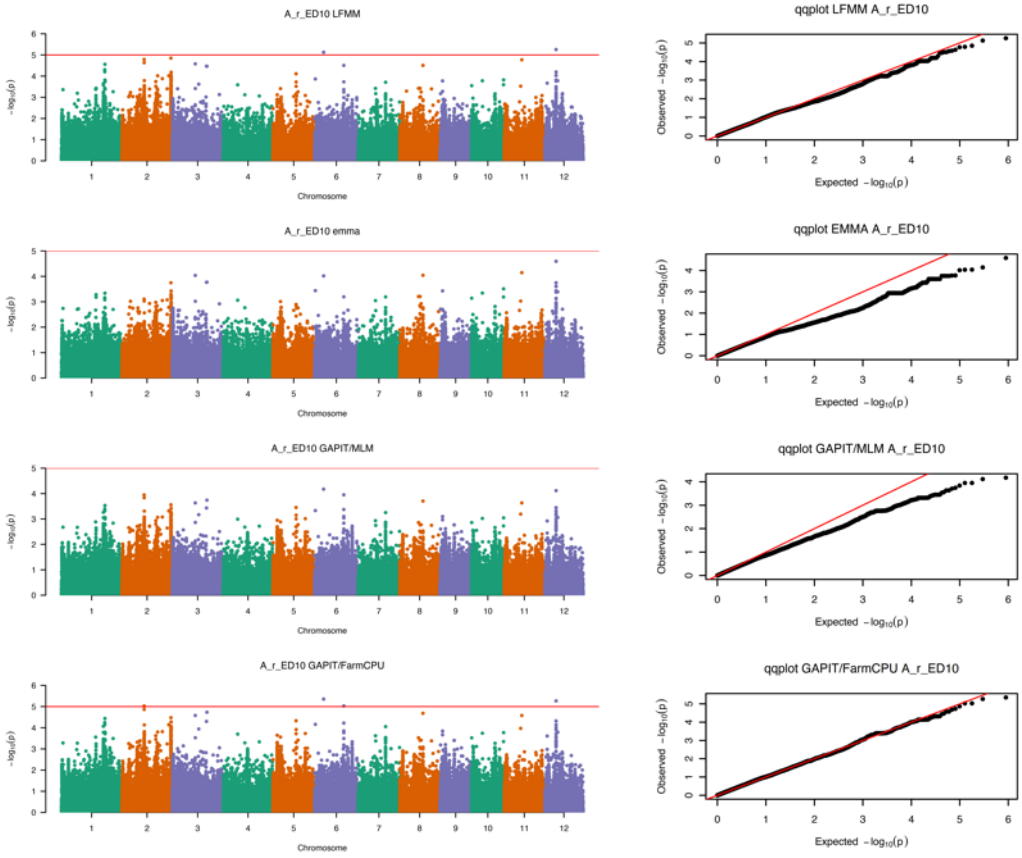
$A_{\text{i rate}}$



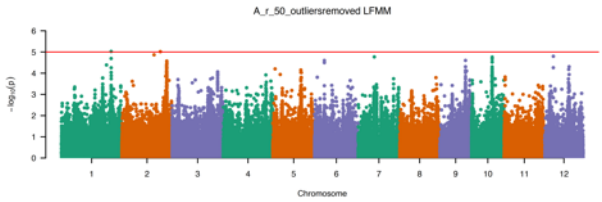
$A_{\text{i slope}}$

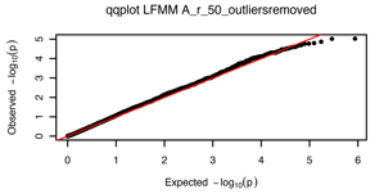


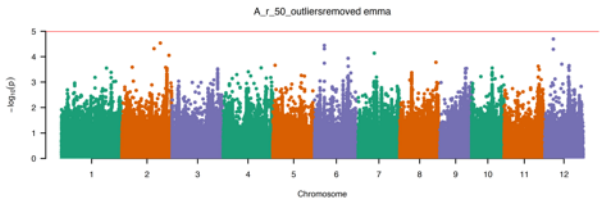
$A_{\text{r 10}}$

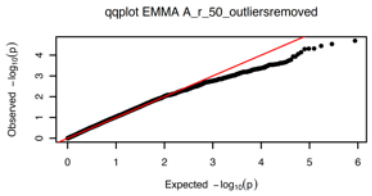


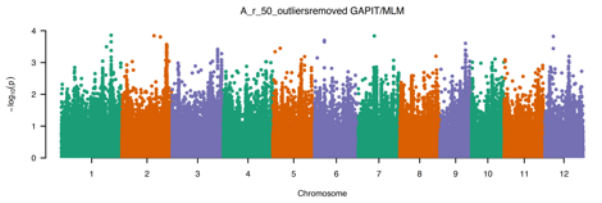
$A_{r\,50}$

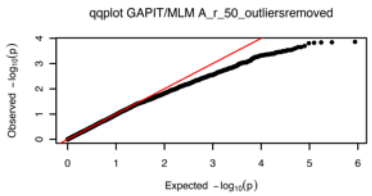

A_r_50_outliersremoved LFMM
Chromosome

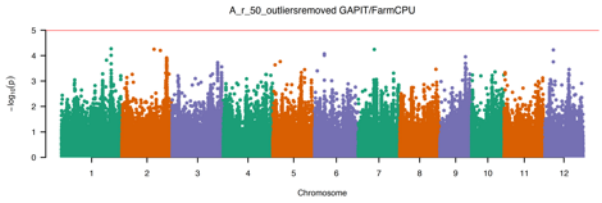

qqplot LFMM A_r_50_outliersremoved
Observed −log10(p)
Expected −log10(p)

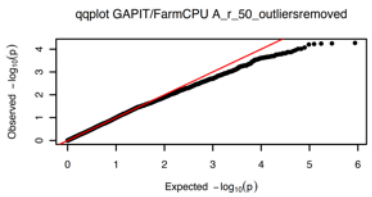

A_r_50_outliersremoved emma
Chromosome


qqplot EMMA A_r_50_outliersremoved
Observed −log10(p)
Expected −log10(p)

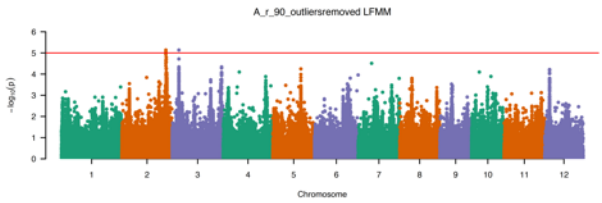

A_r_50_outliersremoved GAPIT/MLM
Chromosome

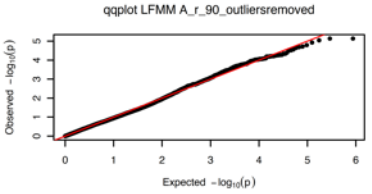

qqplot GAPIT/MLM A_r_50_outliersremoved
Observed −log10(p)
Expected −log10(p)

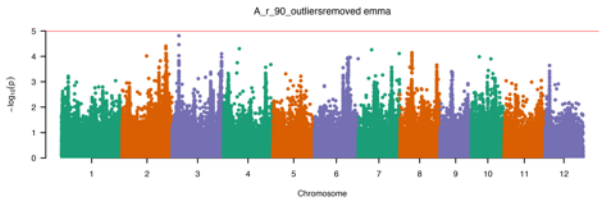

A_r_50_outliersremoved GAPIT/FarmCPU
Chromosome

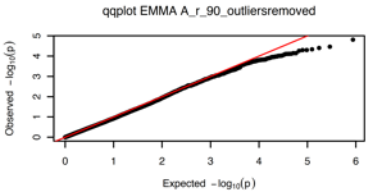

qqplot GAPIT/FarmCPU A_r_50_outliersremoved
Observed −log10(p)
Expected −log10(p)

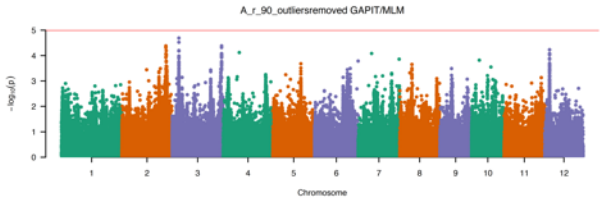
$A_{r\,90}$

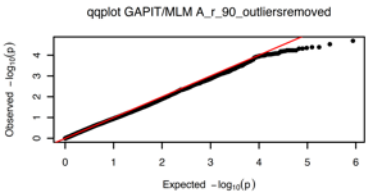

A_r_90_outliersremoved LFMM
Chromosome

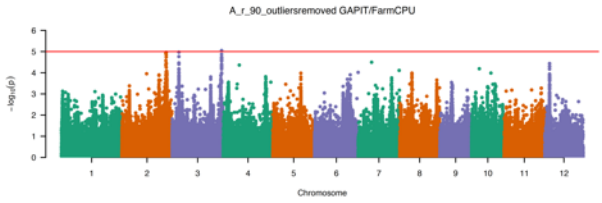

qqplot LFMM A_r_90_outliersremoved
Observed −log10(p)
Expected −log10(p)

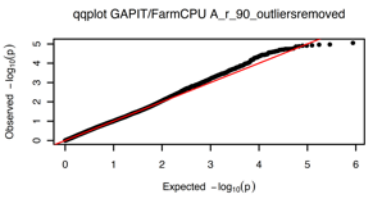

A_r_90_outliersremoved emma
Chromosome


qqplot EMMA A_r_90_outliersremoved
Observed −log10(p)
Expected −log10(p)

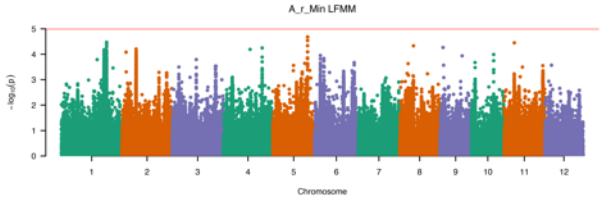

A_r_90_outliersremoved GAPIT/MLM
Chromosome

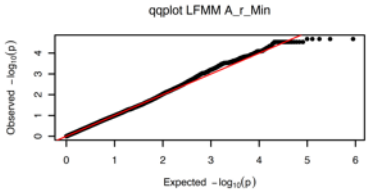

qqplot GAPIT/MLM A_r_90_outliersremoved
Observed −log10(p)
Expected −log10(p)

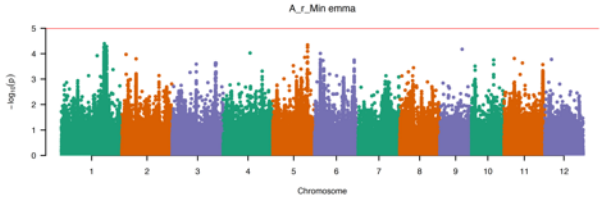

A_r_90_outliersremoved GAPIT/FarmCPU
Chromosome

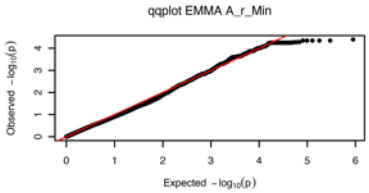

qqplot GAPIT/FarmCPU A_r_90_outliersremoved
Observed −log10(p)
Expected −log10(p)

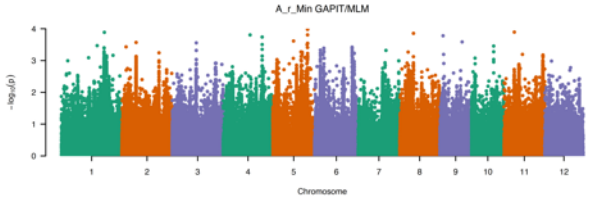
$A_{r\ min}$

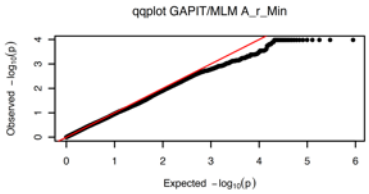


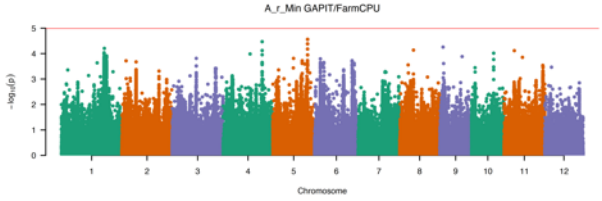


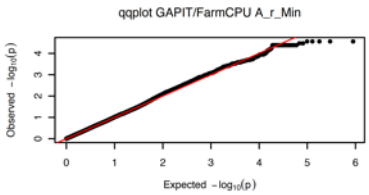




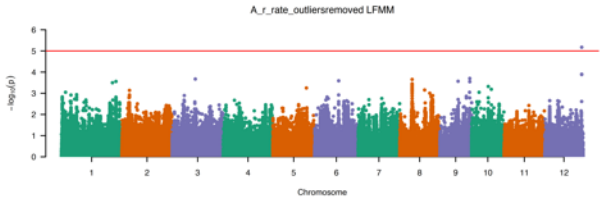


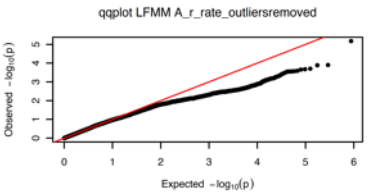


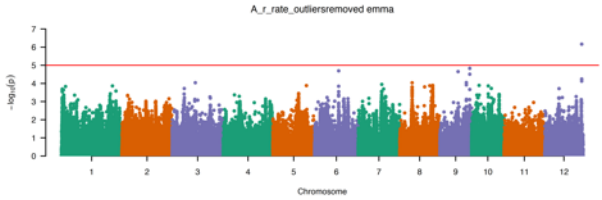


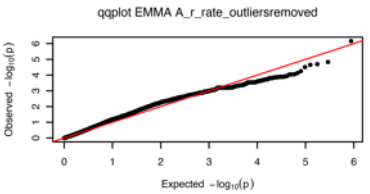


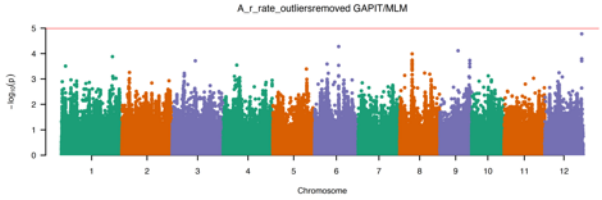
$A_{r\ rate}$

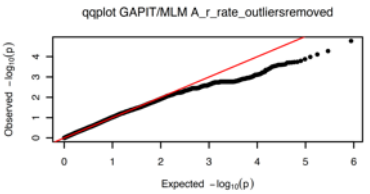


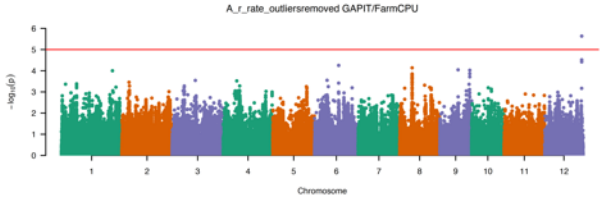


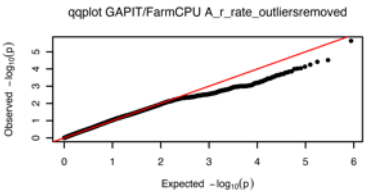




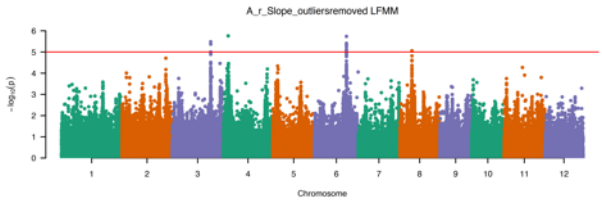


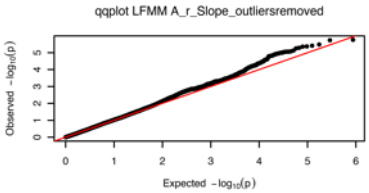


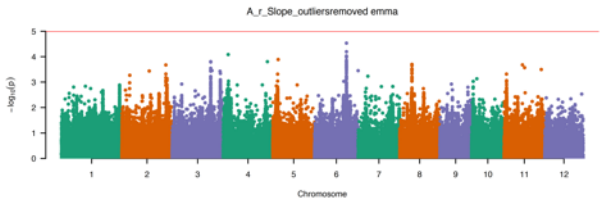


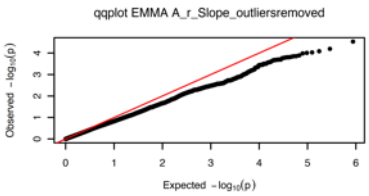


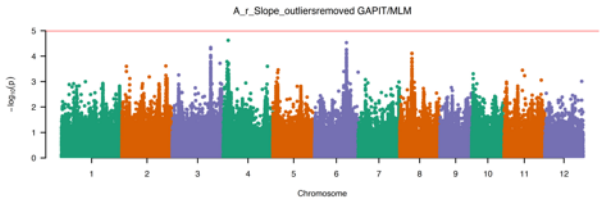
$A_{r\ slope}$

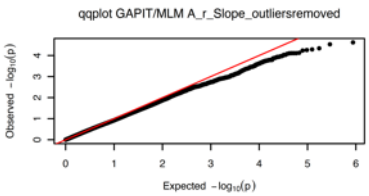


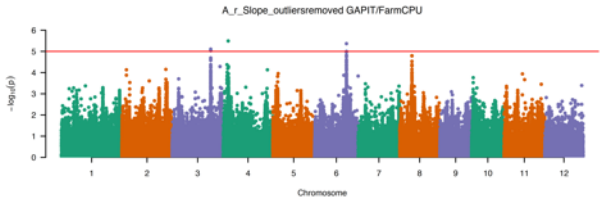


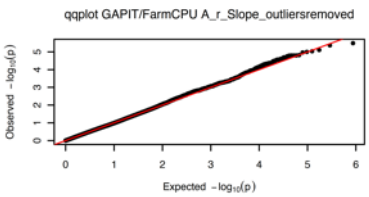




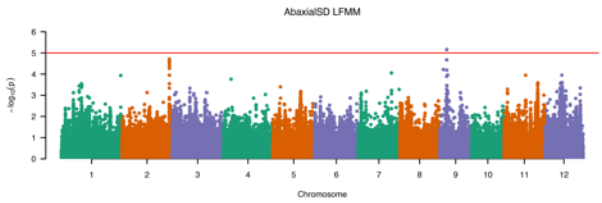


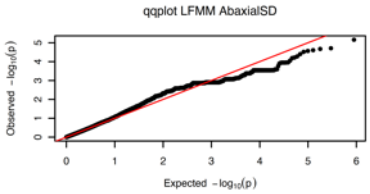


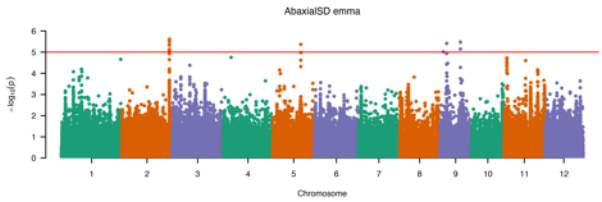


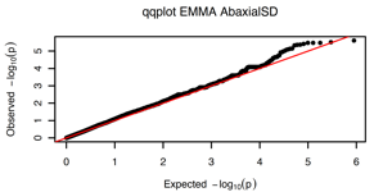


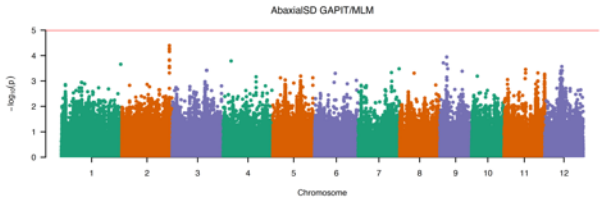
Abaxial stomatal density

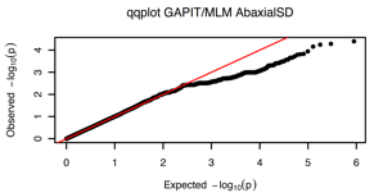


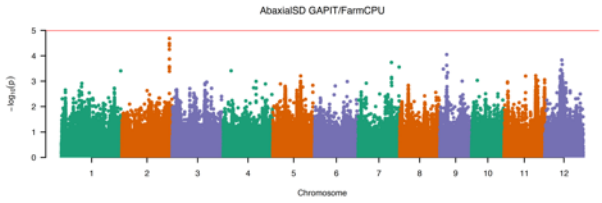


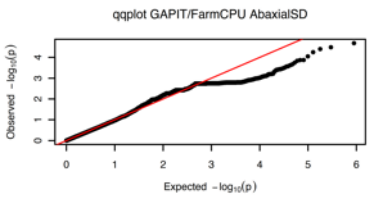




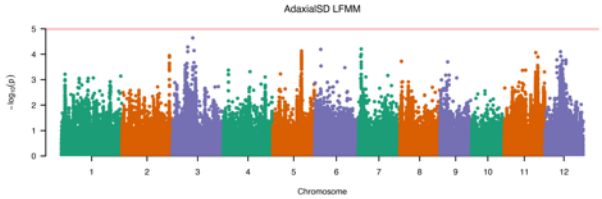


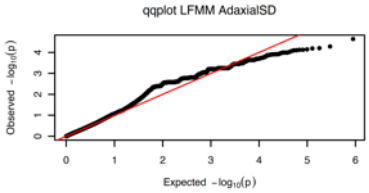


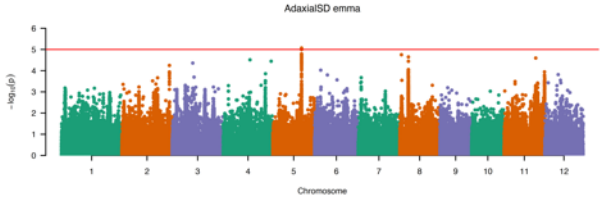


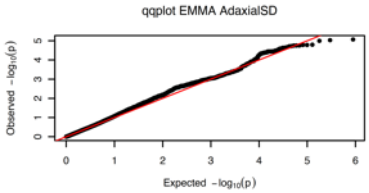


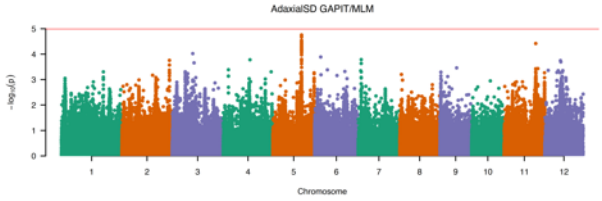
Adaxial stomatal density

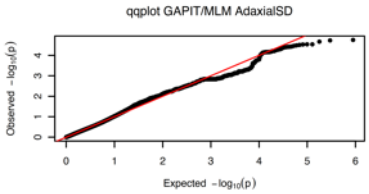


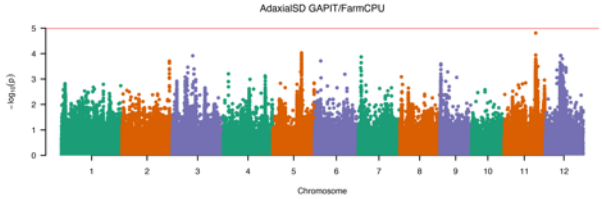


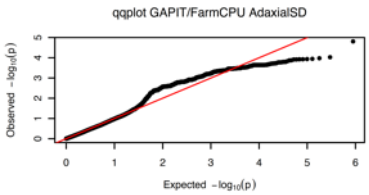




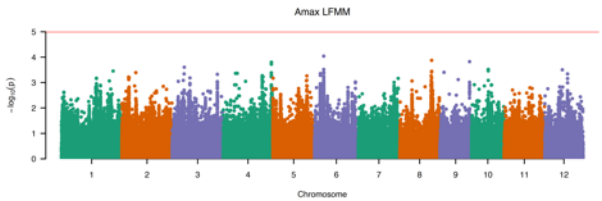


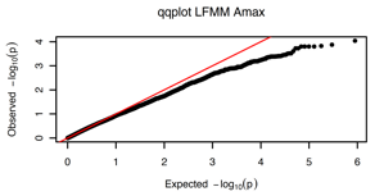


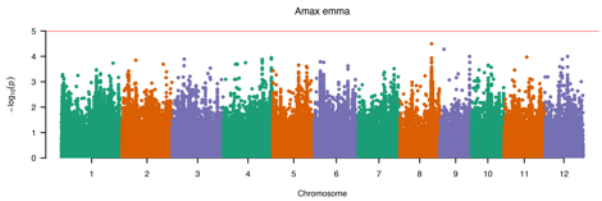


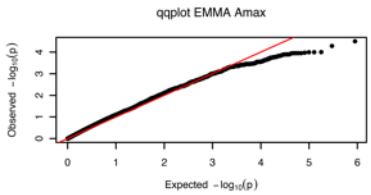


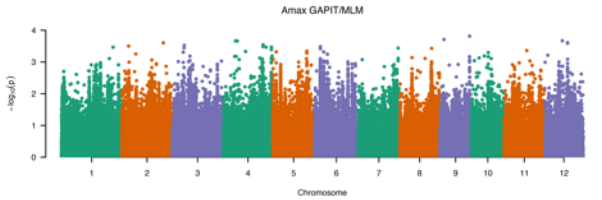
A_{max}

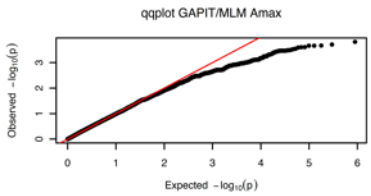


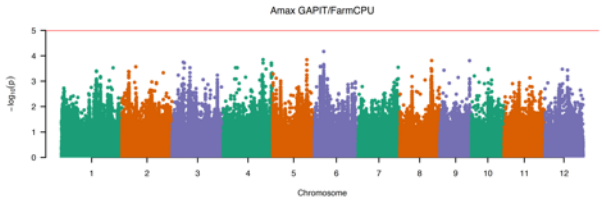


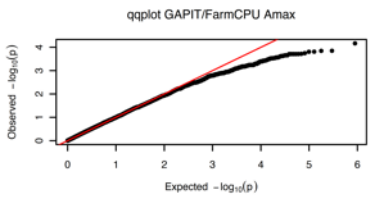




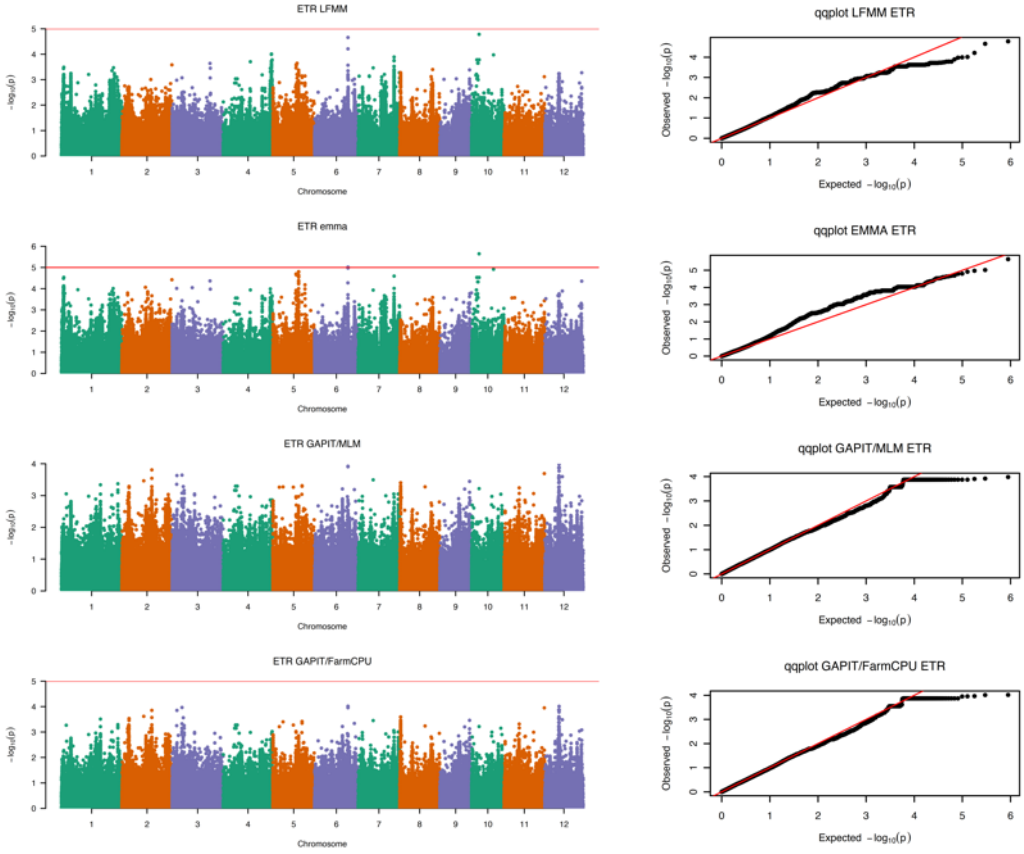




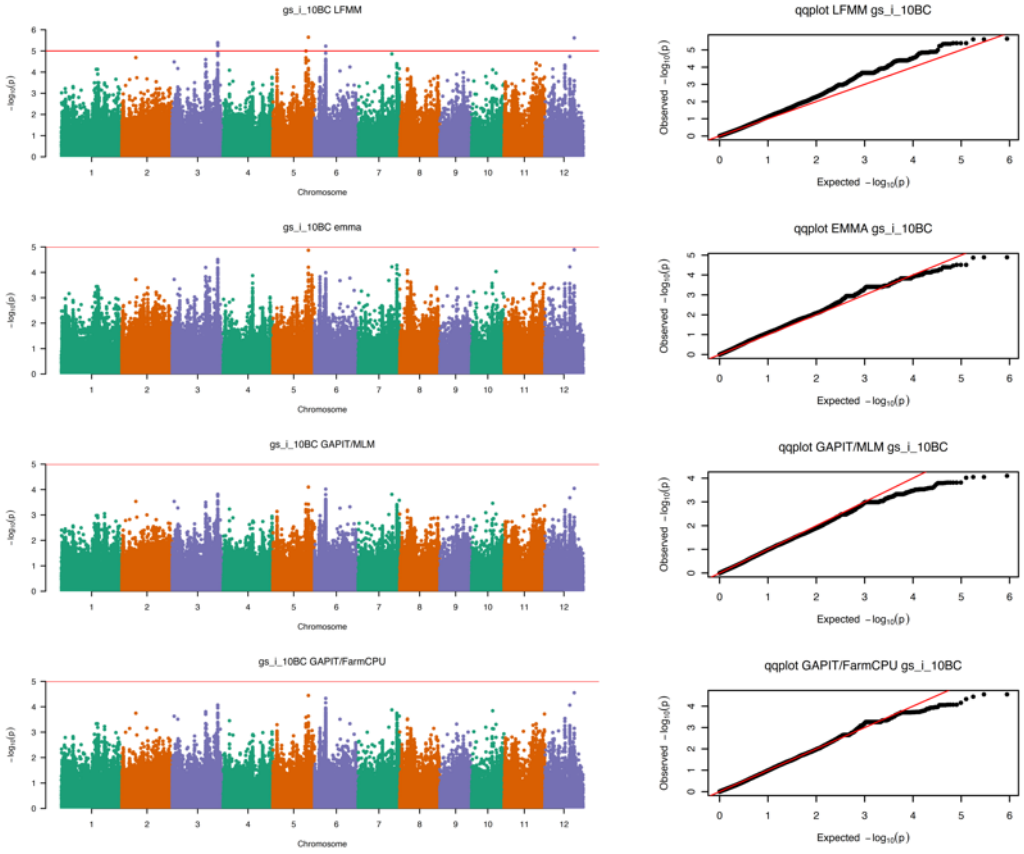




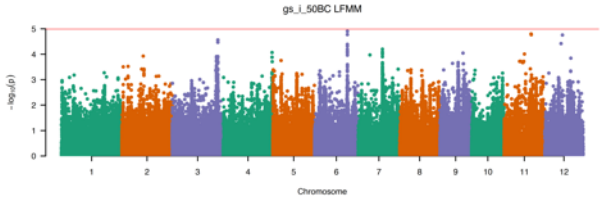
ETR*max*

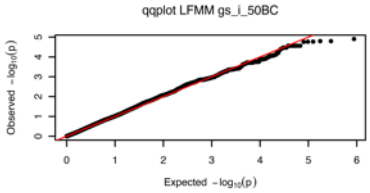


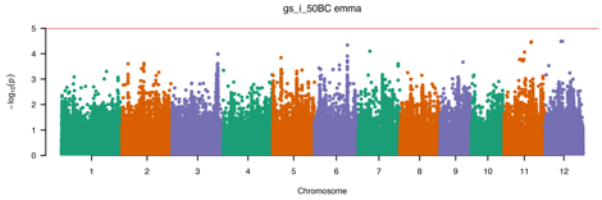
$gs_{i\ 10}$

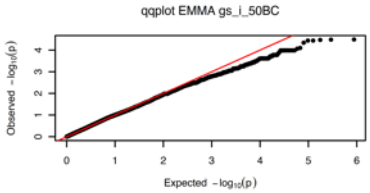


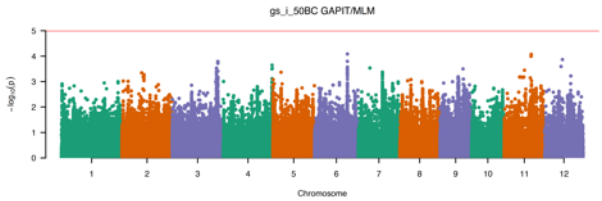
$gs_{i\ 50}$

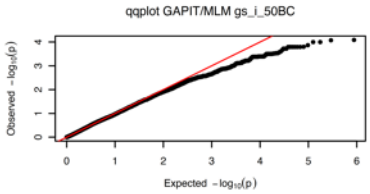

gs_i_50BC LFMM
Chromosome

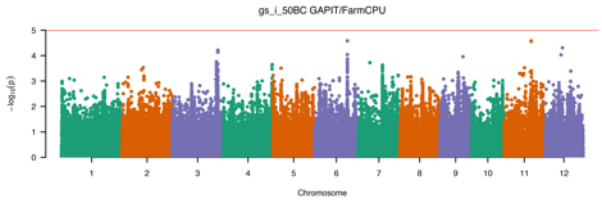

qqplot LFMM gs_i_50BC
Observed −log10(p)
Expected −log10(p)

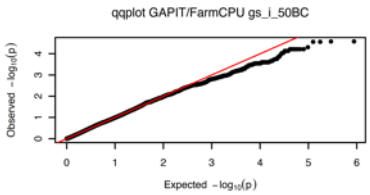

gs_i_50BC emma
Chromosome


qqplot EMMA gs_i_50BC
Observed −log10(p)
Expected −log10(p)

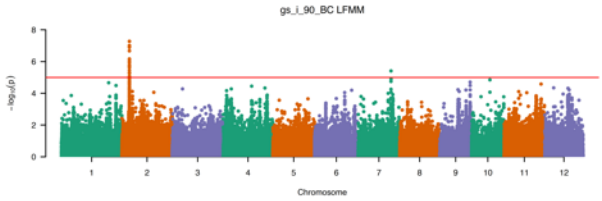

gs_i_50BC GAPIT/MLM
Chromosome

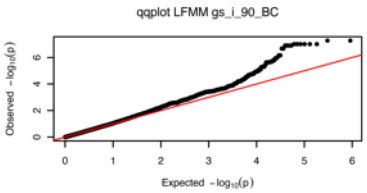

qqplot GAPIT/MLM gs_i_50BC
Observed −log10(p)
Expected −log10(p)

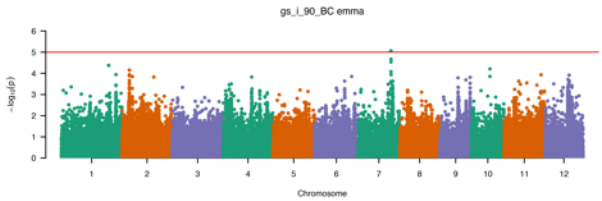

gs_i_50BC GAPIT/FarmCPU
Chromosome

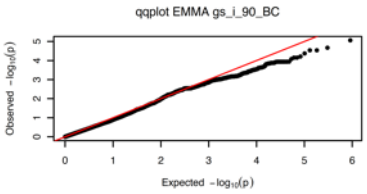

qqplot GAPIT/FarmCPU gs_i_50BC
Observed −log10(p)
Expected −log10(p)

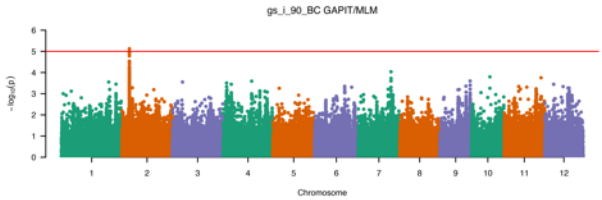
$gs_{i\ 90}$

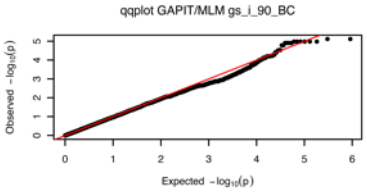

gs_i_90_BC LFMM
Chromosome

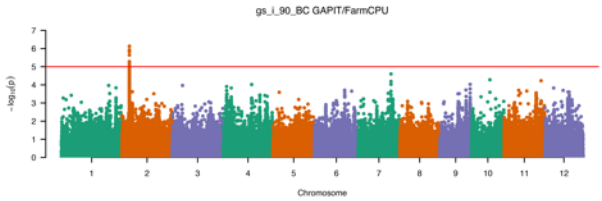

qqplot LFMM gs_i_90_BC
Observed −log10(p)
Expected −log10(p)

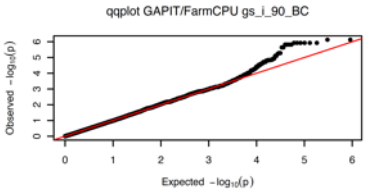

gs_i_90_BC emma
Chromosome


qqplot EMMA gs_i_90_BC
Observed −log10(p)
Expected −log10(p)

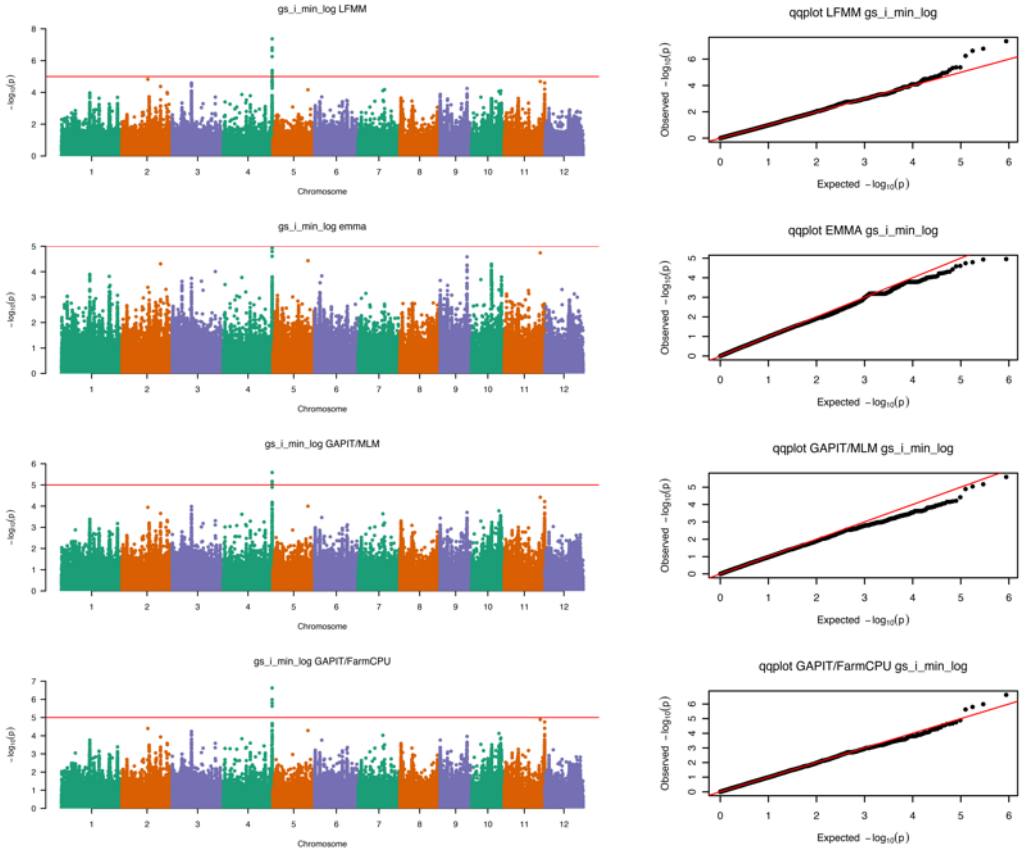

gs_i_90_BC GAPIT/MLM
Chromosome


qqplot GAPIT/MLM gs_i_90_BC
Observed −log10(p)
Expected −log10(p)

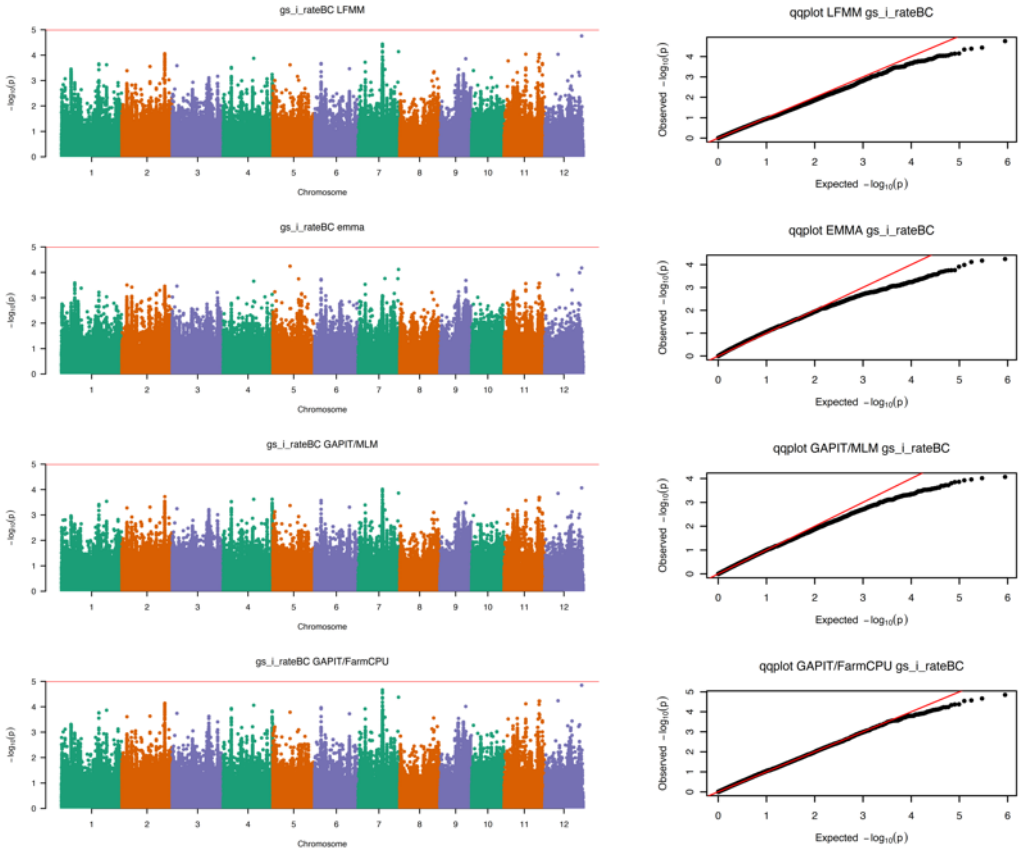

gs_i_90_BC GAPIT/FarmCPU
Chromosome


qqplot GAPIT/FarmCPU gs_i_90_BC
Observed −log10(p)
Expected −log10(p)

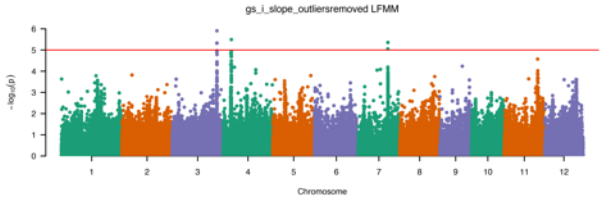
$gs_{\text{i min}}$

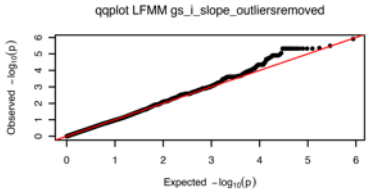


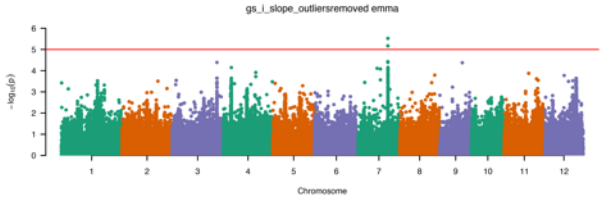
$gs_{\text{i rate}}$

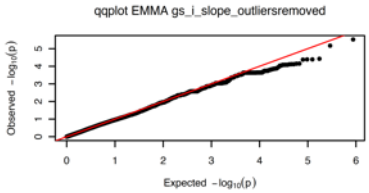


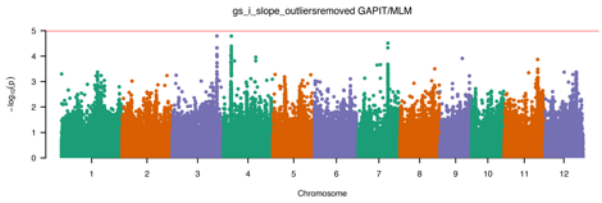
$gs_{\text{i slope}}$

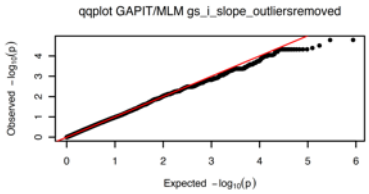


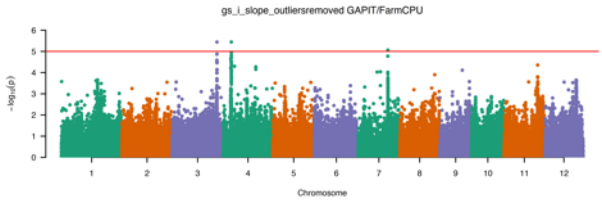


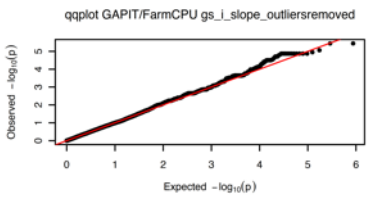




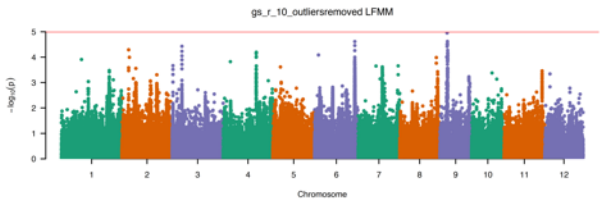


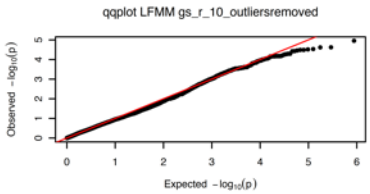


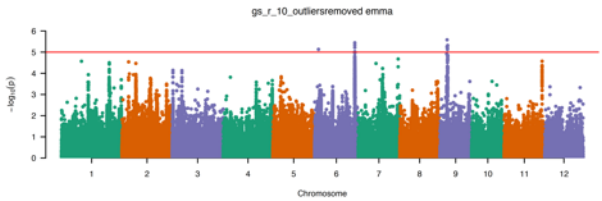


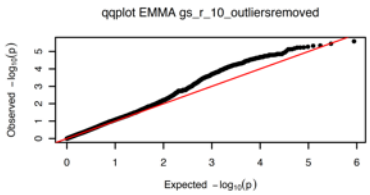


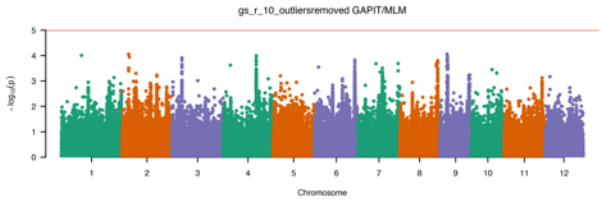
$gs_{\text{r 10}}$

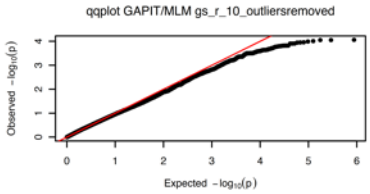


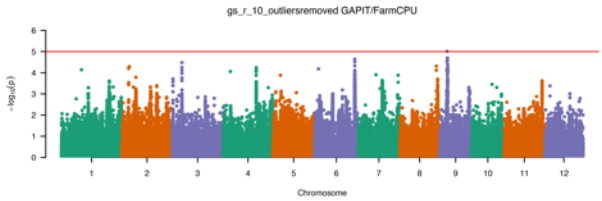


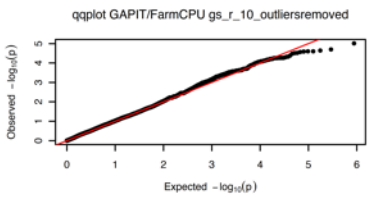




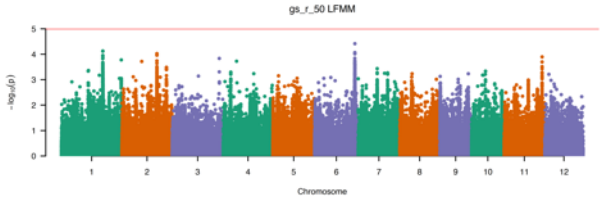


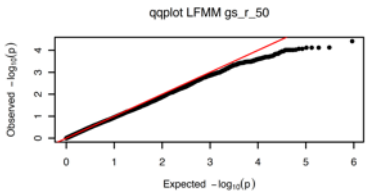


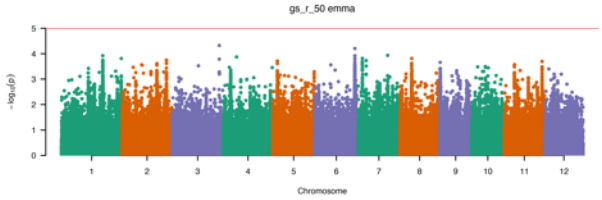


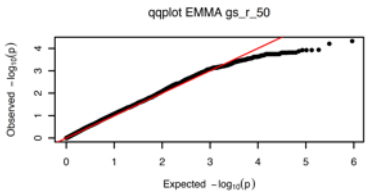


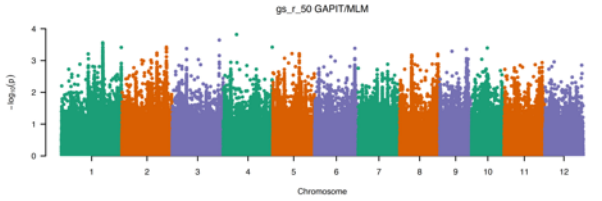
$gs_{r\ 50}$

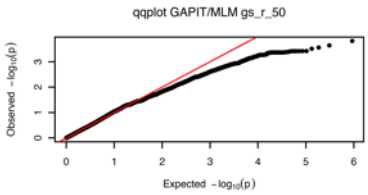

gs_r_50 LFMM
Chromosome

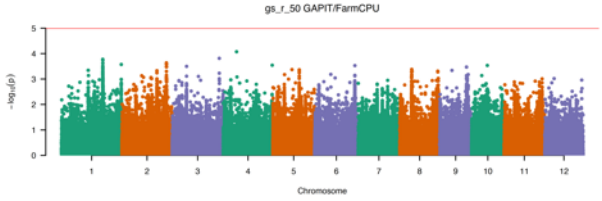

qqplot LFMM gs_r_50
Observed −log10(p)
Expected −log10(p)

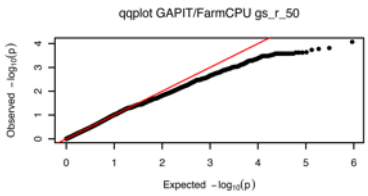

gs_r_50 emma
Chromosome


qqplot EMMA gs_r_50
Observed −log10(p)
Expected −log10(p)

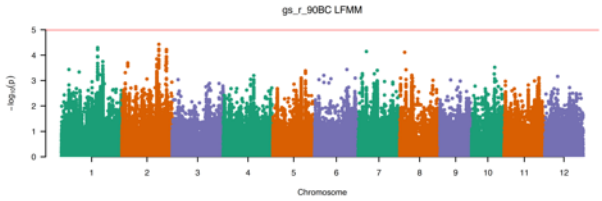

gs_r_50 GAPIT/MLM
Chromosome

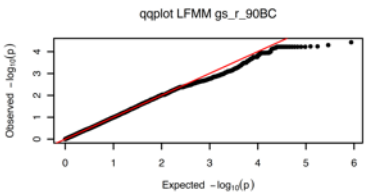

qqplot GAPIT/MLM gs_r_50
Observed −log10(p)
Expected −log10(p)

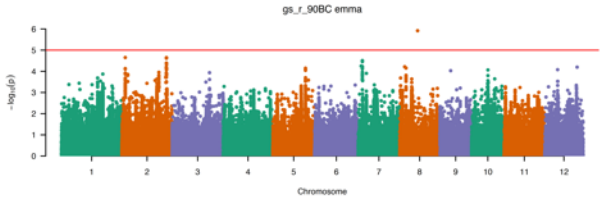

gs_r_50 GAPIT/FarmCPU
Chromosome

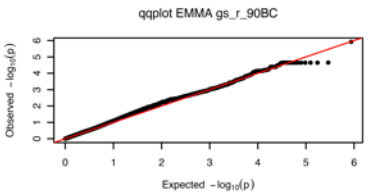

qqplot GAPIT/FarmCPU gs_r_50
Observed −log10(p)
Expected −log10(p)

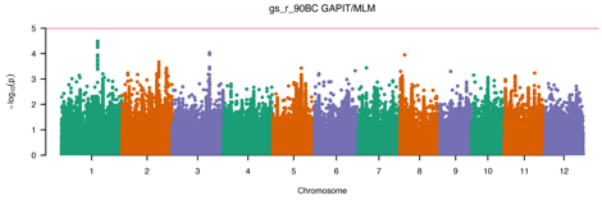
$gs_{r\ 90}$

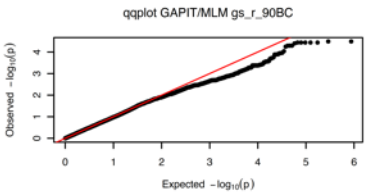

gs_r_90BC LFMM
Chromosome

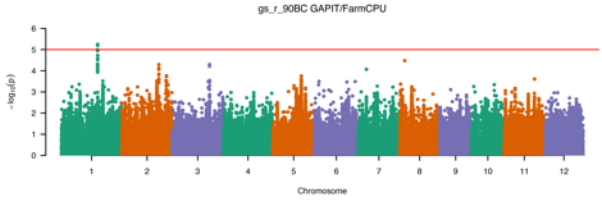

qqplot LFMM gs_r_90BC
Observed −log10(p)
Expected −log10(p)

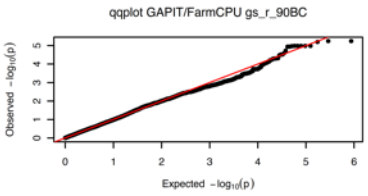

gs_r_90BC emma
Chromosome


qqplot EMMA gs_r_90BC
Observed −log10(p)
Expected −log10(p)

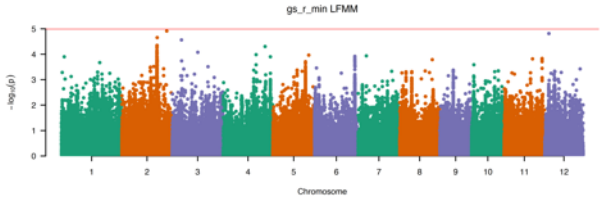

gs_r_90BC GAPIT/MLM
Chromosome

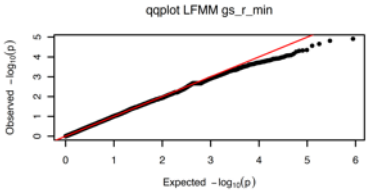

qqplot GAPIT/MLM gs_r_90BC
Observed −log10(p)
Expected −log10(p)

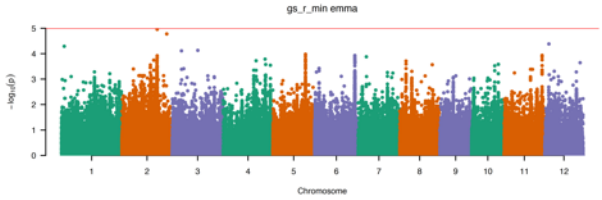

gs_r_90BC GAPIT/FarmCPU
Chromosome

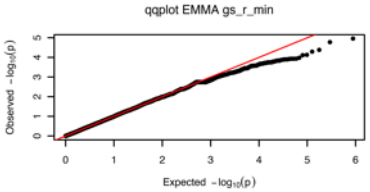

qqplot GAPIT/FarmCPU gs_r_90BC
Observed −log10(p)
Expected −log10(p)

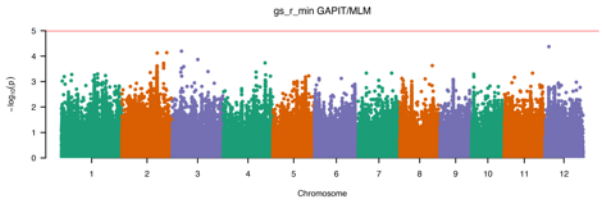
$gs_{r\ min}$

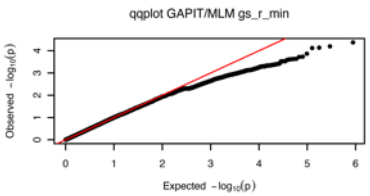


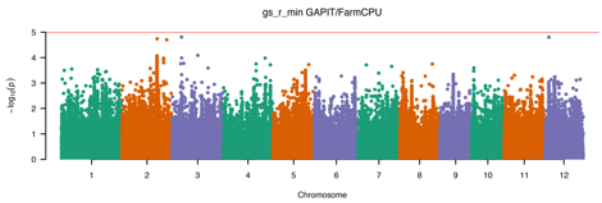


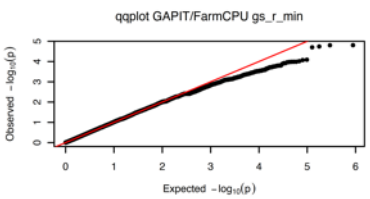




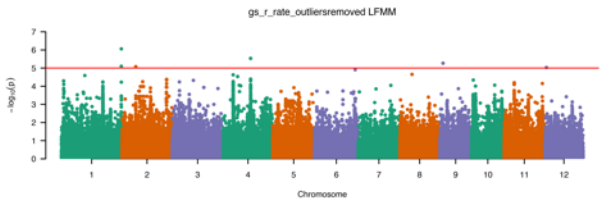


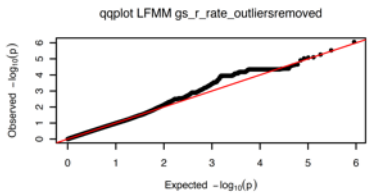


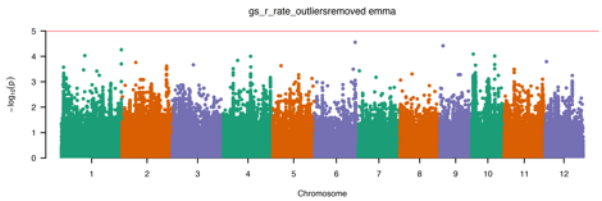


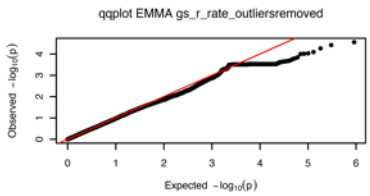


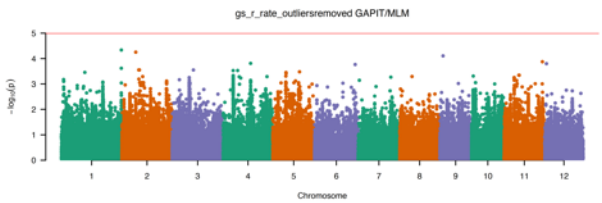
$gs_{r\ rate}$

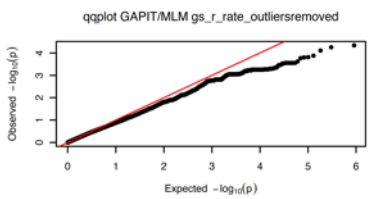


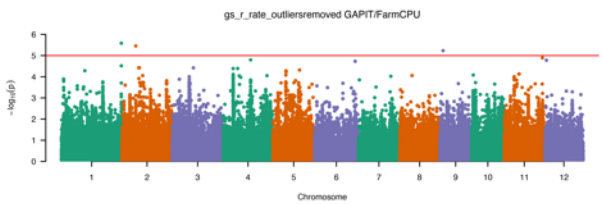


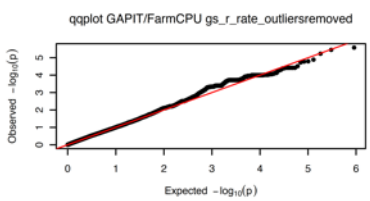




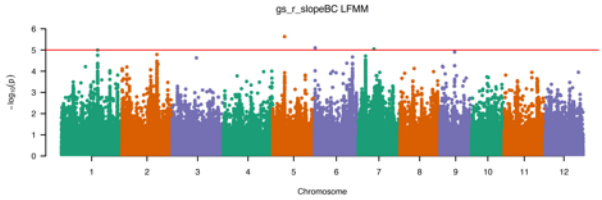


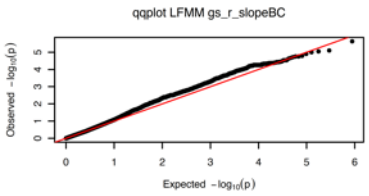


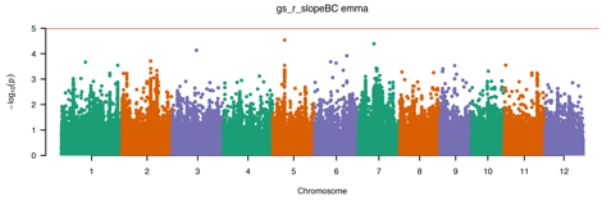


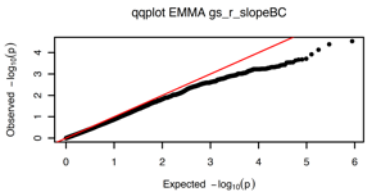


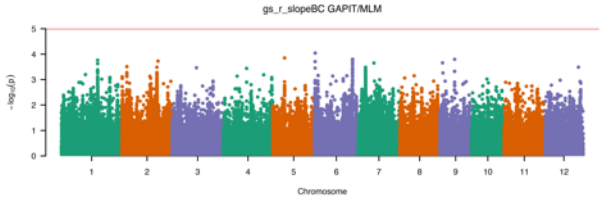
$gs_{\text{r slope}}$

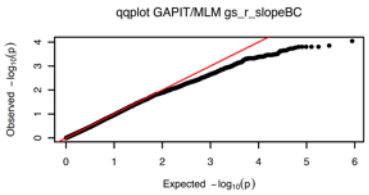


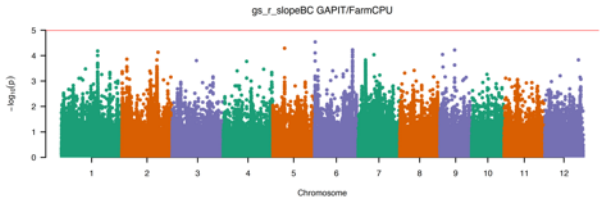


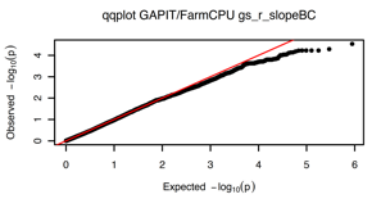




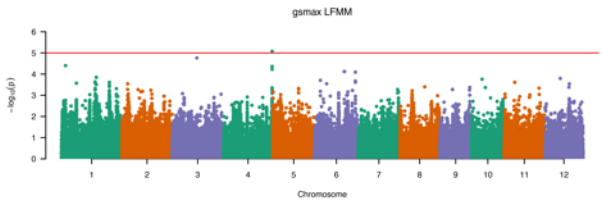


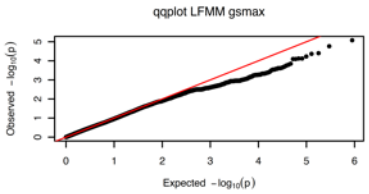


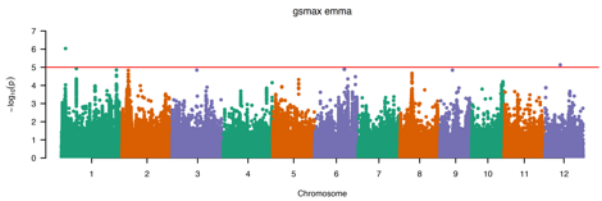


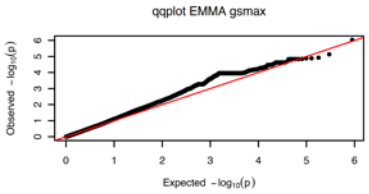


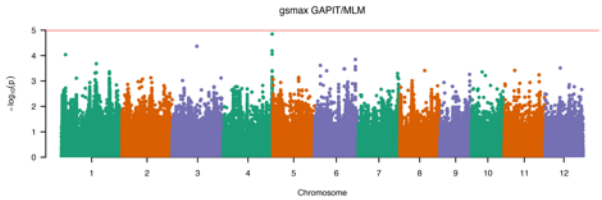
gs_{max}

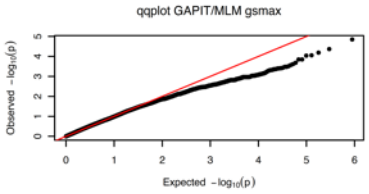


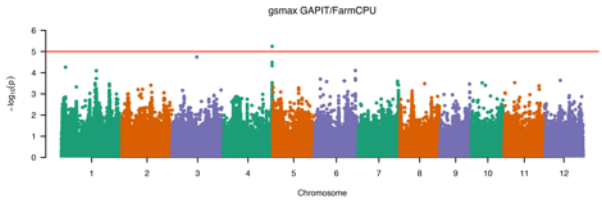


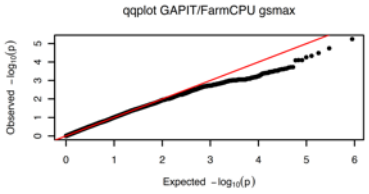




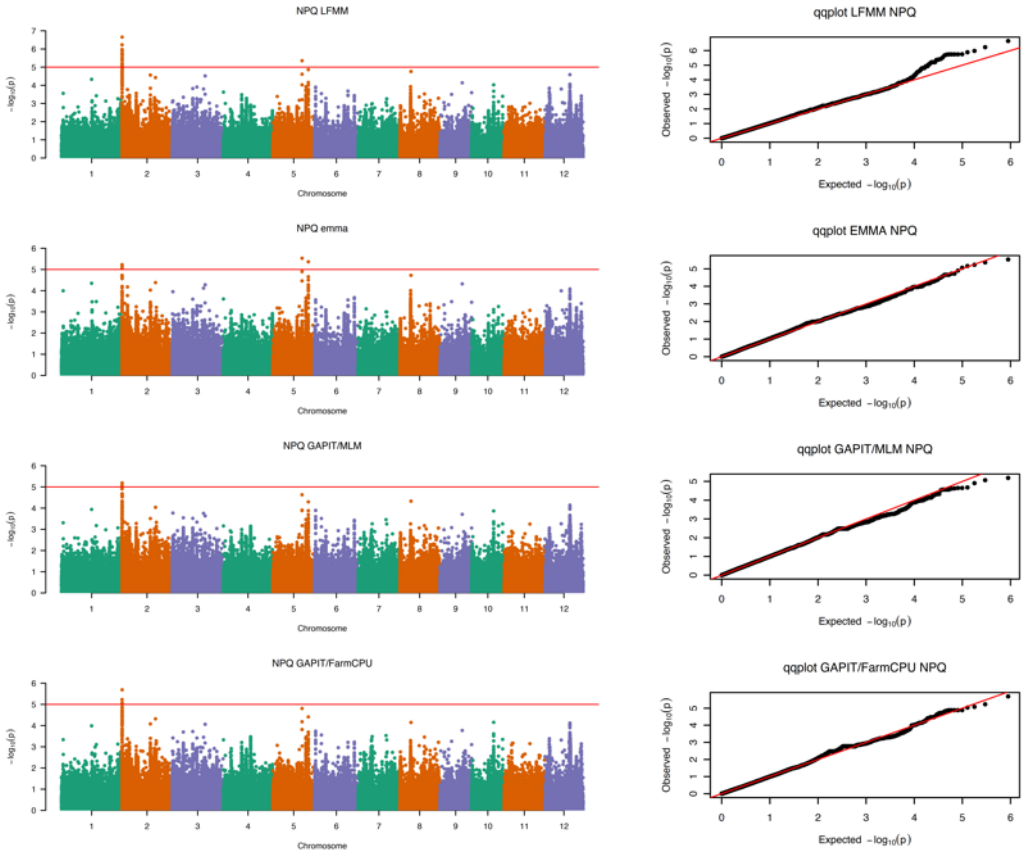




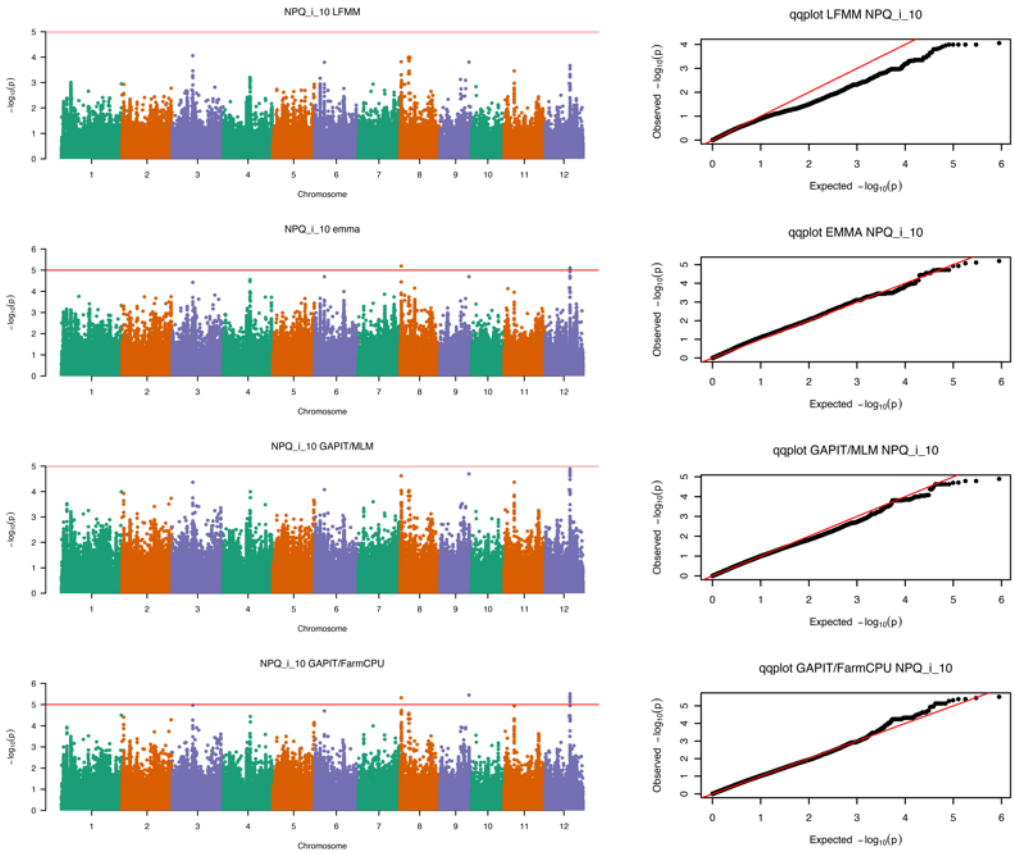




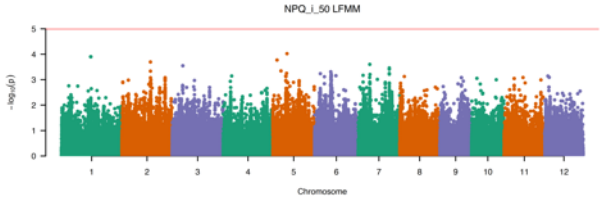
NPQ*max*

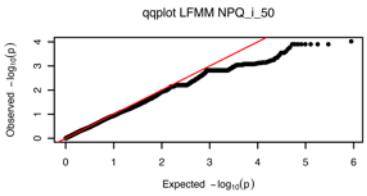


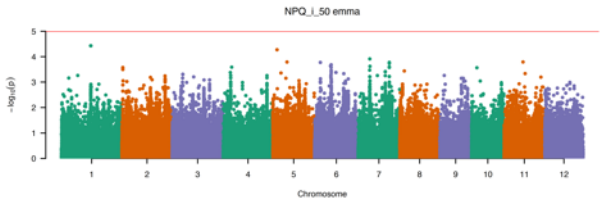
NPQi 10

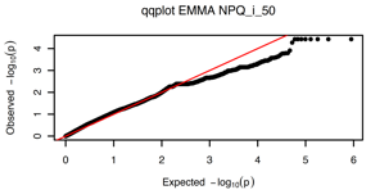


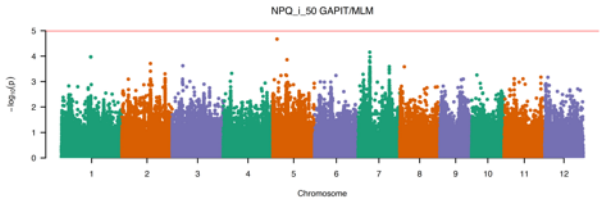
$NPQ_{i\ 50}$

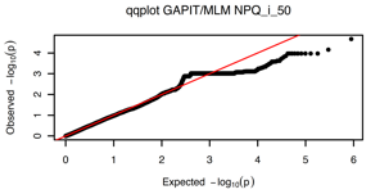


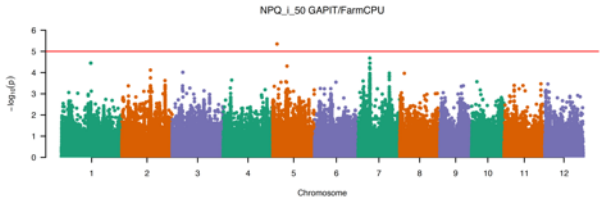


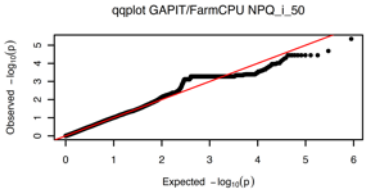




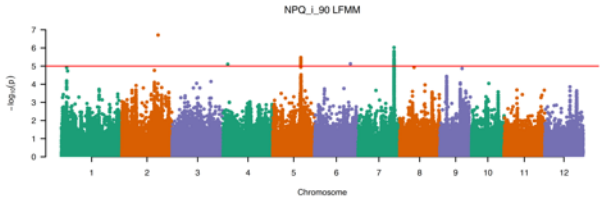


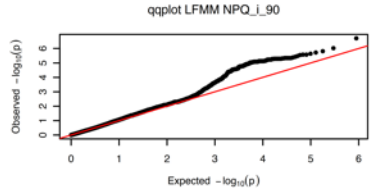


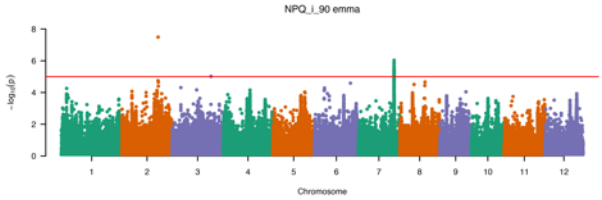


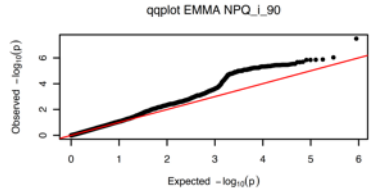


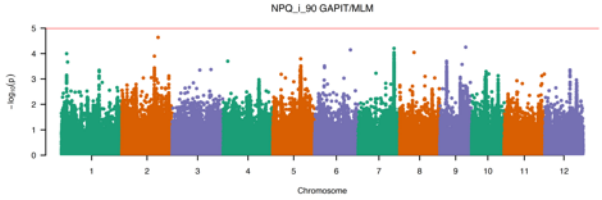
$NPQ_{i\ 90}$

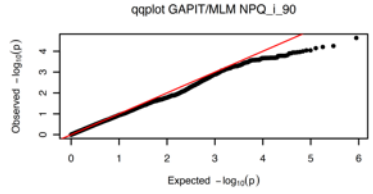


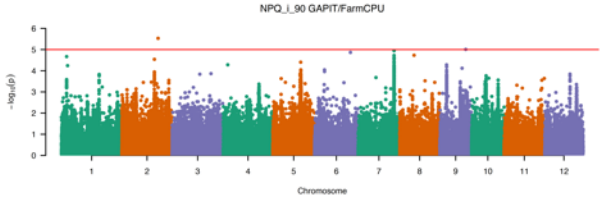


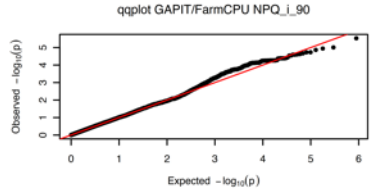




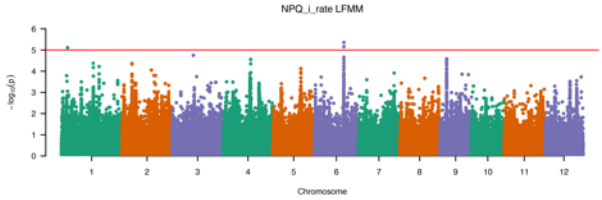


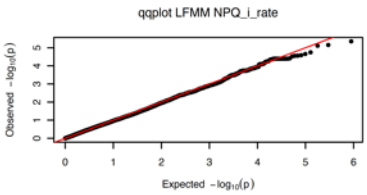


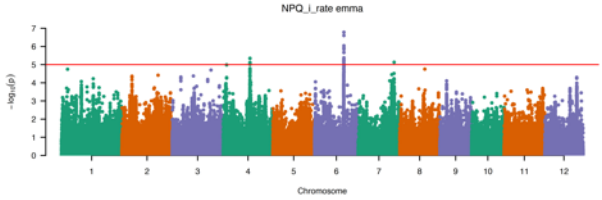


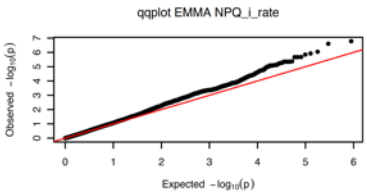


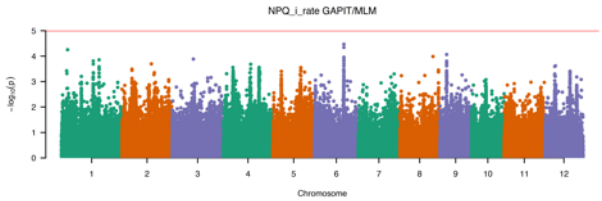
$NPQ_{i\ rate}$

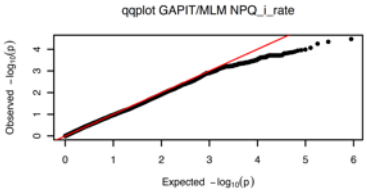


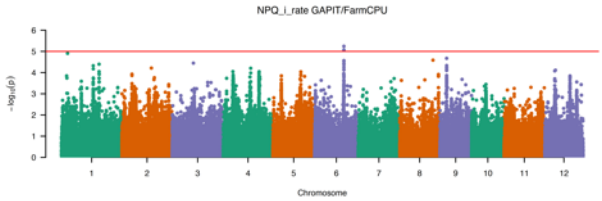


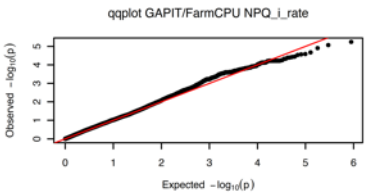




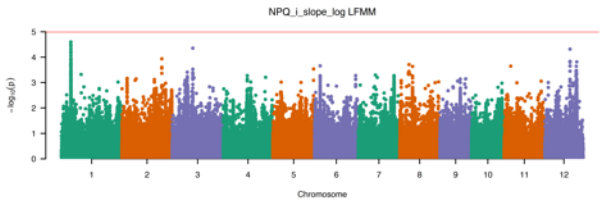


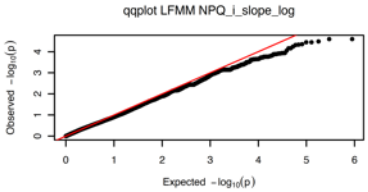


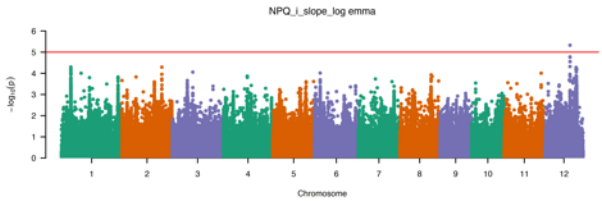


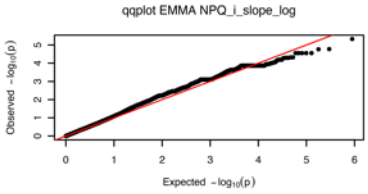


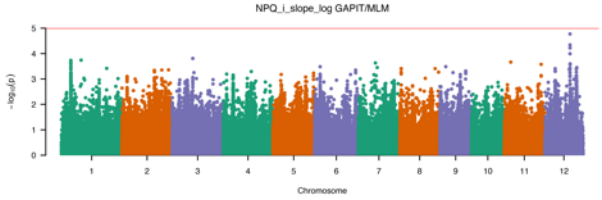
$NPQ_{i\ slope}$

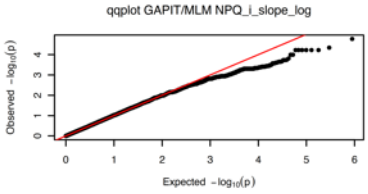


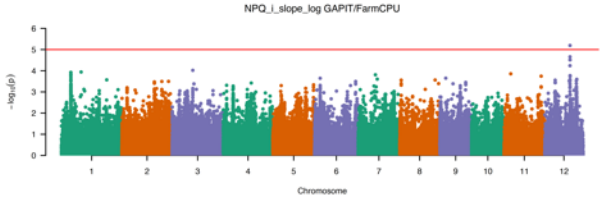


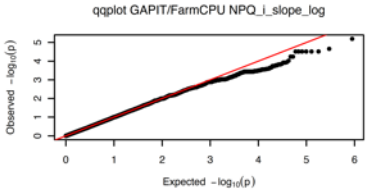




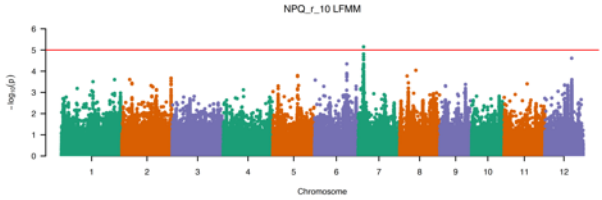


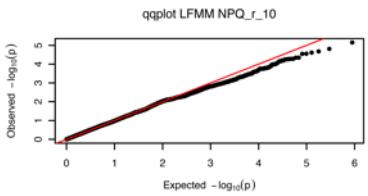


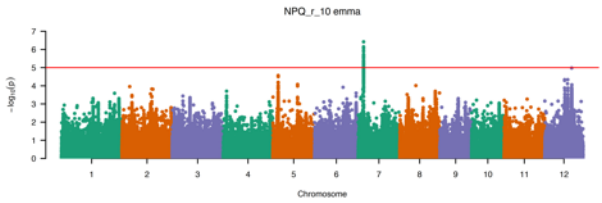


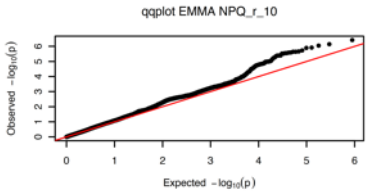


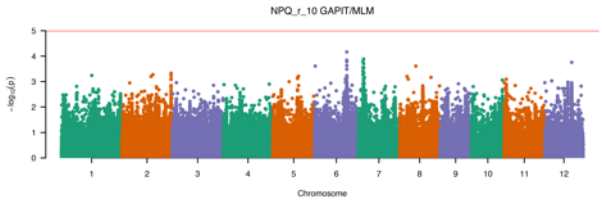
$NPQ_{r\ 10}$

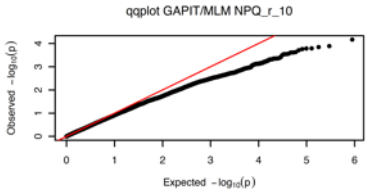


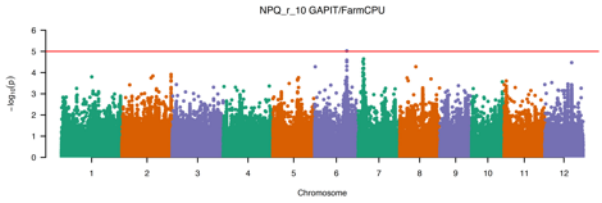


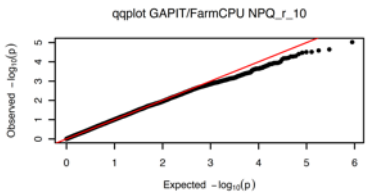




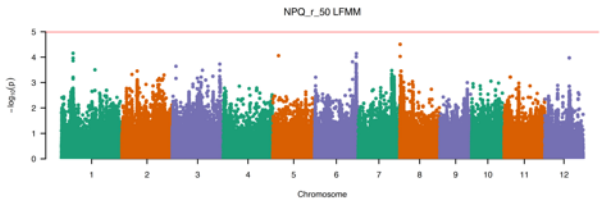


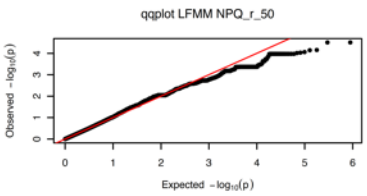


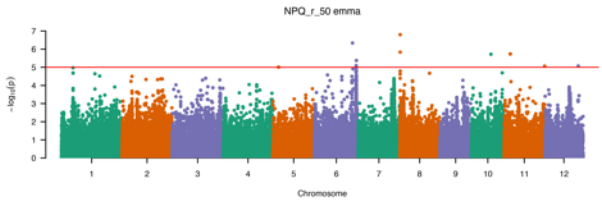


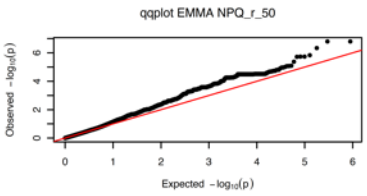


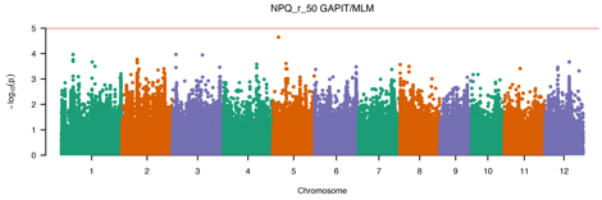
$NPQ_{r\ 50}$

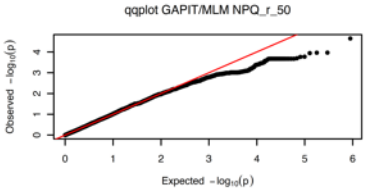


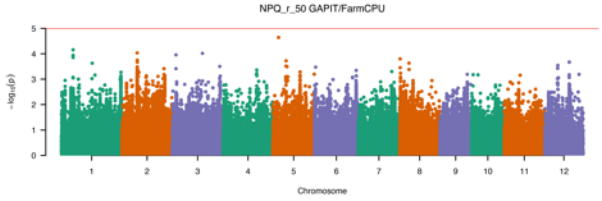


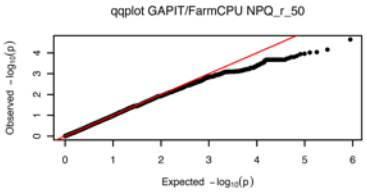




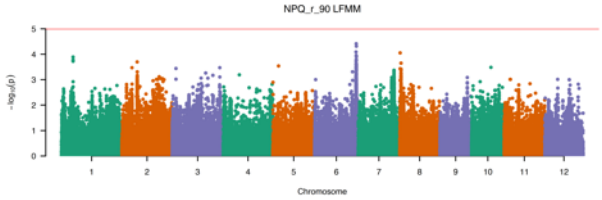


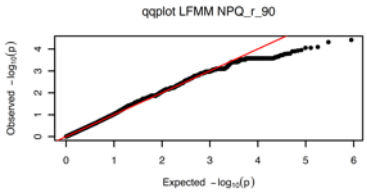


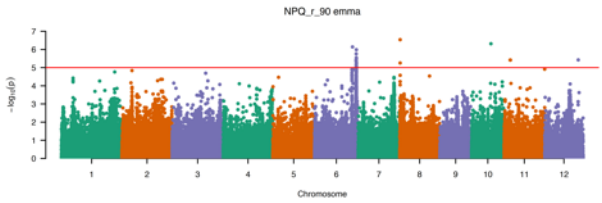


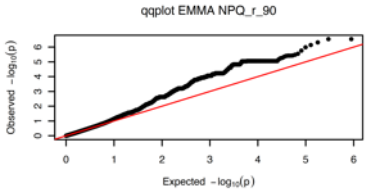


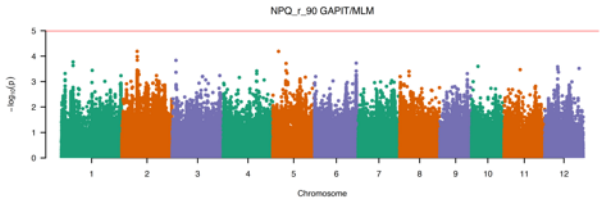
$NPQ_{r\ 90}$

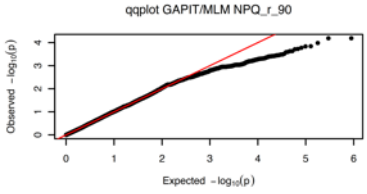


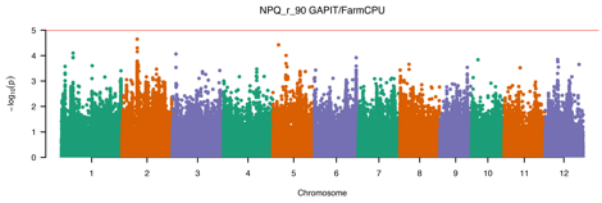


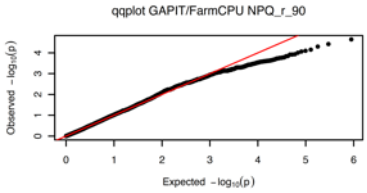




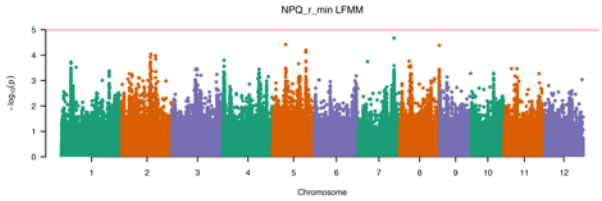


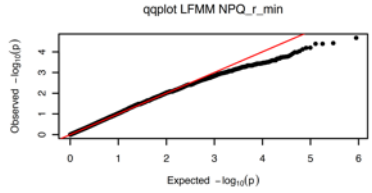


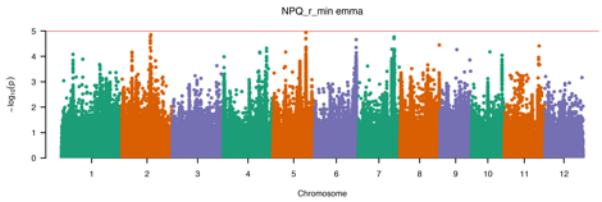


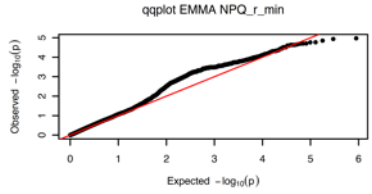


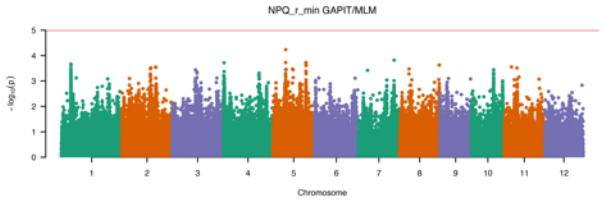
$NPQ_{r\ min}$

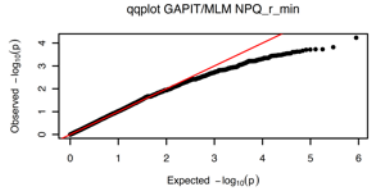


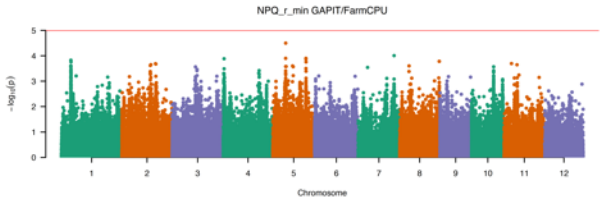


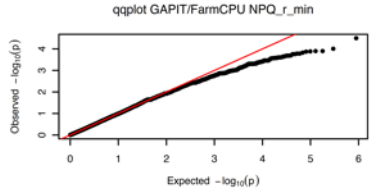




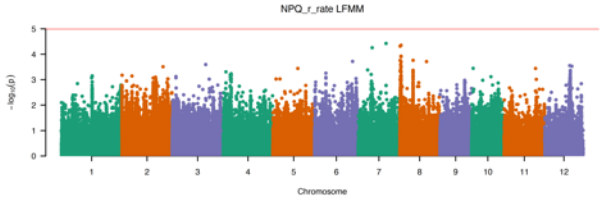


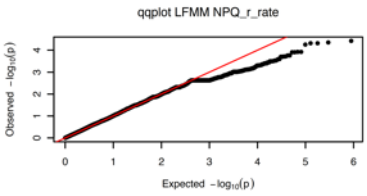


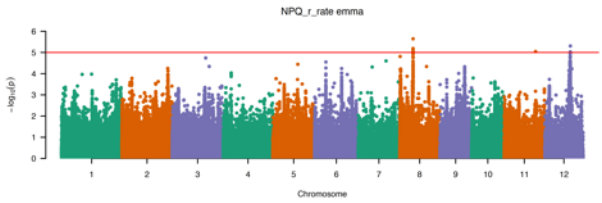


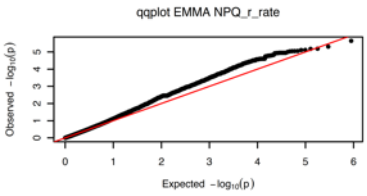


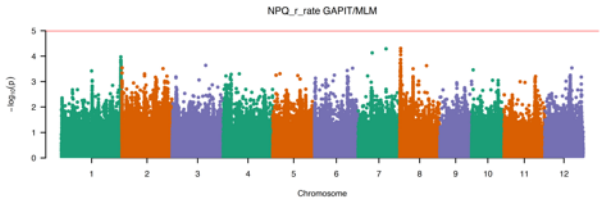
$NPQ_{r\ rate}$

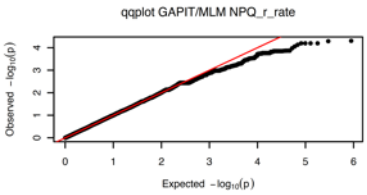


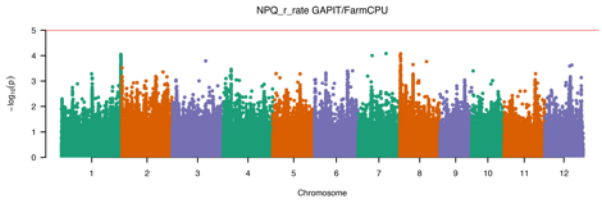


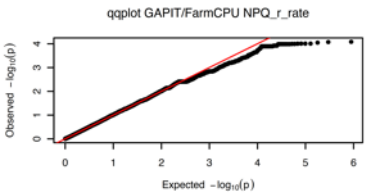




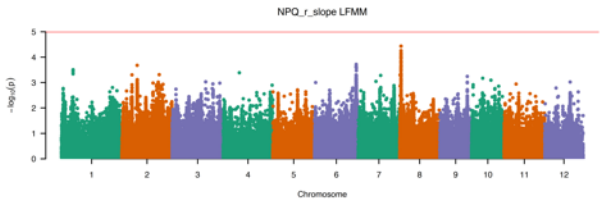


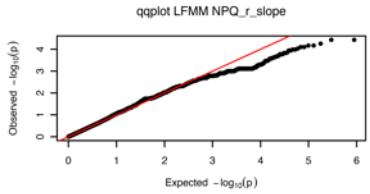


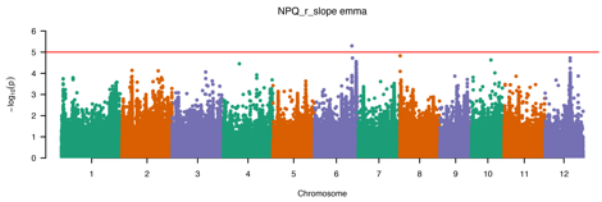


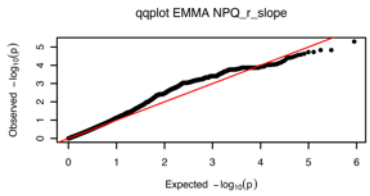


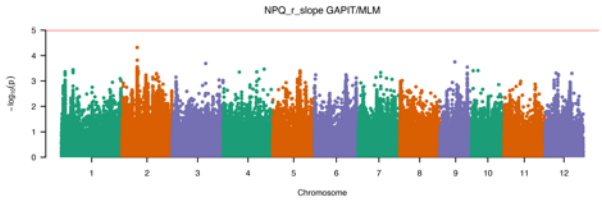
$NPQ_{r\ slope}$

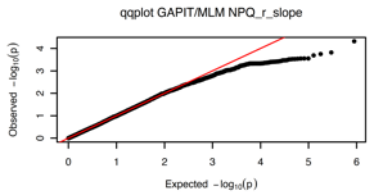


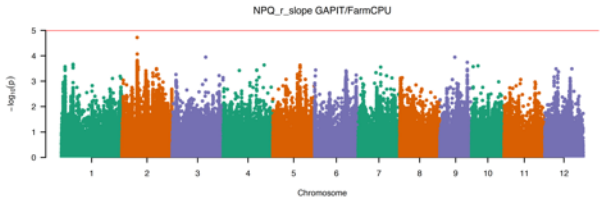


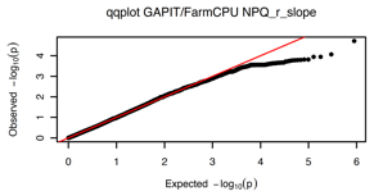




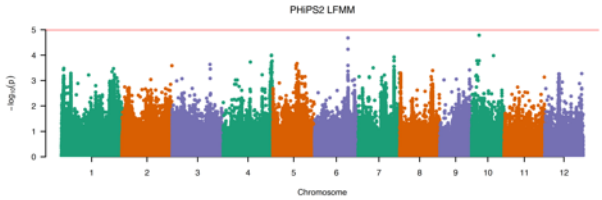


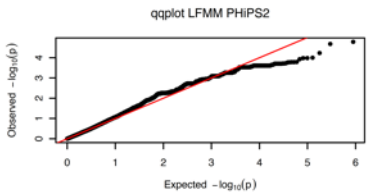


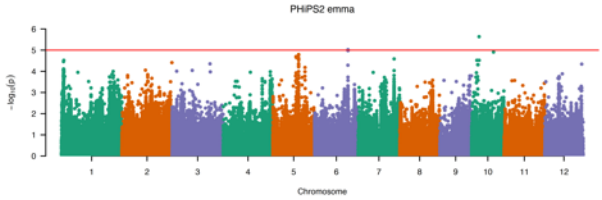


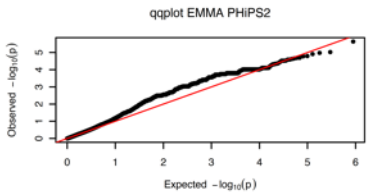


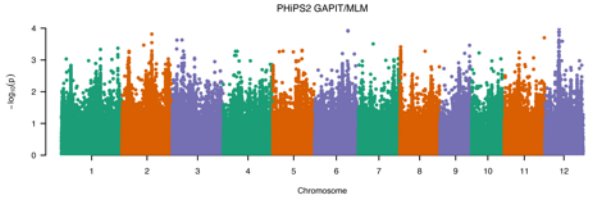
ΦPSII$_{max}$

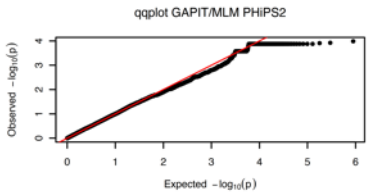


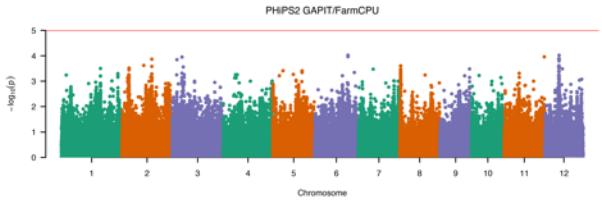


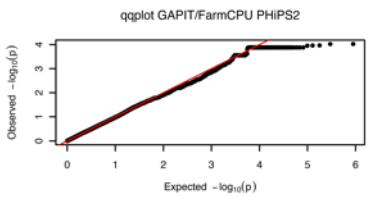




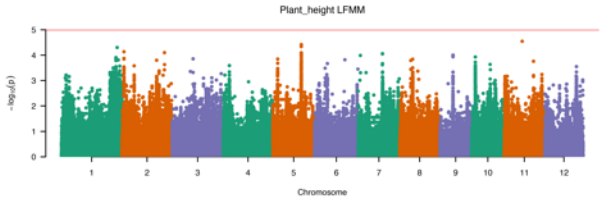


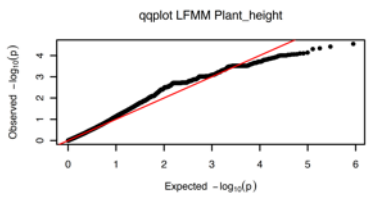


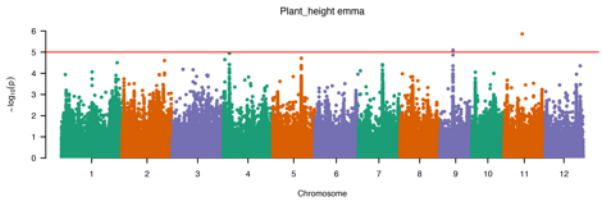


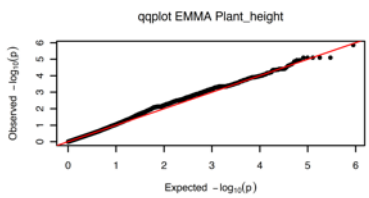


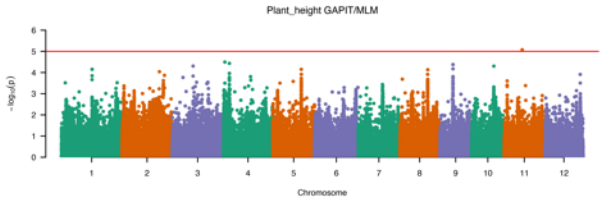
Plant height

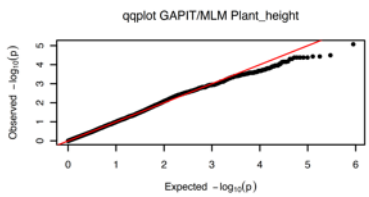


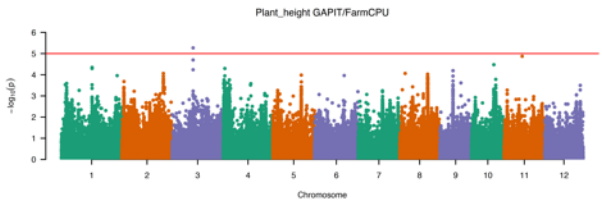


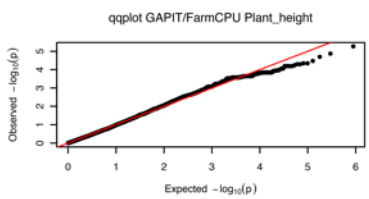




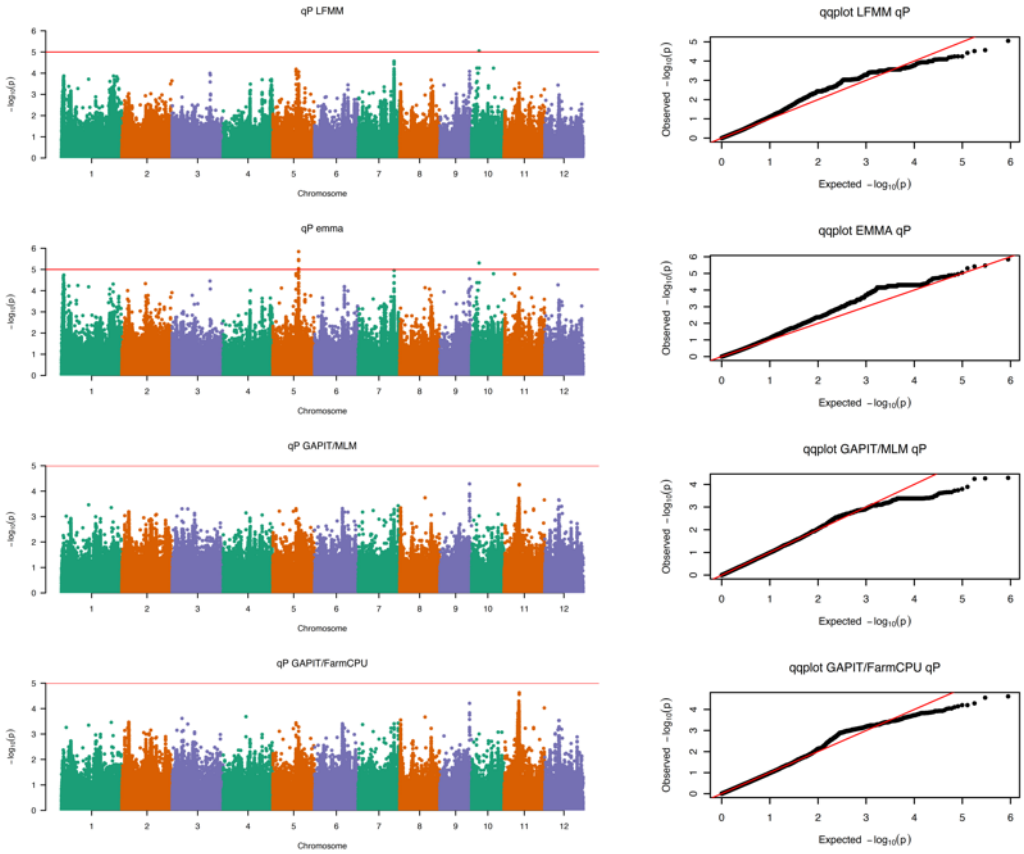




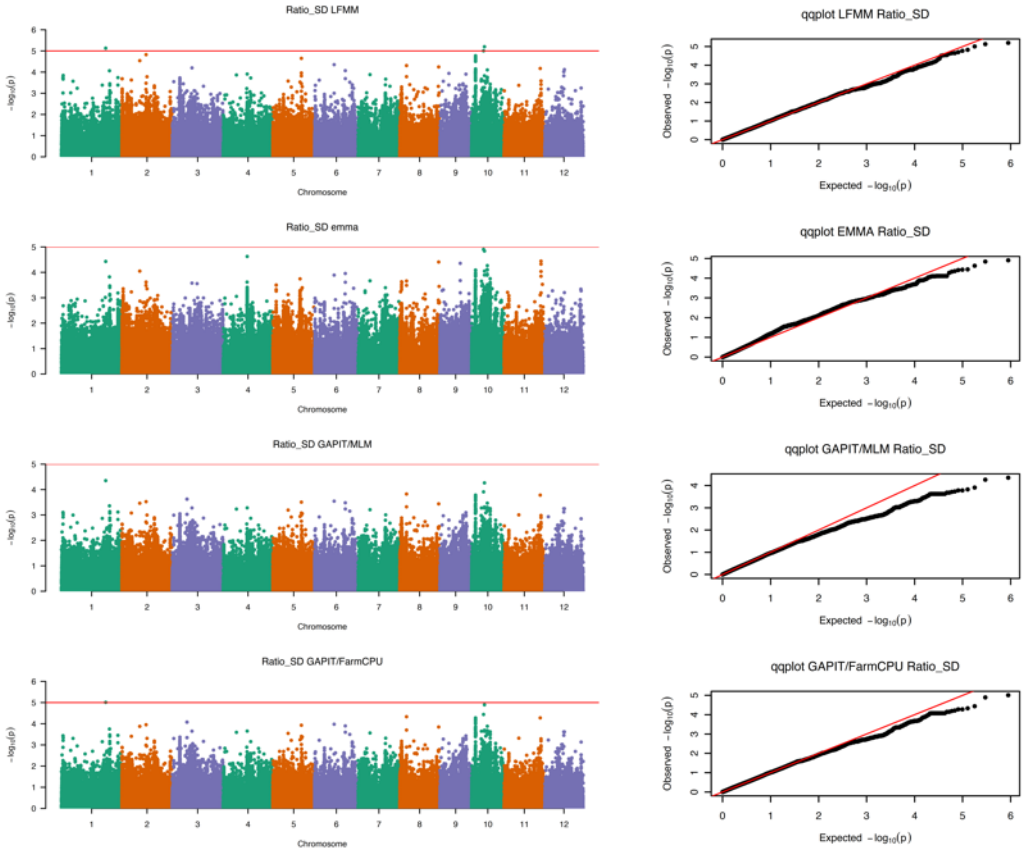




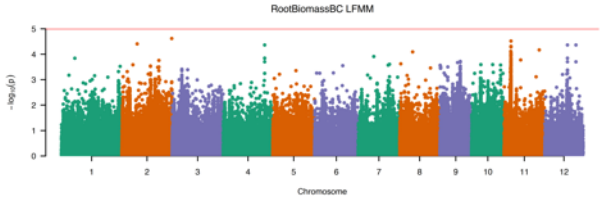
qP*max*

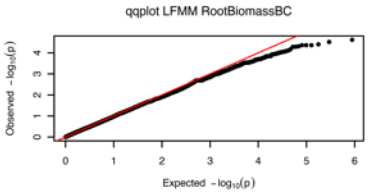


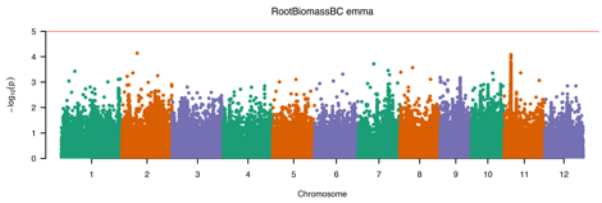
Abaxial:adaxial stomatal density ratio

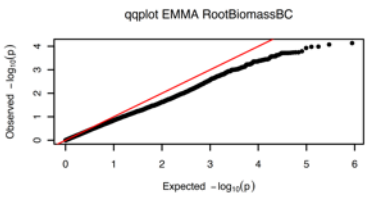


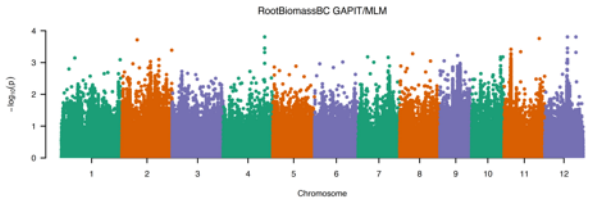
Root biomass

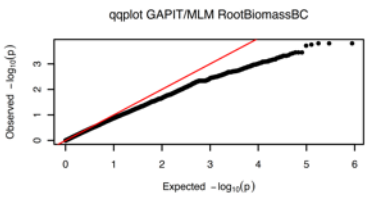


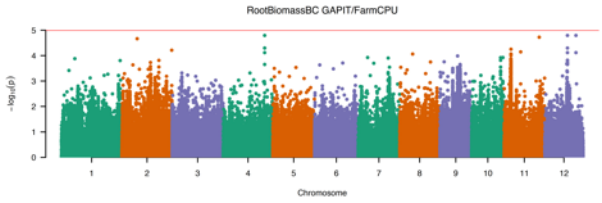


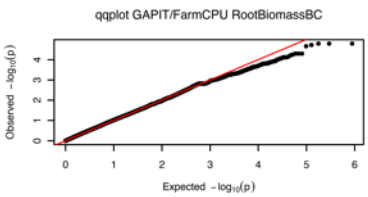




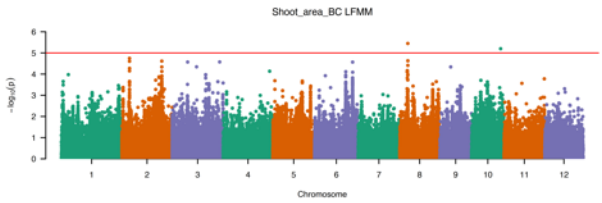


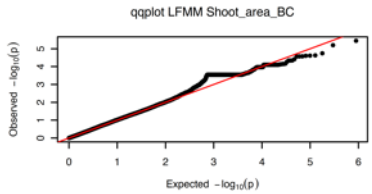


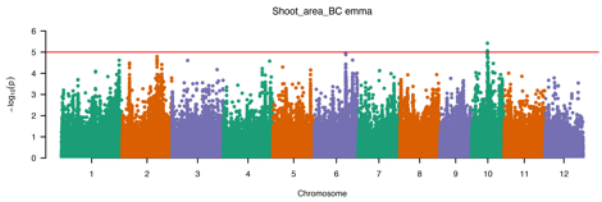


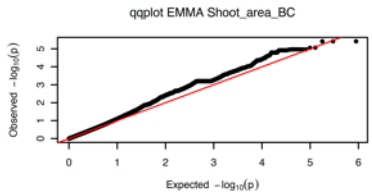


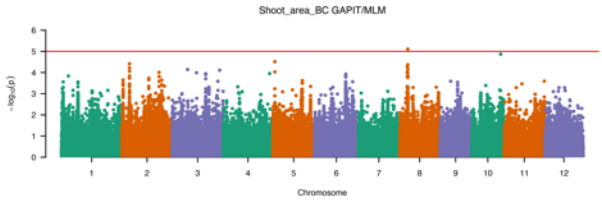
Shoot area

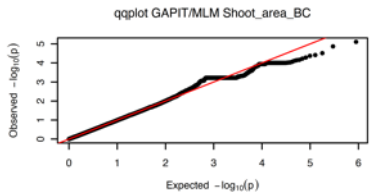


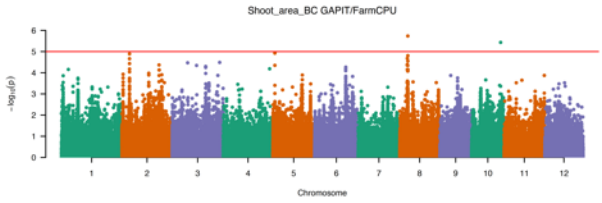


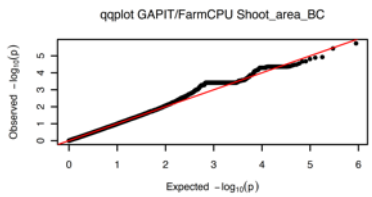




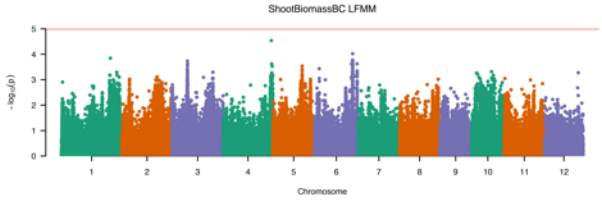


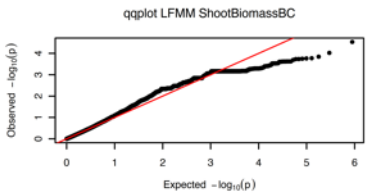


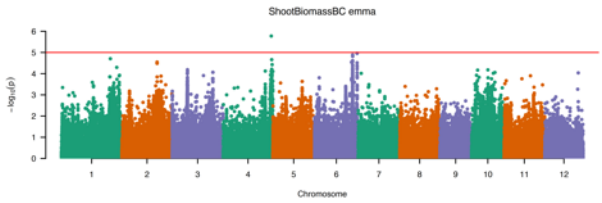


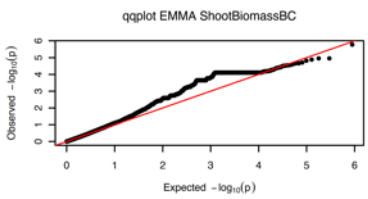


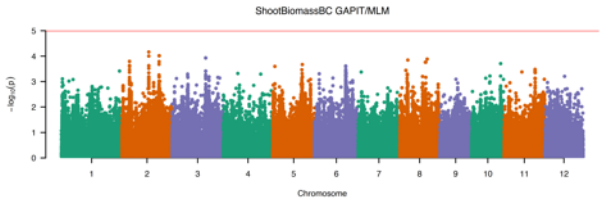
Shoot biomass

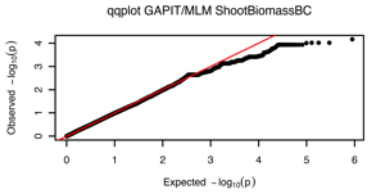


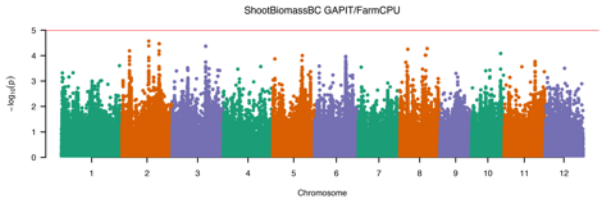


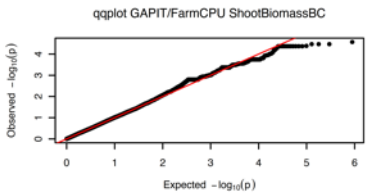




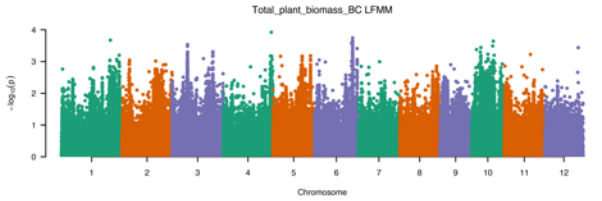


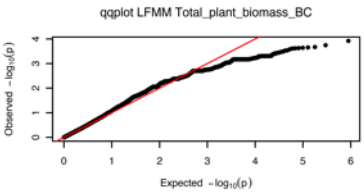


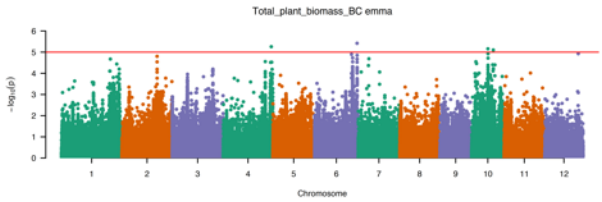


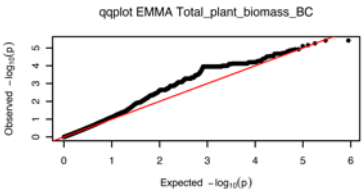


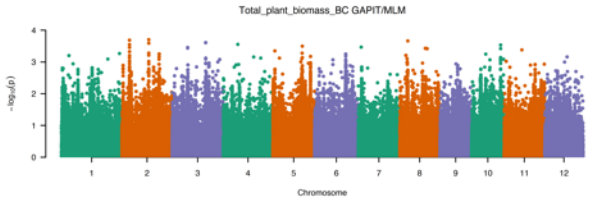
Total plant biomass

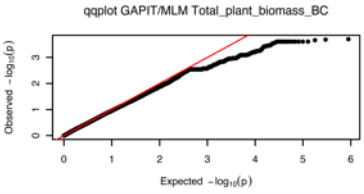


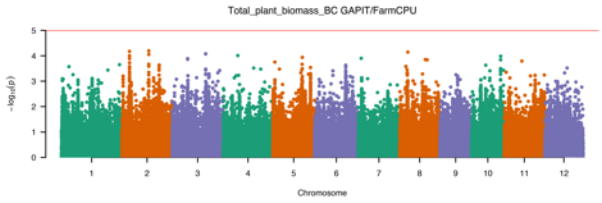


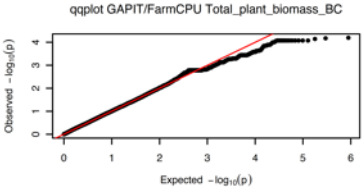




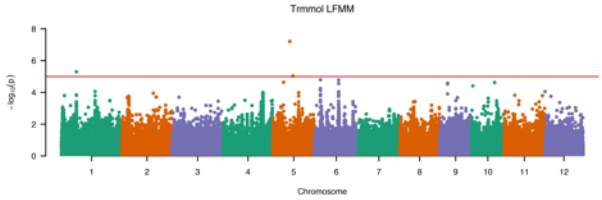


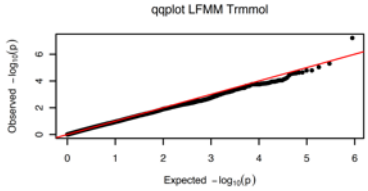


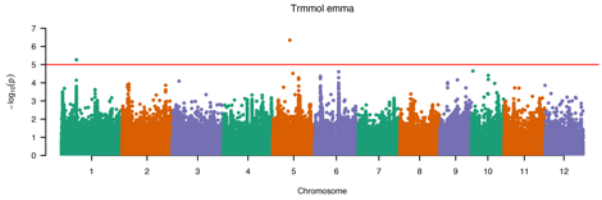


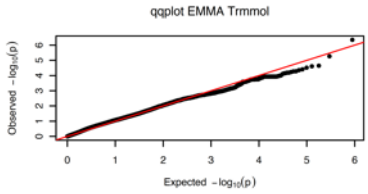


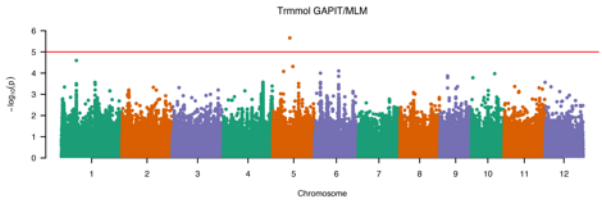
Trmmol$_{max}$

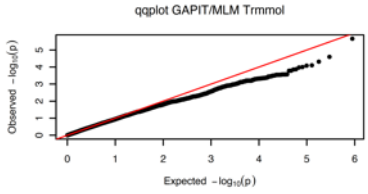


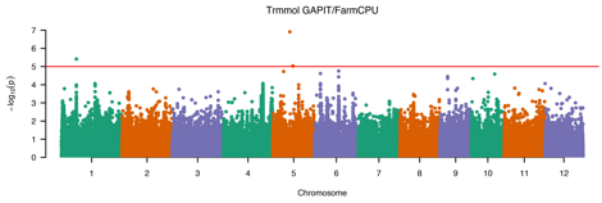


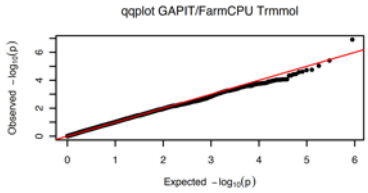




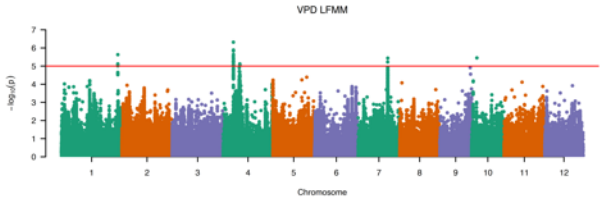


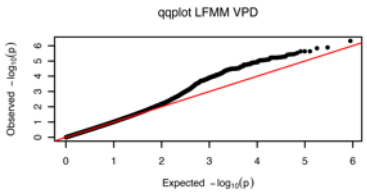


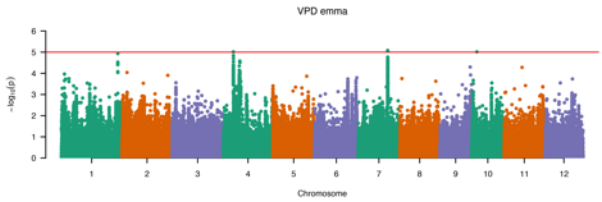


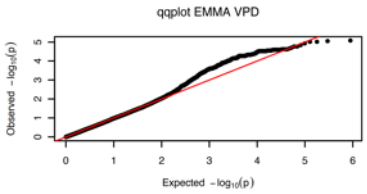


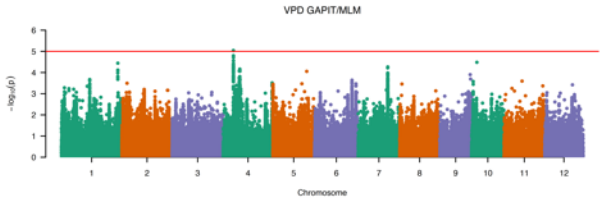
VPD$_{max}$

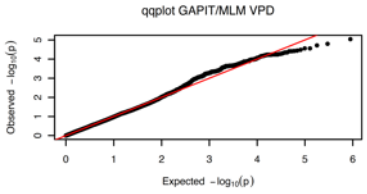


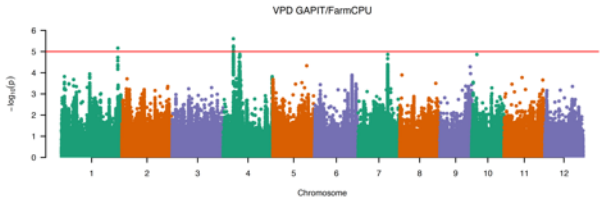


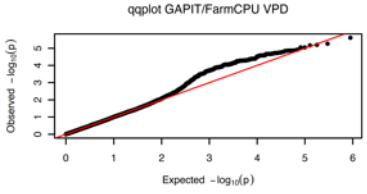




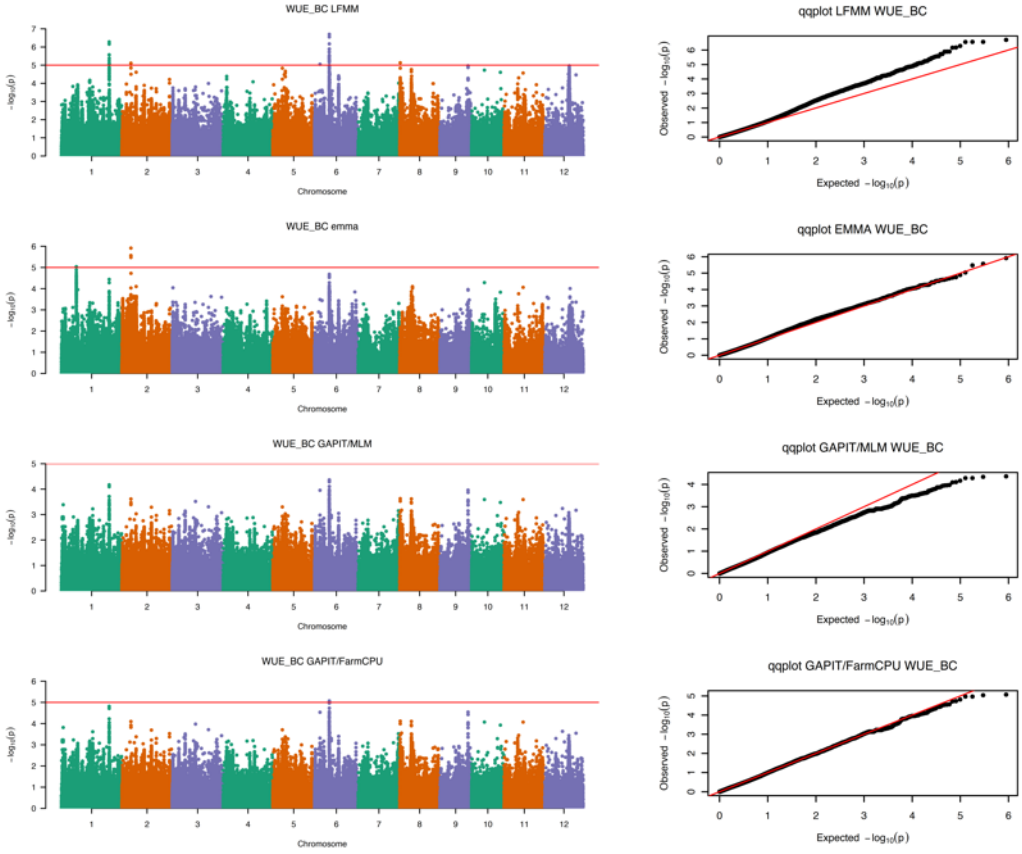








WUE*max*



Supplementary Table S4.3: Gene list for 30 phenotypic traits from the GWAS analysis detailed in chapter 4. The traits listed here were identified in 2 or more GWAS methods, SNPs were filtered based on a $p<10^{-5}$ threshold. '#CHROM' refers to the chromosome where the SNP association occurs. 'POS' is the SNP position on the chromosome. 'pValue' is the p-value given to the statistical association between the trait of interest and SNP in question 'GeneID' is the unique identifier for the candidate gene in question, where 'LOC_Os' refers to an *O. sativa* locus, the following two letters describes the chromosome number, 'g' demonstrates a gene is detected and the following five digits describe the gene order on the chromosome. 'Distance' describes how far the candidate gene is from the significant SNP. 'GeneNotes' provide extra information on the candidate gene name and function.

Trait	#CHROM	POS	pValue	GeneID	Distance	GeneNotes
A_i_10	Chr4	777545	6.97E-06	LOC_Os04g02280	-15993	OsFBX114_-_F-box_domain_containing_protein_C_expressed
A_i_90	Chr1	20980454	9.65E-06	LOC_Os01g37520	22435	expressed_protein
A_i_90	Chr1	20980454	9.65E-06	LOC_Os01g37530	19585	expressed_protein
A_i_90	Chr1	20980454	9.65E-06	LOC_Os01g37510	24939	peptide_deformylase_C_putative_C_expressed
A_i_90	Chr1	20980454	9.65E-06	LOC_Os01g37560	-7306	chaperone_protein_dnaJ_49_C_putative_C_expressed
A_i_min	Chr3	4722220	5.78E-06	LOC_Os03g09110	-15478	mitochondrial_carrier_protein_C_putative_C_expressed
A_i_min	Chr2	15424545	6.97E-06	LOC_Os02g26220	20294	hypothetical_protein
A_i_min	Chr2	15424545	6.97E-06	LOC_Os02g26260	1701	expressed_protein
A_i_min	Chr2	15424545	6.97E-06	LOC_Os02g26290	-12024	fasciclin-like_arabinogalactan_protein_8_precursor_C_putative_C_expressed
A_i_min	Chr2	15424545	6.97E-06	LOC_Os02g26294	-16630	expressed_protein
A_i_min	Chr2	15424545	6.97E-06	LOC_Os02g26300	-22666	expressed_protein
A_i_min	Chr11	25587204	9.32E-06	LOC_Os11g42470	4435	expressed_protein
A_i_min	Chr11	25587204	9.32E-06	LOC_Os11g42480	1834	transferase_family_domain_containing_protein_C_expressed
A_i_min	Chr11	25587204	9.32E-06	LOC_Os11g42500	-6551	dirigent_C_putative_C_expressed
A_i_min	Chr11	25587204	9.32E-06	LOC_Os11g42510	-9233	tyrosine_aminotransferase_C_putative_C_expressed
A_i_min	Chr11	25587204	9.32E-06	LOC_Os11g42520	-21204	DEFL34_-_Defensin_and_Defensin-like_DEFL_family_C_expressed
A_i_min	Chr11	25587204	9.32E-06	LOC_Os11g42525	-24581	hypothetical_protein
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49670	20785	cytidylyltransferase_domain_containing_protein_C_expressed
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49680	12382	DNA_repair_helicase_XPB2_C_putative_C_expressed
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49710	-416	glutathione_S-transferase_C_putative_C_expressed
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49720	-2512	glutathione_S-transferase_C_putative_C_expressed
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49730	-5828	expressed_protein
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49750	-9605	expressed_protein
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49760	-11891	expressed_protein
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49770	-16634	zinc_finger_C_C3HC4_type_domain_containing_protein_C_expressed
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49660	24955	reticulon_domain_containing_protein_C_putative_C_expressed
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49690	5510	Ser_FThr_protein_phosphatase_family_protein_C_putative_C_expressed
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49780	-23740	hypothetical_protein
A_i_min	Chr1	28566209	9.39E-06	LOC_Os01g49740	-6924	expressed_protein
A_i_min	Chr1	28571284	9.39E-06	LOC_Os01g49790	-23663	expressed_protein
A_i_min	Chr1	28572883	9.39E-06	LOC_Os01g49800	-24098	expressed_protein
A_i_min	Chr3	4707808	9.55E-06	LOC_Os03g09020	22678	dehydrogenase_C_putative_C_expressed
A_i_min	Chr3	4707808	9.55E-06	LOC_Os03g09030	17739	expressed_protein
A_i_min	Chr3	4707808	9.55E-06	LOC_Os03g09040	14284	hypothetical_protein
A_i_min	Chr3	4707808	9.55E-06	LOC_Os03g09060	7622	prenyltransferase_C_putative_C_expressed
A_i_min	Chr3	4707808	9.55E-06	LOC_Os03g09090	-12837	expressed_protein
A_i_min	Chr3	4707808	9.55E-06	LOC_Os03g09100	-18331	calmodulin-binding_transcription_activator_C_putative_C_expressed
A_i_min	Chr3	4707808	9.55E-06	LOC_Os03g09070	2313	leucine_rich_repeat_domain_containing_protein_C_putative_C_expressed
A_i_min	Chr3	4707808	9.55E-06	LOC_Os03g09080	-7288	ubiquitin_carboxyl-terminal_hydrolase_domain_containing_protein_C_expresse
A_i_min	Chr3	4753591	9.95E-06	LOC_Os03g09120	3338	expressed_protein
A_i_min	Chr3	4753591	9.95E-06	LOC_Os03g09130	-7020	expressed_protein
A_i_min	Chr3	4753591	9.95E-06	LOC_Os03g09140	-10165	ras-related_protein_C_putative_C_expressed
A_i_min	Chr3	4753591	9.95E-06	LOC_Os03g09150	-14973	pumilio-family_RNA_binding_repeat_domain_containing_protein_C_expressed
A_i_min	Chr3	4753591	9.95E-06	LOC_Os03g09160	-21974	hydroxyproline-rich_glycoprotein_family_protein_C_putative_C_expressed
A_i_rate	Chr11	15520831	1.21E-06	LOC_Os11g26950	22270	UDP-glucoronosyl_and_UDP-glucosyl_transferase_domain_containing_protein_C
A_i_rate	Chr11	15520831	1.21E-06	LOC_Os11g26956	11569	hypothetical_protein
A_i_rate	Chr11	15520831	1.21E-06	LOC_Os11g26970	840	hypothetical_protein
A_i_rate	Chr11	15520831	1.21E-06	LOC_Os11g26990	-10549	expressed_protein
A_i_rate	Chr3	26106375	7.40E-06	LOC_Os03g46140	11020	OsFBX101_-_F-box_domain_containing_protein_C_expressed
A_i_rate	Chr3	26106375	7.40E-06	LOC_Os03g46150	9364	LTPL72_-_Protease_inhibitor_Fseed_storage_FLTP_family_protein_precursor_C_
A_i_rate	Chr3	26106375	7.40E-06	LOC_Os03g46200	-13593	acetyltransferase_C_GNAT_family_C_putative_C_expressed
A_i_rate	Chr3	26106375	7.40E-06	LOC_Os03g46180	-8591	LTPL71_-_Protease_inhibitor_Fseed_storage_FLTP_family_protein_precursor_C_
A_i_rate	Chr3	26106375	7.40E-06	LOC_Os03g46190	-9473	parafibromin_C_putative_C_expressed
A_i_rate	Chr11	15554065	8.61E-06	LOC_Os11g27000	0	expressed_protein
A_i_rate	Chr11	15554065	8.61E-06	LOC_Os11g27010	-3129	hypothetical_protein
A_i_rate	Chr11	15554065	8.61E-06	LOC_Os11g27020	-6778	expressed_protein
A_i_rate	Chr11	15554065	8.61E-06	LOC_Os11g27030	-10476	expressed_protein
A_i_slope	Chr1	31110005	4.88E-07	LOC_Os01g54140	-24247	expressed_protein
A_i_slope	Chr1	31133795	7.27E-07	LOC_Os01g54170	-23598	expressed_protein

A_i_slope	Chr1	31231740	1.30E-06	LOC_Os01g54310	-24038	hypothetical_protein
A_i_slope	Chr1	31179135	2.75E-06	LOC_Os01g54240	-24308	nucleoporin_C_putative_C_expressed
A_i_slope	Chr1	31044158	3.46E-06	LOC_Os01g53970	12226	circumsporozoite_protein_precursor_C_putative_C_expressed
A_i_slope	Chr1	31044158	3.46E-06	LOC_Os01g53980	-3352	IQ_calmodulin-binding_motif_family_protein_C_expressed
A_i_slope	Chr3	26705625	3.63E-06	LOC_Os03g47140	23437	growth_regulating_factor_protein_C_putative_C_expressed
A_i_slope	Chr3	26705625	3.63E-06	LOC_Os03g47160	13672	expressed_protein
A_i_slope	Chr3	26705625	3.63E-06	LOC_Os03g47169	10338	Peptidase_family_C50_C_putative_C_expressed
A_i_slope	Chr3	26705625	3.63E-06	LOC_Os03g47178	9577	hypothetical_protein
A_i_slope	Chr3	26705625	3.63E-06	LOC_Os03g47190	-614	expressed_protein
A_i_slope	Chr3	26705625	3.63E-06	LOC_Os03g47200	-10661	bZIP_transcription_factor_domain_containing_protein_C_expressed
A_i_slope	Chr3	26705625	3.63E-06	LOC_Os03g47149	15637	expressed_protein
A_i_slope	Chr3	26705625	3.63E-06	LOC_Os03g47210	-16395	expressed_protein
A_i_slope	Chr1	30989467	3.74E-06	LOC_Os01g53860	24253	expressed_protein
A_i_slope	Chr1	30989467	3.74E-06	LOC_Os01g53880	9697	OsIAA6_-_Auxin-responsive_Aux_IAA_gene_family_member_C_expressed
A_i_slope	Chr1	30989467	3.74E-06	LOC_Os01g53890	2676	RNA_methyltransferase_C_TrmH_family_protein_C_putative_C_expressed
A_i_slope	Chr1	30989467	3.74E-06	LOC_Os01g53900	-2944	elongation_factor_C_putative_C_expressed
A_i_slope	Chr1	30989467	3.74E-06	LOC_Os01g53920	-16559	receptor-like_protein_kinase_5_precursor_C_putative_C_expressed
A_i_slope	Chr1	30989467	3.74E-06	LOC_Os01g53930	-19539	hexokinase_C_putative_C_expressed
A_i_slope	Chr1	30989467	3.74E-06	LOC_Os01g53910	-8518	dehydrogenase_C_putative_C_expressed
A_i_slope	Chr1	31097018	4.93E-06	LOC_Os01g54100	-14273	CK1_CaseinKinase_1a.2_-_CK1_includes_the_casein_kinase_1_kinases_C_expre
A_i_slope	Chr1	31168835	5.41E-06	LOC_Os01g54180	3683	DUF640_domain_containing_protein_C_putative_C_expressed
A_i_slope	Chr1	31168835	5.41E-06	LOC_Os01g54190	60	expressed_protein
A_i_slope	Chr1	31168835	5.41E-06	LOC_Os01g54210	-11361	GATA_zinc_finger_domain_containing_protein_C_expressed
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g53990	23440	pectinesterase_C_putative_C_expressed
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g54000	18745	expressed_protein
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g54010	11767	peptidase_C_putative_C_expressed
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g54020	5636	tRNA_synthetase_C_putative_C_expressed
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g54050	-10117	histidine-containing_phosphotransfer_protein_C_putative_C_expressed
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g54070	-15594	expressed_protein
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g54080	-17478	kinesin_motor_protein-related_C_putative_C_expressed
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g54030	538	NADP-dependent_malic_enzyme_C_putative_C_expressed
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g54040	0	ATP-dependent_protease_La_C_putative_C_expressed
A_i_slope	Chr1	31081915	5.74E-06	LOC_Os01g54060	-14673	expressed_protein
A_i_slope	Chr1	31227044	7.02E-06	LOC_Os01g54280	-3043	conserved_hypothetical_protein
A_i_slope	Chr1	31227044	7.02E-06	LOC_Os01g54300	-21275	OsMan02_-_Endo-Beta-Mannanase_C_expressed
A_i_slope	Chr1	31085537	8.03E-06	LOC_Os01g54090	-22780	expressed_protein
A_i_slope	Chr1	30990911	8.95E-06	LOC_Os01g53940	-24962	expressed_protein
A_i_slope	Chr1	31173515	9.02E-06	LOC_Os01g54230	-24614	expressed_protein
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12090	22495	miro_C_putative_C_expressed
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12100	20588	mTERF_family_protein_C_expressed
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12110	15891	mTERF_family_protein_C_expressed
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12129	9857	expressed_protein
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12140	6854	expressed_protein
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12150	3658	shikimate_kinase_C_putative_C_expressed
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12170	-13871	expressed_protein
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12180	-20419	jacalin-like_lectin_domain_containing_protein_C_expressed
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12120	13015	BRASSINOSTEROID_INSENSITIVE_1-associated_receptor_kinase_1_precursor_C_
A_r_10	Chr6	6501568	4.43E-06	LOC_Os06g12160	-2167	AAA-type_ATPase_family_protein_C_putative_C_expressed
A_r_10	Chr12	8156440	5.39E-06	LOC_Os12g14310	2183	expressed_protein
A_r_10	Chr12	8156440	5.39E-06	LOC_Os12g14320	588	expressed_protein
A_r_10	Chr12	8156440	5.39E-06	LOC_Os12g14330	-5032	disease_resistance_protein_RPM1_C_putative_C_expressed
A_r_10	Chr12	8156440	5.39E-06	LOC_Os12g14340	-13041	hypothetical_protein
A_r_10	Chr6	20947063	9.31E-06	LOC_Os06g35860	19945	expressed_protein
A_r_10	Chr6	20947063	9.31E-06	LOC_Os06g35870	15901	lectin_protein_kinase_family_protein_C_putative_C_expressed
A_r_10	Chr6	20947063	9.31E-06	LOC_Os06g35880	12594	hypothetical_protein
A_r_10	Chr6	20947063	9.31E-06	LOC_Os06g35890	7374	expressed_protein
A_r_10	Chr6	20947063	9.31E-06	LOC_Os06g35900	2788	BES1_FBZR1_homolog_protein_C_putative_C_expressed
A_r_10	Chr6	20947063	9.31E-06	LOC_Os06g35910	0	FYVE_zinc_finger_domain_containing_protein_C_expressed
A_r_10	Chr6	20947063	9.31E-06	LOC_Os06g35920	-6377	expressed_protein
A_r_10	Chr6	20947063	9.31E-06	LOC_Os06g35930	-13643	aquaporin_protein_C_putative_C_expressed

A_r_10	Chr6	20947063	9.31E-06	LOC_Os06g35940	-24563	osFTL12__FT-Like12_homologous_to_Flowering_Locus_T__gene%3B_contains
A_r_10	Chr2	15877698	9.37E-06	LOC_Os02g27020	-5629	expressed_protein
A_r_10	Chr2	15877698	9.37E-06	LOC_Os02g27040	-22027	expressed_protein
A_r_10	Chr2	15877698	9.37E-06	LOC_Os02g26984	21475	expressed_protein
A_r_10	Chr2	15877698	9.37E-06	LOC_Os02g27000	1209	ATP-binding_region_C_ATPase-like_domain_containing_protein_C_expressed
A_r_10	Chr2	15877698	9.37E-06	LOC_Os02g27030	-13830	cysteine_proteinase_1_precursor_C_putative_C_expressed
A_r_90	Chr2	31418445	7.27E-06	LOC_Os02g51340	-24493	hypothetical_protein
A_r_90	Chr3	4785670	7.27E-06	LOC_Os03g09130	22569	expressed_protein
A_r_90	Chr3	4785670	7.27E-06	LOC_Os03g09140	19301	ras-related_protein_C_putative_C_expressed
A_r_90	Chr3	4785670	7.27E-06	LOC_Os03g09150	12596	pumilio-family_RNA_binding_repeat_domain_containing_protein_C_expressed
A_r_90	Chr3	4785670	7.27E-06	LOC_Os03g09160	9290	hydroxyproline-rich_glycoprotein_family_protein_C_putative_C_expressed
A_r_90	Chr3	4785670	7.27E-06	LOC_Os03g09170	4100	ethylene-responsive_transcription_factor_C_putative_C_expressed
A_r_90	Chr3	4785670	7.27E-06	LOC_Os03g09180	-348	trpH_C_putative_C_expressed
A_r_90	Chr3	4785670	7.27E-06	LOC_Os03g09190	-3955	OsSCP11_-_Putative_Serine_Carboxypeptidase_homologue_C_expressed
A_r_90	Chr3	4785670	7.27E-06	LOC_Os03g09200	-6382	domain_of_unknown_function_DUF966_domain_containing_protein_C_expresse
A_r_90	Chr3	4785670	7.27E-06	LOC_Os03g09210	-22253	NADH_dehydrogenase_1_alpha_subcomplex_subunit_13_C_putative_C_express
A_r_90	Chr3	35307605	8.85E-06	LOC_Os03g62280	20216	expressed_protein
A_r_90	Chr3	35307605	8.85E-06	LOC_Os03g62290	18204	expressed_protein
A_r_90	Chr3	35307605	8.85E-06	LOC_Os03g62300	7909	expressed_protein
A_r_90	Chr3	35307605	8.85E-06	LOC_Os03g62330	-2251	expressed_protein
A_r_90	Chr3	35307605	8.85E-06	LOC_Os03g62340	-4846	protein_kinase_family_protein_C_putative_C_expressed
A_r_90	Chr3	35307605	8.85E-06	LOC_Os03g62360	-22406	expressed_protein
A_r_90	Chr3	35307605	8.85E-06	LOC_Os03g62370	-24447	WD40-like_Beta_Propeller_Repeat_family_protein_C_expressed
A_r_90	Chr2	31417249	9.00E-06	LOC_Os02g51280	18144	TCP-domain_protein_C_putative_C_expressed
A_r_90	Chr2	31417249	9.00E-06	LOC_Os02g51300	3065	AP2_domain_containing_protein_C_expressed
A_r_90	Chr2	31417249	9.00E-06	LOC_Os02g51310	0	TCP_family_transcription_factor_C_putative_C_expressed
A_r_90	Chr2	31417249	9.00E-06	LOC_Os02g51330	-15343	MAG2_C_putative_C_expressed
A_r_90	Chr2	31417249	9.00E-06	LOC_Os02g51290	7525	HVA22_C_putative_C_expressed
A_r_90	Chr2	31417249	9.00E-06	LOC_Os02g51320	-6694	helix-loop-helix_DNA-binding_domain_containing_protein_C_expressed
A_r_rate	Chr12	26468692	6.84E-07	LOC_Os12g42590	15081	expressed_protein
A_r_rate	Chr12	26468692	6.84E-07	LOC_Os12g42610	-8941	YABBY_domain_containing_protein_C_putative_C_expressed
A_r_rate	Chr12	26468692	6.84E-07	LOC_Os12g42620	-21694	hypothetical_protein
A_r_rate	Chr12	26468692	6.84E-07	LOC_Os12g42600	7354	ubiquitin_carboxyl-terminal_hydrolase_domain_containing_protein_C_expresse
A_r_slope	Chr4	3613262	3.22E-06	LOC_Os04g06830	13112	hypothetical_protein
A_r_slope	Chr4	3613262	3.22E-06	LOC_Os04g06840	4657	expressed_protein
A_r_slope	Chr4	3613262	3.22E-06	LOC_Os04g06850	0	expressed_protein
A_r_slope	Chr4	3613262	3.22E-06	LOC_Os04g06860	-4087	peroxidase_28_precursor_C_putative_C_expressed
A_r_slope	Chr4	3613262	3.22E-06	LOC_Os04g06879	-15524	expressed_protein
A_r_slope	Chr4	3613262	3.22E-06	LOC_Os04g06900	-21910	expressed_protein
A_r_slope	Chr6	22848947	4.29E-06	LOC_Os06g38520	22628	pectate_lyase_family_protein_C_expressed
A_r_slope	Chr6	22848947	4.29E-06	LOC_Os06g38564	7263	expressed_protein
A_r_slope	Chr6	22848947	4.29E-06	LOC_Os06g38580	-7468	expressed_protein
A_r_slope	Chr6	22848947	4.29E-06	LOC_Os06g38590	-14227	receptor-like_protein_kinase_precursor_C_putative_C_expressed
A_r_slope	Chr6	22848947	4.29E-06	LOC_Os06g38594	-17863	expressed_protein
A_r_slope	Chr6	22848947	4.29E-06	LOC_Os06g38600	-21833	expressed_protein
A_r_slope	Chr6	22848947	4.29E-06	LOC_Os06g38550	9407	expressed_protein
A_r_slope	Chr3	27522816	7.82E-06	LOC_Os03g48310	5495	plasma_membrane_ATPase_C_putative_C_expressed
A_r_slope	Chr3	27522816	7.82E-06	LOC_Os03g48320	1324	disease_resistance_RPP13-like_protein_1_C_putative_C_expressed
A_r_slope	Chr3	27522816	7.82E-06	LOC_Os03g48380	-23295	expressed_protein
A_r_slope	Chr3	27522816	7.82E-06	LOC_Os03g48300	18852	histidine_acid_phosphatase_C_putative_C_expressed
A_r_slope	Chr3	27522816	7.82E-06	LOC_Os03g48370	-13918	disease_resistance_RPP13-like_protein_1_C_putative_C_expressed
A_r_slope	Chr6	22812794	8.29E-06	LOC_Os06g38460	20926	expressed_protein
A_r_slope	Chr6	22812794	8.29E-06	LOC_Os06g38470	12573	histone_deacetylase_19_C_putative_C_expressed
A_r_slope	Chr6	22812794	8.29E-06	LOC_Os06g38500	-5037	expressed_protein
A_r_slope	Chr6	22812794	8.29E-06	LOC_Os06g38510	-8323	pectate_lyase_precursor_C_putative_C_expressed
A_r_slope	Chr8	8711723	8.73E-06	LOC_Os08g14490	2735	dehydrogenase-phosphopantetheinyltransferase_C_putative_C_expressed
A_r_slope	Chr8	8711723	8.73E-06	LOC_Os08g14500	-4007	hypothetical_protein
A_r_slope	Chr3	27535910	9.21E-06	LOC_Os03g48390	-16247	expressed_protein
A_r_slope	Chr3	27535910	9.21E-06	LOC_Os03g48400	-18611	hypothetical_protein
Abaxial_SD	Chr9	14943211	3.36E-06	LOC_Os09g24970	24299	expressed_protein

Abaxial_SD	Chr9	14943211	3.36E-06	LOC_Os09g24990	10814	CAF1_family_ribonuclease_containing_protein_C_putative_C_expressed
Abaxial_SD	Chr9	14943211	3.36E-06	LOC_Os09g25000	3325	spotted_leaf_11_C_putative_C_expressed
Abaxial_SD	Chr9	14943211	3.36E-06	LOC_Os09g25010	-2634	expressed_protein
Abaxial_SD	Chr9	14943211	3.36E-06	LOC_Os09g25030	-13956	expressed_protein
Abaxial_SD	Chr9	14943211	3.36E-06	LOC_Os09g25040	-16156	joka8_C_putative_C_expressed
Abaxial_SD	Chr9	14943211	3.36E-06	LOC_Os09g24980	16905	vesicle_transport_v-SNARE_protein_C_putative_C_expressed
Abaxial_SD	Chr9	5135862	3.86E-06	LOC_Os09g09500	13289	lectin-like_receptor_kinase_C_putative_C_expressed
Abaxial_SD	Chr9	5135862	3.86E-06	LOC_Os09g09510	8535	legume_lectins_beta_domain_containing_protein_C_expressed
Abaxial_SD	Chr9	5135862	3.86E-06	LOC_Os09g09520	-2022	transporter_family_protein_C_putative_C_expressed
Abaxial_SD	Chr9	5135862	3.86E-06	LOC_Os09g09550	-20474	plant_protein_of_unknown_function_domain_containing_protein_C_expressed
Abaxial_SD	Chr9	5135862	3.86E-06	LOC_Os09g09530	-7653	expressed_protein
Abaxial_SD	Chr9	5135862	3.86E-06	LOC_Os09g09560	-23598	expressed_protein
Abaxial_SD	Chr9	5135862	3.86E-06	LOC_Os09g09490	22904	disease_resistance_protein_RPM1_C_putative_C_expressed
Abaxial_SD	Chr5	20017174	4.28E-06	LOC_Os05g33890	23084	microtubule_associated_protein_C_putative_C_expressed
Abaxial_SD	Chr5	20017174	4.28E-06	LOC_Os05g33900	17607	auxin-induced_protein_5NG4_C_putative_C_expressed
Abaxial_SD	Chr5	20017174	4.28E-06	LOC_Os05g33910	13934	MATE_C_putative_C_expressed
Abaxial_SD	Chr5	20017174	4.28E-06	LOC_Os05g33920	-10924	pentatricopeptide_C_putative_C_expressed
Abaxial_SD	Chr5	20017174	4.28E-06	LOC_Os05g33930	-14777	expressed_protein
Abaxial_SD	Chr5	20017174	4.28E-06	LOC_Os05g33940	-20674	alpha_Fbeta_hydrolase_fold_C_putative_C_expressed
Abaxial_SD	Chr2	33937702	4.38E-06	LOC_Os02g55430	-15578	alginate_regulatory_protein_AlgP_C_putative_C_expressed
Abaxial_SD	Chr2	33937702	4.38E-06	LOC_Os02g55440	-18589	transmembrane_9_superfamily_member_C_putative_C_expressed
Abaxial_SD	Chr2	33946653	7.34E-06	LOC_Os02g55470	-24717	alpha-1_C6-mannosyl-glycoprotein_2-beta-N-acetylglucosaminyltransferase_C_p
Abaxial_SD	Chr2	33922369	8.46E-06	LOC_Os02g55360	13058	hypothetical_protein
Abaxial_SD	Chr2	33922369	8.46E-06	LOC_Os02g55370	7727	60S_ribosomal_protein_L39_C_putative_C_expressed
Abaxial_SD	Chr2	33922369	8.46E-06	LOC_Os02g55380	1351	AP2_domain_containing_protein_C_expressed
Abaxial_SD	Chr2	33922369	8.46E-06	LOC_Os02g55400	-9449	ATPase_8_C_plasma_membrane-type_C_putative_C_expressed
Abaxial_SD	Chr2	33922369	8.46E-06	LOC_Os02g55410	-13819	MCM5_-_Putative_minichromosome_maintenance_MCM_complex_subunit_5_C
Abaxial_SD	Chr2	33922369	8.46E-06	LOC_Os02g55420	-19655	aminotransferase_C_classes_I_and_II_C_domain_containing_protein_C_expresse
Abaxial_SD	Chr9	2730690	9.61E-06	LOC_Os09g05040	18176	expressed_protein
Abaxial_SD	Chr9	2730690	9.61E-06	LOC_Os09g05030	23342	armadillo_Fbeta-catenin-like_repeat_family_protein_C_expressed
gs_i_90	Chr2	5349758	7.43E-07	LOC_Os02g10260	-23298	outer_membrane_protein_C_OMP85_family_protein_C_expressed
gs_i_90	Chr2	5344490	1.18E-06	LOC_Os02g10250	-24787	expressed_protein
gs_i_90	Chr2	5324384	2.04E-06	LOC_Os02g10130	17948	tubulin_binding_cofactor_C_C_putative_C_expressed
gs_i_90	Chr2	5342135	5.24E-06	LOC_Os02g10240	-22244	ZOS2-05_-_C2H2_zinc_finger_protein_C_expressed
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10140	16170	bZIP_transcription_factor_domain_containing_protein_C_expressed
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10150	11769	bZIP_transcription_factor_C_putative_C_expressed
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10170	3669	expressed_protein
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10180	0	dnaJ_homolog_subfamily_C_member_C_putative_C_expressed
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10200	-11399	zinc_finger_A20_and_AN1_domain-containing_stress-associated_protein_C_puta
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10210	-14529	expressed_protein
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10220	-17276	heat_shock_protein_DnaJ_C_putative_C_expressed
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10160	5890	AGAP003371-PA_C_putative_C_expressed
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10190	-4817	DUF1680_domain_containing_protein_C_putative_C_expressed
gs_i_90	Chr2	5334191	7.90E-06	LOC_Os02g10230	-23119	metal_cation_transporter_C_putative_C_expressed
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39070	24483	5-formyltetrahydrofolate_cyclo-ligase_C_putative_C_expressed
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39080	23032	expressed_protein
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39100	14556	expressed_protein
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39110	12083	AP2_FEREBP_transcription_factor_BABY_BOOM_C_putative_C_expressed
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39114	9611	expressed_protein
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39120	5392	expressed_protein
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39130	3817	expressed_protein
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39140	2398	expressed_protein
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39180	-24391	expressed_protein
gs_i_90	Chr7	23446937	8.69E-06	LOC_Os07g39090	16051	pentatricopeptide_repeat_domain_containing_protein_C_putative_C_expressed
gs_i_min	Chr4	35003817	2.37E-07	LOC_Os04g58880	-17798	exo70_exocyst_complex_subunit_C_putative_C_expressed
gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58790	22186	expressed_protein
gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58800	16700	ubiquitin-conjugating_enzyme_C_putative_C_expressed
gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58810	12276	CAF1_family_ribonuclease_containing_protein_C_putative_C_expressed
gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58820	4353	ATOFP18_FOFP18_C_putative_C_expressed
gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58830	0	ribosome_biogenesis_regulatory_protein_C_putative_C_expressed

gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58840	-2592	Eukaryotic_aspartyl_protease_domain_containing_protein_C_expressed
gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58850	-6490	harpin-induced_protein_1_domain_containing_protein_C_expressed
gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58860	-16461	harpin-induced_protein_1_domain_containing_protein_C_expressed
gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58870	-22110	exo70_exocyst_complex_subunit_C_putative_C_expressed
gs_i_min	Chr4	34994479	1.03E-06	LOC_Os04g58780	24367	pentatricopeptide_repeat_protein_C_putative_C_expressed
gs_i_min	Chr4	35046577	4.19E-06	LOC_Os04g58920	-7886	U-box_domain-containing_protein_C_putative_C_expressed
gs_i_min	Chr4	35046577	4.19E-06	LOC_Os04g58940	-17430	expressed_protein
gs_i_min	Chr4	35046577	4.19E-06	LOC_Os04g58960	-20704	regulator_of_chromosome_condensation_C_putative_C_expressed
gs_i_min	Chr4	35017571	6.60E-06	LOC_Os04g58890	-13616	expressed_protein
gs_i_min	Chr4	35017571	6.60E-06	LOC_Os04g58900	-17255	hydrolase_C_NUDIX_family_C_domain_containing_protein_C_expressed
gs_i_min	Chr4	35017571	6.60E-06	LOC_Os04g58910	-22344	receptor_protein_kinase_TMK1_precursor_C_putative_C_expressed
gs_i_slope	Chr7	21227397	3.00E-06	LOC_Os07g35470	13046	expressed_protein
gs_i_slope	Chr7	21227397	3.00E-06	LOC_Os07g35480	1486	glucan_endo-1_C3-beta-glucosidase_precursor_C_putative_C_expressed
gs_i_slope	Chr7	21227397	3.00E-06	LOC_Os07g35490	-2520	expressed_protein
gs_i_slope	Chr7	21227397	3.00E-06	LOC_Os07g35510	-10872	glucan_endo-1_C3-beta-glucosidase_precursor_C_putative_C_expressed
gs_i_slope	Chr7	21227397	3.00E-06	LOC_Os07g35520	-17475	glucan_endo-1_C3-beta-glucosidase_precursor_C_putative_C_expressed
gs_i_slope	Chr7	21227397	3.00E-06	LOC_Os07g35500	-4716	mucin-associated_surface_protein_C_putative_C_expressed
gs_i_slope	Chr4	5861543	3.60E-06	LOC_Os04g10760	10182	expressed_protein
gs_i_slope	Chr4	5861543	3.60E-06	LOC_Os04g10770	6355	expressed_protein
gs_i_slope	Chr4	5861543	3.60E-06	LOC_Os04g10780	1671	expressed_protein
gs_i_slope	Chr4	5861543	3.60E-06	LOC_Os04g10790	0	expressed_protein
gs_i_slope	Chr4	5861543	3.60E-06	LOC_Os04g10750	12153	inorganic_phosphate_transporter_C_putative_C_expressed
gs_i_slope	Chr4	5861543	3.60E-06	LOC_Os04g10800	-3633	inorganic_phosphate_transporter_C_putative_C_expressed
gs_i_slope	Chr3	31974267	3.62E-06	LOC_Os03g56090	19638	MYB_family_transcription_factor_C_putative_C_expressed
gs_i_slope	Chr3	31974267	3.62E-06	LOC_Os03g56100	11286	expressed_protein
gs_i_slope	Chr3	31974267	3.62E-06	LOC_Os03g56120	0	expressed_protein
gs_i_slope	Chr3	31974267	3.62E-06	LOC_Os03g56130	-1903	lichenase-2_precursor_C_putative_C_expressed
gs_i_slope	Chr3	31974267	3.62E-06	LOC_Os03g56140	-12435	homeobox_protein_rough_sheath_1_C_putative_C_expressed
gs_i_slope	Chr3	31974267	3.62E-06	LOC_Os03g56110	3740	homeobox_protein_knotted-1_C_putative_C_expressed
gs_i_slope	Chr3	31939292	4.74E-06	LOC_Os03g56050	13779	AP2-like_ethylene-responsive_transcription_factor_AINTEGUMENTA_C_putative
gs_i_slope	Chr3	31939292	4.74E-06	LOC_Os03g56070	-3288	expressed_protein
gs_i_slope	Chr3	31939292	4.74E-06	LOC_Os03g56080	-7848	hypothetical_protein
gs_i_slope	Chr3	31939292	4.74E-06	LOC_Os03g56060	4668	CSLC9_-_cellulose_synthase-like_family_C_C_expressed
gs_i_slope	Chr3	31993245	4.74E-06	LOC_Os03g56160	-15155	lectin-like_receptor_kinase_7_C_putative_C_expressed
gs_i_slope	Chr3	31993245	4.74E-06	LOC_Os03g56170	-18844	expressed_protein
gs_i_slope	Chr3	31993245	4.74E-06	LOC_Os03g56190	-24399	cytochrome_c_oxidase-related_C_putative_C_expressed
gs_i_slope	Chr3	31993245	4.74E-06	LOC_Os03g56180	-19971	legume_lectins_beta_domain_containing_protein_C_expressed
gs_i_slope	Chr3	31995241	4.74E-06	LOC_Os03g56200	-24331	expressed_protein
gs_i_slope	Chr3	32015043	4.74E-06	LOC_Os03g56220	-12848	stress-induced_protein_C_putative_C_expressed
gs_i_slope	Chr3	32015043	4.74E-06	LOC_Os03g56234	-19622	myb_FSANT_domain_protein_C_putative_C_expressed
gs_i_slope	Chr7	21180803	6.77E-06	LOC_Os07g35370	22007	TKL_IRAK_DUF26-lc.15_-_DUF26_kinases_have_homology_to_DUF26_containi
gs_i_slope	Chr7	21180803	6.77E-06	LOC_Os07g35380	16911	TKL_IRAK_DUF26-lc.16_-_DUF26_kinases_have_homology_to_DUF26_containi
gs_i_slope	Chr7	21180803	6.77E-06	LOC_Os07g35400	8153	expressed_protein
gs_i_slope	Chr7	21180803	6.77E-06	LOC_Os07g35410	2240	TKL_IRAK_DUF26-lc.18_-_DUF26_kinases_have_homology_to_DUF26_containi
gs_i_slope	Chr7	21180803	6.77E-06	LOC_Os07g35390	10117	TKL_IRAK_DUF26-lc.17_-_DUF26_kinases_have_homology_to_DUF26_containi
gs_r_10	Chr9	5398647	2.61E-06	LOC_Os09g09900	18969	expressed_protein
gs_r_10	Chr9	5398647	2.61E-06	LOC_Os09g09920	5716	expressed_protein
gs_r_10	Chr9	5398647	2.61E-06	LOC_Os09g09930	1712	heavy_metal_transport_Fdetoxification_protein_C_putative_C_expressed
gs_r_10	Chr9	5398647	2.61E-06	LOC_Os09g09940	-2279	hypothetical_protein
gs_r_10	Chr9	5398647	2.61E-06	LOC_Os09g09944	-4633	expressed_protein
gs_r_10	Chr6	28796771	3.60E-06	LOC_Os06g47590	-21391	AP2_domain_containing_protein_C_expressed
gs_r_10	Chr9	5815694	4.79E-06	LOC_Os09g10630	21722	expressed_protein
gs_r_10	Chr9	5815694	4.79E-06	LOC_Os09g10650	8753	phosphatidylinositol-4-phosphate_5-Kinase_family_protein_C_putative_C_expr
gs_r_10	Chr9	5815694	4.79E-06	LOC_Os09g10660	-2176	expressed_protein
gs_r_10	Chr9	5815694	4.79E-06	LOC_Os09g10670	-4801	hypothetical_protein
gs_r_10	Chr9	5815694	4.79E-06	LOC_Os09g10680	-7090	hypothetical_protein
gs_r_10	Chr9	5815694	4.79E-06	LOC_Os09g10690	-10501	expressed_protein
gs_r_10	Chr9	5815694	4.79E-06	LOC_Os09g10700	-13702	hypothetical_protein
gs_r_10	Chr9	5815694	4.79E-06	LOC_Os09g10710	-20967	R-interacting_factor_1_C_putative_C_expressed
gs_r_10	Chr9	5415274	5.39E-06	LOC_Os09g09980	-15621	glucan_endo-1_C3-beta-glucosidase_precursor_C_putative_C_expressed

gs_r_10	Chr9	5650513	5.98E-06	LOC_Os09g10330	24208	expressed_protein
gs_r_10	Chr9	5650513	5.98E-06	LOC_Os09g10340	4990	cytochrome_P450_C_putative_C_expressed
gs_r_10	Chr9	5650513	5.98E-06	LOC_Os09g10360	-7370	expressed_protein
gs_r_10	Chr6	28788944	6.03E-06	LOC_Os06g47540	9052	expressed_protein
gs_r_10	Chr6	28788944	6.03E-06	LOC_Os06g47544	600	expressed_protein
gs_r_10	Chr6	28788944	6.03E-06	LOC_Os06g47550	-71	cadmium_Fzinc-transporting_ATPase_C_putative_C_expressed
gs_r_10	Chr6	28788944	6.03E-06	LOC_Os06g47560	-10073	expressed_protein
gs_r_10	Chr6	28788944	6.03E-06	LOC_Os06g47580	-17757	REV1_C_putative_C_expressed
gs_r_10	Chr6	28788944	6.03E-06	LOC_Os06g47520	17877	expressed_protein
gs_r_10	Chr6	28788944	6.03E-06	LOC_Os06g47530	14513	TKL_IRAK_DUF26-lh.6_-_DUF26_kinases_have_homology_to_DUF26_containir
gs_r_10	Chr6	28788944	6.03E-06	LOC_Os06g47570	-12807	PPR_repeat_containing_protein_C_expressed
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06100	22831	dihydroneopterin_aldolase_C_putative_C_expressed
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06115	9577	expressed_protein
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06130	0	glutamate_receptor_C_putative_C_expressed
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06140	-5683	hypothetical_protein
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06150	-8856	zinc_finger_C_C3HC4_type_domain_containing_protein_C_expressed
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06160	-15550	IQ_calmodulin-binding_motif_domain_containing_protein_C_expressed
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06170	-19829	expressed_protein
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06180	-21809	transferase_family_protein_C_putative_C_expressed
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06120	3567	expressed_protein
gs_r_10	Chr6	2839891	7.38E-06	LOC_Os06g06190	-23520	uncharacterized_protein_ycf45_C_putative_C_expressed
gs_r_10	Chr6	28801449	9.38E-06	LOC_Os06g47600	-24381	thaumatin_family_domain_containing_protein_C_expressed
gs_r_90	Chr1	25903592	5.66E-06	LOC_Os01g45570	19874	homeobox_associated_leucine_zipper_C_putative_C_expressed
gs_r_90	Chr1	25903592	5.66E-06	LOC_Os01g45600	-4382	expressed_protein
gs_r_90	Chr1	25903592	5.66E-06	LOC_Os01g45624	-19862	oleosin_C_putative_C_expressed
gs_r_90	Chr1	25903592	5.66E-06	LOC_Os01g45640	-22644	tat_pathway_signal_sequence_family_protein_C_expressed
gs_r_90	Chr1	25903592	5.66E-06	LOC_Os01g45650	-24998	OsFBX19_-_F-box_domain_containing_protein_C_expressed
gs_r_90	Chr1	25903592	5.66E-06	LOC_Os01g45620	-14031	CGMC_MAPKCMGC_2.5_-_CGMC_includes_CDA_C_MAPK_C_GSK3_C_and_C
gs_r_90	Chr1	25907676	5.66E-06	LOC_Os01g45659	-23473	expressed_protein
gs_r_90	Chr1	25909752	6.30E-06	LOC_Os01g45670	-23864	expressed_protein
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g73920	22007	hypothetical_protein
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g73950	10306	PPR_repeat_containing_protein_C_expressed
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g73940	13377	expressed_protein
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g73960	4776	drought_induced_19_protein_C_putative_C_expressed
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g73970	0	lysine_ketoglutarate_reductase_trans-splicing_related_1_C_putative_C_expresse
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g73980	-2730	xylem_cysteine_proteinase_2_precursor_C_putative_C_expressed
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g74000	-7781	glycerol-3-phosphate_dehydrogenase_C_putative_C_expressed
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g74010	-17597	cyclase_Fdehydrase_family_protein_C_expressed
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g74020	-21280	MYB_family_transcription_factor_C_putative_C_expressed
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g73910	24066	peptidase_C_putative_C_expressed
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g74030	-23612	pumilio-family_RNA_binding_protein_C_putative_C_expressed
gs_r_rate	Chr1	42852927	2.60E-06	LOC_Os01g73990	-5837	cytochrome_b5-like_Heme_FSteroid_binding_domain_containing_protein_C_ex
gs_r_rate	Chr4	19642144	2.93E-06	LOC_Os04g32610	0	expressed_protein
gs_r_rate	Chr4	19642144	2.93E-06	LOC_Os04g32620	-9515	ethylene-responsive_transcription_factor_ERF114_C_putative_C_expressed
gs_r_rate	Chr4	19642144	2.93E-06	LOC_Os04g32600	12286	expressed_protein
gs_r_rate	Chr2	9893249	3.51E-06	LOC_Os02g17230	13973	flavin_monooxygenase_C_putative_C_expressed
gs_r_rate	Chr2	9893249	3.51E-06	LOC_Os02g17240	4855	ATROPGEF7_FROPGEF7_C_putative_C_expressed
gs_r_rate	Chr2	9893249	3.51E-06	LOC_Os02g17250	0	expressed_protein
gs_r_rate	Chr2	9893249	3.51E-06	LOC_Os02g17270	-5598	expressed_protein
gs_r_rate	Chr2	9893249	3.51E-06	LOC_Os02g17280	-9380	gamma-secretase_subunit_APH-1B_C_putative_C_expressed
gs_r_rate	Chr9	2527147	5.85E-06	LOC_Os09g04710	22726	GDSL-like_lipase_Facylhydrolase_C_putative_C_expressed
gs_r_rate	Chr9	2527147	5.85E-06	LOC_Os09g04730	9870	dehydrogenase_Freductase_SDR_family_member_2_C_putative_C_expressed
gs_r_rate	Chr9	2527147	5.85E-06	LOC_Os09g04760	-8231	hypothetical_protein
gs_r_rate	Chr9	2527147	5.85E-06	LOC_Os09g04790	-21727	PAP_fibrillin_family_domain_containing_protein_C_expressed
gs_r_rate	Chr9	2527147	5.85E-06	LOC_Os09g04720	18853	SWIB_FMDM2_domain_containing_protein_C_expressed
gs_r_rate	Chr9	2527147	5.85E-06	LOC_Os09g04770	-12243	hypothetical_protein
gs_r_rate	Chr1	42828254	7.95E-06	LOC_Os01g73880	12964	eukaryotic_translation_initiation_factor_C_putative_C_expressed
gs_r_rate	Chr1	42828254	7.95E-06	LOC_Os01g73890	7127	transcription_initiation_factor_IIA_gamma_chain_C_putative_C_expressed
gs_r_rate	Chr1	42828254	7.95E-06	LOC_Os01g73900	3129	DEAD-box_ATP-dependent_RNA_helicase_C_putative_C_expressed
gs_r_rate	Chr12	1273195	9.32E-06	LOC_Os12g03260	17139	MATE_efflux_family_protein_C_putative_C_expressed

gs_r_rate	Chr12	1273195	9.32E-06	LOC_Os12g03270	4041	ELMO_FCED-12_family_protein_C_putative_C_expressed
gs_r_rate	Chr12	1273195	9.32E-06	LOC_Os12g03290	0	AP2_domain_containing_protein_C_expressed
gs_r_rate	Chr12	1273195	9.32E-06	LOC_Os12g03320	-11635	expressed_protein
gs_r_rate	Chr12	1273195	9.32E-06	LOC_Os12g03330	-18389	hypothetical_protein
gs_r_rate	Chr12	1273195	9.32E-06	LOC_Os12g03340	-19474	protein_of_unknown_function_C_DUF618_domain_containing_protein_C_expre
gsmax	Chr1	2954356	9.23E-07	LOC_Os01g06210	-158	gibberellin_receptor_GID1L2_C_putative_C_expressed
gsmax	Chr1	2954356	9.23E-07	LOC_Os01g06220	-2710	gibberellin_receptor_GID1L2_C_putative_C_expressed
gsmax	Chr1	2954356	9.23E-07	LOC_Os01g06230	-6838	OsSAUR1_-_Auxin-responsive_SAUR_gene_family_member_C_expressed
gsmax	Chr1	2954356	9.23E-07	LOC_Os01g06240	-8066	protein_kinase_C_putative_C_expressed
gsmax	Chr1	2954356	9.23E-07	LOC_Os01g06250	-11968	dirigent_C_putative_C_expressed
gsmax	Chr1	2954356	9.23E-07	LOC_Os01g06260	-14798	expressed_protein
gsmax	Chr1	2954356	9.23E-07	LOC_Os01g06270	-16769	expressed_protein
gsmax	Chr1	2954356	9.23E-07	LOC_Os01g06150	21464	expressed_protein
gsmax	Chr4	35003817	5.67E-06	LOC_Os04g58810	21614	CAF1_family_ribonuclease_containing_protein_C_putative_C_expressed
gsmax	Chr4	35003817	5.67E-06	LOC_Os04g58820	13691	ATOFP18_FOFP18_C_putative_C_expressed
gsmax	Chr4	35003817	5.67E-06	LOC_Os04g58830	9018	ribosome_biogenesis_regulatory_protein_C_putative_C_expressed
gsmax	Chr4	35003817	5.67E-06	LOC_Os04g58840	4649	Eukaryotic_aspartyl_protease_domain_containing_protein_C_expressed
gsmax	Chr4	35003817	5.67E-06	LOC_Os04g58850	1671	harpin-induced_protein_1_domain_containing_protein_C_expressed
gsmax	Chr4	35003817	5.67E-06	LOC_Os04g58860	-7123	harpin-induced_protein_1_domain_containing_protein_C_expressed
gsmax	Chr4	35003817	5.67E-06	LOC_Os04g58870	-12772	exo70_exocyst_complex_subunit_C_putative_C_expressed
gsmax	Chr4	35003817	5.67E-06	LOC_Os04g58880	-17798	exo70_exocyst_complex_subunit_C_putative_C_expressed
gsmax	Chr12	11142380	7.35E-06	LOC_Os12g19170	8965	expressed_protein
gsmax	Chr12	11142380	7.35E-06	LOC_Os12g19260	-23711	tetratricopeptide-like_helical_C_putative_C_expressed
gsmax	Chr12	11142380	7.35E-06	LOC_Os12g19180	775	ATP_binding_protein_C_putative_C_expressed
NPQ_i_10	Chr9	21018158	3.58E-06	LOC_Os09g36370	23692	ER_lumen_protein_retaining_receptor_containing_protein_C_expressed
NPQ_i_10	Chr9	21018158	3.58E-06	LOC_Os09g36380	22494	expressed_protein
NPQ_i_10	Chr9	21018158	3.58E-06	LOC_Os09g36420	0	hsp90_protein_C_expressed
NPQ_i_10	Chr9	21018158	3.58E-06	LOC_Os09g36430	-8098	expressed_protein
NPQ_i_10	Chr9	21018158	3.58E-06	LOC_Os09g36440	-10808	transcription_initiation_factor_IIB_C_putative_C_expressed
NPQ_i_10	Chr9	21018158	3.58E-06	LOC_Os09g36450	-14897	triosephosphate_isomerase_C_chloroplast_precursor_C_putative_C_expressed
NPQ_i_10	Chr9	21018158	3.58E-06	LOC_Os09g36460	-19956	zinc_RING_finger_protein_C_putative_C_expressed
NPQ_i_10	Chr9	21018158	3.58E-06	LOC_Os09g36400	10716	expressed_protein
NPQ_i_10	Chr12	18090308	4.16E-06	LOC_Os12g30130	20510	hypothetical_protein
NPQ_i_10	Chr8	1055591	6.38E-06	LOC_Os08g02520	15404	OsSAUR31_-_Auxin-responsive_SAUR_gene_family_member_C_expressed
NPQ_i_10	Chr8	1055591	6.38E-06	LOC_Os08g02530	9574	OsSAUR32_-_Auxin-responsive_SAUR_gene_family_member_C_expressed
NPQ_i_10	Chr8	1055591	6.38E-06	LOC_Os08g02550	491	proteasome_subunit_C_putative_C_expressed
NPQ_i_10	Chr8	1055591	6.38E-06	LOC_Os08g02560	-782	DUF1517_domain_containing_protein_C_putative_C_expressed
NPQ_i_10	Chr8	1055591	6.38E-06	LOC_Os08g02580	-13119	expressed_protein
NPQ_i_10	Chr8	1055591	6.38E-06	LOC_Os08g02590	-18580	expressed_protein
NPQ_i_10	Chr8	1055591	6.38E-06	LOC_Os08g02600	-19702	methyltransferase_domain_containing_protein_C_expressed
NPQ_i_10	Chr8	1055591	6.38E-06	LOC_Os08g02570	-6731	DUF1517_domain_containing_protein_C_putative_C_expressed
NPQ_i_10	Chr8	1055591	6.38E-06	LOC_Os08g02540	5582	adenylate_kinase_C_putative_C_expressed
NPQ_i_10	Chr12	18113120	8.46E-06	LOC_Os12g30150	11283	CAMK_CAMK_like.47_-_CAMK_includes_calcium_Fcalmodulin_depedent_prote
NPQ_i_10	Chr12	18113120	8.46E-06	LOC_Os12g30160	3121	hypothetical_protein
NPQ_i_10	Chr12	18113120	8.46E-06	LOC_Os12g30170	0	expressed_protein
NPQ_i_10	Chr12	18113120	8.46E-06	LOC_Os12g30180	-2788	protein_kinase_domain_containing_protein_C_expressed
NPQ_i_10	Chr12	18113120	8.46E-06	LOC_Os12g30190	-8623	expressed_protein
NPQ_i_90	Chr2	25845312	3.25E-08	LOC_Os02g42920	24432	DTA2_C_putative_C_expressed
NPQ_i_90	Chr2	25845312	3.25E-08	LOC_Os02g42930	21163	expressed_protein
NPQ_i_90	Chr2	25845312	3.25E-08	LOC_Os02g42940	13676	MSP_domain_containing_protein_C_expressed
NPQ_i_90	Chr2	25845312	3.25E-08	LOC_Os02g42950	2583	YABBY_domain_containing_protein_C_putative_C_expressed
NPQ_i_90	Chr2	25845312	3.25E-08	LOC_Os02g42960	-10220	thylakoid_lumenal_protein_C_putative_C_expressed
NPQ_i_90	Chr2	25845312	3.25E-08	LOC_Os02g42970	-14192	NAC_domain_containing_protein_C_putative_C_expressed
NPQ_i_90	Chr7	25675834	1.41E-06	LOC_Os07g42910	-16000	cytochrome_c_oxidase_subunit_C_putative_C_expressed
NPQ_i_90	Chr7	25675834	1.41E-06	LOC_Os07g42924	-20966	dehydrogenase_C_putative_C_expressed
NPQ_i_90	Chr7	25638781	2.93E-06	LOC_Os07g42860	-23986	cyclin_C_putative_C_expressed
NPQ_i_90	Chr7	25659807	4.31E-06	LOC_Os07g42900	-23466	helicase_domain-containing_protein_C_putative_C_expressed
NPQ_i_90	Chr5	20030433	4.84E-06	LOC_Os05g33920	1079	pentatricopeptide_C_putative_C_expressed
NPQ_i_90	Chr5	20030433	4.84E-06	LOC_Os05g33930	-1518	expressed_protein
NPQ_i_90	Chr5	20030433	4.84E-06	LOC_Os05g33940	-7415	alpha_Fbeta_hydrolase_fold_C_putative_C_expressed

NPQ_i_90	Chr5	20030433	4.84E-06	LOC_Os05g33960	-14761	peptide_transporter_PTR2_C_putative_C_expressed
NPQ_i_90	Chr5	20030433	4.84E-06	LOC_Os05g33970	-24913	hypothetical_protein
NPQ_i_90	Chr7	25656984	5.68E-06	LOC_Os07g42895	-23849	expressed_protein
NPQ_i_90	Chr7	25622487	5.91E-06	LOC_Os07g42820	-23501	ADP-ribosylation_factor_C_putative_C_expressed
NPQ_i_90	Chr7	25636635	5.94E-06	LOC_Os07g42850	-24935	hypothetical_protein
NPQ_i_90	Chr7	25648503	6.29E-06	LOC_Os07g42880	-23698	PPR_repeat_containing_protein_C_expressed
NPQ_i_90	Chr7	25619020	6.63E-06	LOC_Os07g42790	-14650	caltractin_C_putative_C_expressed
NPQ_i_90	Chr7	25619020	6.63E-06	LOC_Os07g42800	-17683	heat_shock_protein_DnaJ_C_putative_C_expressed
NPQ_i_90	Chr7	25619020	6.63E-06	LOC_Os07g42810	-21827	adaptor_complexes_medium_subunit_family_domain_containing_protein_C_exp
NPQ_i_90	Chr6	25687749	7.49E-06	LOC_Os06g42680	15805	phytosulfokines_1_precursor_C_putative_C_expressed
NPQ_i_90	Chr6	25687749	7.49E-06	LOC_Os06g42690	10259	bZIP_transcription_factor_domain_containing_protein_C_expressed
NPQ_i_90	Chr6	25687749	7.49E-06	LOC_Os06g42700	5474	zinc_finger_C_C3HC4_type_domain_containing_protein_C_expressed
NPQ_i_90	Chr6	25687749	7.49E-06	LOC_Os06g42730	-10992	OsPOP14_-_Putative_Prolyl_Oligopeptidase_homologue_C_expressed
NPQ_i_90	Chr6	25687749	7.49E-06	LOC_Os06g42740	-16292	expressed_protein
NPQ_i_90	Chr6	25687749	7.49E-06	LOC_Os06g42754	-20928	expressed_protein
NPQ_i_90	Chr6	25687749	7.49E-06	LOC_Os06g42770	-24119	type_II_intron_maturase_protein_C_putative_C_expressed
NPQ_i_90	Chr6	25687749	7.49E-06	LOC_Os06g42720	-2348	amino_acid_transporter_C_putative_C_expressed
NPQ_i_90	Chr4	3298032	7.89E-06	LOC_Os04g06300	0	expressed_protein
NPQ_i_90	Chr4	3298032	7.89E-06	LOC_Os04g06340	-12303	MEGL11_-_Maternally_expressed_gene_MEG_family_protein_precursor_C_expr
NPQ_i_90	Chr4	3298032	7.89E-06	LOC_Os04g06350	-14088	expressed_protein
NPQ_i_90	Chr4	3298032	7.89E-06	LOC_Os04g06280	20181	expressed_protein
NPQ_i_90	Chr5	19941411	8.23E-06	LOC_Os05g33810	8949	OsSPL9_-_SBP-box_gene_family_member_C_expressed
NPQ_i_90	Chr5	19941411	8.23E-06	LOC_Os05g33820	3167	lipase_C_putative_C_expressed
NPQ_i_90	Chr5	19941411	8.23E-06	LOC_Os05g33830	-11971	zinc_finger_C_C3HC4_type_domain_containing_protein_C_expressed
NPQ_i_90	Chr5	19941411	8.23E-06	LOC_Os05g33840	-15728	transketolase_C_putative_C_expressed
NPQ_i_90	Chr7	25652799	8.35E-06	LOC_Os07g42890	-24996	GRAM_domain_containing_protein_C_expressed
NPQ_i_90	Chr7	25602948	8.44E-06	LOC_Os07g42730	16714	EF_hand_family_protein_C_expressed
NPQ_i_90	Chr7	25602948	8.44E-06	LOC_Os07g42750	0	DDT_domain_containing_protein_C_putative_C_expressed
NPQ_i_90	Chr7	25602948	8.44E-06	LOC_Os07g42760	-8977	expressed_protein
NPQ_i_90	Chr7	25602948	8.44E-06	LOC_Os07g42770	-14216	CAMK_CAMK_like.35_-_CAMK_includes_calcium_Fcalmodulin_depedent_prot
NPQ_i_90	Chr7	25602948	8.44E-06	LOC_Os07g42780	-24655	expressed_protein
NPQ_i_90	Chr7	25602948	8.44E-06	LOC_Os07g42740	10138	DUF1645_domain_containing_protein_C_putative_C_expressed
NPQ_i_90	Chr3	27729841	9.63E-06	LOC_Os03g48600	19343	domain_of_unknown_function_DUF966_domain_containing_protein_C_expresse
NPQ_i_90	Chr3	27729841	9.63E-06	LOC_Os03g48610	9665	galactosyltransferase_family_protein_C_putative_C_expressed
NPQ_i_90	Chr3	27729841	9.63E-06	LOC_Os03g48626	787	expressed_protein
NPQ_i_90	Chr3	27729841	9.63E-06	LOC_Os03g48642	-732	hypothetical_protein
NPQ_i_90	Chr3	27729841	9.63E-06	LOC_Os03g48660	-6450	DUF1336_domain_containing_protein_C_expressed
NPQ_i_90	Chr9	18720950	9.81E-06	LOC_Os09g31100	19018	expressed_protein
NPQ_i_90	Chr9	18720950	9.81E-06	LOC_Os09g31120	10635	pirin_C_putative_C_expressed
NPQ_i_90	Chr9	18720950	9.81E-06	LOC_Os09g31130	2086	citrate_transporter_C_putative_C_expressed
NPQ_i_90	Chr9	18720950	9.81E-06	LOC_Os09g31140	-14065	ZOS9-16_-_C2H2_zinc_finger_protein_C_expressed
NPQ_i_90	Chr9	18720950	9.81E-06	LOC_Os09g31160	-19582	glutamate_receptor_precursor_C_putative_C_expressed
NPQ_i_rate	Chr6	21007563	2.16E-06	LOC_Os06g35950	24315	expressed_protein
NPQ_i_rate	Chr6	21007563	2.16E-06	LOC_Os06g35960	8696	HSF-type_DNA-binding_domain_containing_protein_C_expressed
NPQ_i_rate	Chr6	21007563	2.16E-06	LOC_Os06g35970	0	meiosis_5_C_putative_C_expressed
NPQ_i_rate	Chr6	21007563	2.16E-06	LOC_Os06g35980	-7896	expressed_protein
NPQ_i_rate	Chr6	21007563	2.16E-06	LOC_Os06g35990	-14274	expressed_protein
NPQ_i_rate	Chr6	20920460	5.30E-06	LOC_Os06g35814	18744	ras-related_protein_C_putative_C_expressed
NPQ_i_rate	Chr6	20920460	5.30E-06	LOC_Os06g35830	13249	expressed_protein
NPQ_i_rate	Chr6	20920460	5.30E-06	LOC_Os06g35850	1963	lectin_protein_kinase_family_protein_C_putative_C_expressed
NPQ_i_rate	Chr6	20920460	5.30E-06	LOC_Os06g35860	-4076	expressed_protein
NPQ_i_rate	Chr6	20920460	5.30E-06	LOC_Os06g35870	-8780	lectin_protein_kinase_family_protein_C_putative_C_expressed
NPQ_i_rate	Chr6	20920460	5.30E-06	LOC_Os06g35880	-11172	hypothetical_protein
NPQ_i_rate	Chr6	20920460	5.30E-06	LOC_Os06g35890	-18875	expressed_protein
NPQ_i_rate	Chr6	20920460	5.30E-06	LOC_Os06g35900	-20660	BES1_FBZR1_homolog_protein_C_putative_C_expressed
NPQ_i_rate	Chr6	20929176	6.44E-06	LOC_Os06g35920	-24264	expressed_protein
NPQ_i_rate	Chr7	25680187	7.39E-06	LOC_Os07g42850	17865	hypothetical_protein
NPQ_i_rate	Chr7	25680187	7.39E-06	LOC_Os07g42860	14871	cyclin_C_putative_C_expressed
NPQ_i_rate	Chr7	25680187	7.39E-06	LOC_Os07g42890	339	GRAM_domain_containing_protein_C_expressed
NPQ_i_rate	Chr7	25680187	7.39E-06	LOC_Os07g42895	-646	expressed_protein

NPQ_i_rate	Chr7	25680187	7.39E-06	LOC_Os07g42900	-3086	helicase_domain-containing_protein_C_putative_C_expressed
NPQ_i_rate	Chr7	25680187	7.39E-06	LOC_Os07g42910	-11647	cytochrome_c_oxidase_subunit_C_putative_C_expressed
NPQ_i_rate	Chr7	25680187	7.39E-06	LOC_Os07g42924	-16613	dehydrogenase_C_putative_C_expressed
NPQ_i_rate	Chr7	25680187	7.39E-06	LOC_Os07g42880	6112	PPR_repeat_containing_protein_C_expressed
NPQ_i_rate	Chr4	19171062	7.59E-06	LOC_Os04g31971	19041	expressed_protein
NPQ_i_rate	Chr4	19171062	7.59E-06	LOC_Os04g32000	9266	expressed_protein
NPQ_i_rate	Chr4	19171062	7.59E-06	LOC_Os04g32004	6029	expressed_protein
NPQ_i_rate	Chr4	19171062	7.59E-06	LOC_Os04g32020	-4641	2-oxoglutarate_dehydrogenase_E1_component_C_mitochondrial_precursor_C_pu
NPQ_i_rate	Chr4	19171062	7.59E-06	LOC_Os04g32030	-18060	heavy_metal-associated_domain_containing_protein_C_expressed
NPQ_i_rate	Chr4	19171062	7.59E-06	LOC_Os04g32010	0	thiamine_pyrophosphate_enzyme_C_C-terminal_TPP_binding_domain_containin
NPQ_i_rate	Chr1	4403414	7.73E-06	LOC_Os01g08760	6414	choline_Fethanolamine_kinase_C_putative_C_expressed
NPQ_i_rate	Chr1	4403414	7.73E-06	LOC_Os01g08770	1333	WD_domain_C_G-beta_repeat_domain_containing_protein_C_expressed
NPQ_i_rate	Chr1	4403414	7.73E-06	LOC_Os01g08780	0	endonuclease_Fexonuclease_Fphosphatase_family_domain_containing_protein_C
NPQ_i_rate	Chr1	4403414	7.73E-06	LOC_Os01g08790	-12002	histone-like_transcription_factor_and_archaeal_histone_C_putative_C_expressed
NPQ_i_rate	Chr1	4403414	7.73E-06	LOC_Os01g08800	-13869	cytochrome_P450_C_putative_C_expressed
NPQ_i_rate	Chr1	4403414	7.73E-06	LOC_Os01g08810	-17904	cytochrome_P450_C_putative_C_expressed
NPQ_i_rate	Chr1	4403414	7.73E-06	LOC_Os01g08814	-20478	expressed_protein
NPQ_i_rate	Chr1	4403414	7.73E-06	LOC_Os01g08750	9331	expressed_protein
NPQ_i_rate	Chr6	20924713	7.79E-06	LOC_Os06g35910	-21431	FYVE_zinc_finger_domain_containing_protein_C_expressed
NPQ_i_slope	Chr12	18122989	4.69E-06	LOC_Os12g30150	21152	CAMK_CAMK_like.47_-_CAMK_includes_calcium_Fcalmodulin_depedent_prote
NPQ_i_slope	Chr12	18122989	4.69E-06	LOC_Os12g30160	12990	hypothetical_protein
NPQ_i_slope	Chr12	18122989	4.69E-06	LOC_Os12g30170	9254	expressed_protein
NPQ_i_slope	Chr12	18122989	4.69E-06	LOC_Os12g30180	2309	protein_kinase_domain_containing_protein_C_expressed
NPQ_i_slope	Chr12	18122989	4.69E-06	LOC_Os12g30190	995	expressed_protein
NPQ_r_10	Chr7	3834831	1.21E-06	LOC_Os07g07690	-24924	PHD-finger_domain_containing_protein_C_expressed
NPQ_r_10	Chr7	3848041	1.27E-06	LOC_Os07g07709	-24093	ribosomal_protein_S13p_FS18e_C_putative_C_expressed
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07580	22298	ROOT_HAIRLESS_1_C_putative_C_expressed
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07590	17194	expressed_protein
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07610	5800	RNA_polymerase_subunit_C_putative_C_expressed
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07620	2412	pentatricopeptide_C_putative_C_expressed
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07630	0	expressed_protein
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07640	-2055	hypothetical_protein
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07646	-7370	expressed_protein
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07654	-11089	BT1_family_protein_C_putative_C_expressed
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07670	-20433	expressed_protein
NPQ_r_10	Chr7	3823141	2.82E-06	LOC_Os07g07680	-23230	frigida_C_putative_C_expressed
NPQ_r_10	Chr7	3855998	4.15E-06	LOC_Os07g07715	-24798	expressed_protein
NPQ_r_10	Chr6	23087961	9.46E-06	LOC_Os06g38870	19810	expressed_protein
NPQ_r_10	Chr6	23087961	9.46E-06	LOC_Os06g38880	17816	hypothetical_protein
NPQ_r_10	Chr6	23087961	9.46E-06	LOC_Os06g38910	0	expressed_protein
NPQ_r_10	Chr6	23087961	9.46E-06	LOC_Os06g38930	-6057	leucine-rich_repeat_receptor_protein_kinase_EXS_precursor_C_putative_C_expre
NPQ_r_10	Chr6	23087961	9.46E-06	LOC_Os06g38950	-17264	ABC_transporter_C_ATP-binding_protein_C_putative_C_expressed
NPQ_r_10	Chr6	23087961	9.46E-06	LOC_Os06g38940	-10896	RMD5_homolog_A_C_putative_C_expressed
NPQmax	Chr2	104656	1.78E-06	LOC_Os02g01200	-7772	glutaredoxin_C_putative_C_expressed
NPQmax	Chr2	104656	1.78E-06	LOC_Os02g01210	-14735	amino_acid_transporter_C_putative_C_expressed
NPQmax	Chr5	20911445	2.95E-06	LOC_Os05g35190	13554	powdery_mildew_resistant_protein_5_C_putative_C_expressed
NPQmax	Chr5	20911445	2.95E-06	LOC_Os05g35200	7804	glycosyl_transferase_C_putative_C_expressed
NPQmax	Chr5	20911445	2.95E-06	LOC_Os05g35210	-2470	hypothetical_protein
NPQmax	Chr5	20911445	2.95E-06	LOC_Os05g35220	-18072	hypothetical_protein
NPQmax	Chr5	20911445	2.95E-06	LOC_Os05g35230	-19669	expressed_protein
NPQmax	Chr5	20911445	2.95E-06	LOC_Os05g35240	-24777	expressed_protein
NPQmax	Chr2	202415	4.05E-06	LOC_Os02g01380	-17162	expressed_protein
NPQmax	Chr5	25399165	4.33E-06	LOC_Os05g43650	13532	expressed_protein
NPQmax	Chr5	25399165	4.33E-06	LOC_Os05g43660	11197	expressed_protein
NPQmax	Chr5	25399165	4.33E-06	LOC_Os05g43670	6644	IQ_calmodulin-binding_motif_family_protein_C_putative_C_expressed
NPQmax	Chr5	25399165	4.33E-06	LOC_Os05g43690	-76	glucan_endo-1_C3-beta-glucosidase-like_protein_3_precursor_C_putative_C_ex
NPQmax	Chr5	25399165	4.33E-06	LOC_Os05g43700	-3250	expressed_protein
NPQmax	Chr5	25399165	4.33E-06	LOC_Os05g43680	3111	expressed_protein
NPQmax	Chr2	81643	5.93E-06	LOC_Os02g01110	16917	translation_initiation_factor_IF-3_C_putative_C_expressed
NPQmax	Chr2	81643	5.93E-06	LOC_Os02g01120	14690	expressed_protein

NPQmax	Chr2	81643	5.93E-06	LOC_Os02g01130	7779	expressed_protein
NPQmax	Chr2	81643	5.93E-06	LOC_Os02g01140	2064	GDSL-like_lipase_Facylhydrolase_C_putative_C_expressed
NPQmax	Chr2	81643	5.93E-06	LOC_Os02g01160	-3689	skp1_family_C_tetramerisation_domain_containing_protein_C_expressed
NPQmax	Chr2	81643	5.93E-06	LOC_Os02g01170	-4821	HECT-domain_domain_containing_protein_C_expressed
NPQmax	Chr2	81643	5.93E-06	LOC_Os02g01150	0	erythronate-4-phosphate_dehydrogenase_domain_containing_protein_C_expressed
NPQmax	Chr2	81643	5.93E-06	LOC_Os02g01180	-16355	kinesin_motor_domain_containing_protein_C_putative_C_expressed
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01220	19955	rhodanese-like_domain_containing_protein_C_putative_C_expressed
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01230	17403	ribosomal_protein_C_putative_C_expressed
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01245	12437	expressed_protein
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01240	12437	CPuORF11_-_conserved_peptide_uORF-containing_transcript_C_expressed
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01250	9495	LSM_domain_containing_protein_C_expressed
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01275	-5196	expressed_protein
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01280	-11070	T-complex_protein_C_putative_C_expressed
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01300	-17901	invertase_Fpectin_methylesterase_inhibitor_family_protein_C_putative_C_expre
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01310	-20849	invertase_Fpectin_methylesterase_inhibitor_family_protein_C_putative_C_expre
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01270	0	START_domain_containing_protein_C_expressed
NPQmax	Chr2	147592	6.99E-06	LOC_Os02g01290	-15719	expressed_protein
NPQmax	Chr2	185826	8.38E-06	LOC_Os02g01326	6456	auxin-induced_protein_5NG4_C_putative_C_expressed
NPQmax	Chr2	185826	8.38E-06	LOC_Os02g01332	2384	ribosomal_protein_L6_C_putative_C_expressed
NPQmax	Chr2	185826	8.38E-06	LOC_Os02g01340	499	ferredoxin--NADP_reductase_C_chloroplast_precursor_C_putative_C_expressed
NPQmax	Chr2	185826	8.38E-06	LOC_Os02g01350	-1652	expressed_protein
NPQmax	Chr2	185826	8.38E-06	LOC_Os02g01365	-13398	expressed_protein
NPQmax	Chr2	185826	8.38E-06	LOC_Os02g01355	-13398	MADS-box_transcription_factor_C_putative_C_expressed
NPQmax	Chr2	185826	8.38E-06	LOC_Os02g01360	-19598	OsMADS60_-_MADS-box_family_gene_with_MIKCc_type-box_C_expressed
NPQmax	Chr2	185826	8.38E-06	LOC_Os02g01370	-21924	hypothetical_protein
NPQmax	Chr2	86472	8.76E-06	LOC_Os02g01190	-24291	POEI25_-_Pollen_Ole_e_I_allergen_and_extensin_family_protein_precursor_C_e
Plant_height	Chr11	12938384	1.38E-06	LOC_Os11g22490	9636	hypothetical_protein
Plant_height	Chr3	14952545	5.39E-06	LOC_Os03g26060	16787	expressed_protein
Plant_height	Chr3	14952545	5.39E-06	LOC_Os03g26070	14526	expressed_protein
Plant_height	Chr3	14952545	5.39E-06	LOC_Os03g26090	0	GPI-anchored_wall_transfer_protein_1_C_putative_C_expressed
Plant_height	Chr3	14952545	5.39E-06	LOC_Os03g26120	-11182	hypothetical_protein
Plant_height	Chr3	14952545	5.39E-06	LOC_Os03g26130	-15054	MYB_family_transcription_factor_C_putative_C_expressed
Plant_height	Chr3	14952545	5.39E-06	LOC_Os03g26044	22149	CSLA5_-_cellulose_synthase-like_family_A%3B_mannan_synthase_C_expressed
Plant_height	Chr3	14952545	5.39E-06	LOC_Os03g26080	10242	nucleoside-triphosphatase_C_putative_C_expressed
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15740	13907	expressed_protein
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15750	5270	proteasome_Fcyclosome_repeat_containing_protein_C_expressed
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15760	2165	expressed_protein
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15770	-390	CPuORF13_-_conserved_peptide_uORF-containing_transcript_C_expressed
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15775	-2348	expressed_protein
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15780	-4805	expressed_protein
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15790	-9468	ras-related_protein_C_putative_C_expressed
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15810	-18117	bifunctional_protein_folD_C_putative_C_expressed
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15820	-21363	aldose_1-epimerase_C_putative_C_expressed
Plant_height	Chr9	9634061	8.04E-06	LOC_Os09g15800	-13656	tetratricopeptide-like_helical_C_putative_C_expressed
Plant_height	Chr9	9637938	8.04E-06	LOC_Os09g15830	-23835	expressed_protein
Plant_height	Chr9	9640513	8.04E-06	LOC_Os09g15835	-22943	OBP32pep_C_putative_C_expressed
qP	Chr5	18451100	1.43E-06	LOC_Os05g31730	-24423	transporter_C_monovalent_cation%3Aproton_antiporter-2_family_C_putative_C
qP	Chr5	18449522	3.38E-06	LOC_Os05g31650	23103	expressed_protein
qP	Chr5	18449522	3.38E-06	LOC_Os05g31660	20689	expressed_protein
qP	Chr5	18449522	3.38E-06	LOC_Os05g31670	19229	AWPM-19-like_membrane_family_protein_C_putative_C_expressed
qP	Chr5	18449522	3.38E-06	LOC_Os05g31680	11826	expressed_protein
qP	Chr5	18449522	3.38E-06	LOC_Os05g31720	-17441	stromal_membrane-associated_protein_C_putative_C_expressed
qP	Chr5	18449522	3.38E-06	LOC_Os05g31690	0	expressed_protein
qP	Chr5	18449522	3.38E-06	LOC_Os05g31710	-13859	expressed_protein
qP	Chr10	5335554	4.94E-06	LOC_Os10g09820	16919	no_apical_meristem_protein_C_putative_C_expressed
qP	Chr10	5335554	4.94E-06	LOC_Os10g09830	14603	expressed_protein
qP	Chr10	5335554	4.94E-06	LOC_Os10g09850	5583	EF_hand_family_protein_C_putative_C_expressed
qP	Chr10	5335554	4.94E-06	LOC_Os10g09860	0	chalcone_synthase_C_putative_C_expressed
qP	Chr10	5335554	4.94E-06	LOC_Os10g09870	-5412	HVA22_C_putative_C_expressed
SD_ratio	Chr10	9171041	6.31E-06	LOC_Os10g18080	3189	expressed_protein

SD_ratio	Chr10	9171041	6.31E-06	LOC_Os10g18084	823	expressed_protein
SD_ratio	Chr10	9171041	6.31E-06	LOC_Os10g18090	0	expressed_protein
SD_ratio	Chr10	9171041	6.31E-06	LOC_Os10g18070	6219	expressed_protein
SD_ratio	Chr10	9171041	6.31E-06	LOC_Os10g18099	-13584	WRKY18_C_expressed
SD_ratio	Chr1	31665069	9.75E-06	LOC_Os01g55030	24584	DNA-binding_protein-related_C_putative_C_expressed
SD_ratio	Chr1	31665069	9.75E-06	LOC_Os01g55040	16313	1_C3-beta-glucan_synthase_component_C_putative_C_expressed
SD_ratio	Chr1	31665069	9.75E-06	LOC_Os01g55060	5279	expressed_protein
SD_ratio	Chr1	31665069	9.75E-06	LOC_Os01g55070	776	pentatricopeptide_C_putative_C_expressed
SD_ratio	Chr1	31665069	9.75E-06	LOC_Os01g55080	-4659	expressed_protein
SD_ratio	Chr1	31665069	9.75E-06	LOC_Os01g55050	8538	protein_of_unknown_function_DUF1421_domain_containing_protein_C_expres
SD_ratio	Chr1	31665069	9.75E-06	LOC_Os01g55094	-13242	expressed_protein
SD_ratio	Chr10	8518947	9.85E-06	LOC_Os10g16990	9320	expressed_protein
SD_ratio	Chr10	8518947	9.85E-06	LOC_Os10g17000	4389	expressed_protein
SD_ratio	Chr10	8518947	9.85E-06	LOC_Os10g17010	0	expressed_protein
SD_ratio	Chr10	8518947	9.85E-06	LOC_Os10g16974	14618	cytochrome_P450_C_putative_C_expressed
shoot_area	Chr8	5771687	1.84E-06	LOC_Os08g09950	17940	acyl-desaturase_C_chloroplast_precursor_C_putative_C_expressed
shoot_area	Chr8	5771687	1.84E-06	LOC_Os08g09960	16094	expressed_protein
shoot_area	Chr8	5771687	1.84E-06	LOC_Os08g09970	9515	expressed_protein
shoot_area	Chr8	5771687	1.84E-06	LOC_Os08g09990	-2346	expressed_protein
shoot_area	Chr8	5771687	1.84E-06	LOC_Os08g10000	-9119	expressed_protein
shoot_area	Chr8	5771687	1.84E-06	LOC_Os08g10010	-15795	acyl-desaturase_C_chloroplast_precursor_C_putative_C_expressed
shoot_area	Chr8	5771687	1.84E-06	LOC_Os08g09940	19201	GTP-binding_protein_C_putative_C_expressed
shoot_area	Chr10	20738531	3.69E-06	LOC_Os10g38880	22898	DUF623_domain_containing_protein_C_expressed
shoot_area	Chr10	20738531	3.69E-06	LOC_Os10g38890	15575	expressed_protein
shoot_area	Chr10	20738531	3.69E-06	LOC_Os10g38910	7531	expressed_protein
shoot_area	Chr10	20738531	3.69E-06	LOC_Os10g38920	4477	protein_kinase_domain_containing_protein_C_expressed
shoot_area	Chr10	20738531	3.69E-06	LOC_Os10g38930	218	shikimate_Fquinate_5-dehydrogenase_C_putative_C_expressed
shoot_area	Chr10	20738531	3.69E-06	LOC_Os10g38940	-5890	fatty_acid_hydroxylase_C_putative_C_expressed
shoot_area	Chr10	20738531	3.69E-06	LOC_Os10g38954	-21071	hypothetical_protein
shoot_area	Chr10	20738531	3.69E-06	LOC_Os10g38900	7531	erythronate-4-phosphate_dehydrogenase_C_putative_C_expressed
shoot_area	Chr10	20738531	3.69E-06	LOC_Os10g38950	-17558	CGMC_MAPKCMGC_2_ERK.14_-_CGMC_includes_CDA_C_MAPK_C_GSK3_C_
shoot_area	Chr10	11276061	3.80E-06	LOC_Os10g21890	20472	glycosyl_transferase_family_8_C_putative_C_expressed
shoot_area	Chr10	11276061	3.80E-06	LOC_Os10g21910	0	acetyl-CoA_carboxylase_C_putative_C_expressed
shoot_area	Chr10	11276061	3.80E-06	LOC_Os10g21920	-13157	PPR_repeat_containing_protein_C_expressed
shoot_area	Chr10	11276061	3.80E-06	LOC_Os10g21930	-24669	OsFBK21_-_F-box_domain_and_kelch_repeat_containing_protein_C_expressed
shoot_area	Chr10	11282193	3.80E-06	LOC_Os10g21940	-22798	expressed_protein
shoot_area	Chr10	11321265	8.58E-06	LOC_Os10g21970	-21267	expressed_protein
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18880	18021	ubiquitin_carboxyl-terminal_hydrolase_C_family_1_C_putative_C_expressed
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18890	15661	peroxidase_precursor_C_putative_C_expressed
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18900	10360	stigma_specific_peroxidase_precursor_C_putative_C_expressed
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18910	5355	peroxidase_precursor_C_putative_C_expressed
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18920	1935	hAT_dimerisation_domain-containing_protein_C_putative
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18930	-3777	peroxidase_precursor_C_putative_C_expressed
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18940	-5934	expressed_protein
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18950	-15290	peroxidase_precursor_C_putative_C_expressed
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18960	-17907	hypothetical_protein
trmmol	Chr1	10697849	5.41E-06	LOC_Os01g18970	-22570	peroxidase_precursor_C_putative_C_expressed
trmmol	Chr5	14344075	9.26E-06	LOC_Os05g24760	-11121	harpin-induced_protein_1_domain_containing_protein_C_expressed
trmmol	Chr5	14344075	9.26E-06	LOC_Os05g24770	-12005	reticulon_domain_containing_protein_C_putative_C_expressed
trmmol	Chr5	14344075	9.26E-06	LOC_Os05g24780	-16659	OsCML21_-_Calmodulin-related_calcium_sensor_protein_C_expressed
VPD	Chr4	7370371	5.19E-06	LOC_Os04g13300	-21264	hypothetical_protein
VPD	Chr4	7356808	6.37E-06	LOC_Os04g13280	-3832	expressed_protein
VPD	Chr4	7356808	6.37E-06	LOC_Os04g13260	0	expressed_protein
VPD	Chr1	40330703	6.84E-06	LOC_Os01g69830	-1030	OsSPL2_-_SBP-box_gene_family_member_C_expressed
VPD	Chr1	40330703	6.84E-06	LOC_Os01g69840	-6030	expressed_protein
VPD	Chr1	40330703	6.84E-06	LOC_Os01g69850	-13626	OsMADS65_-_MADS-box_family_gene_with_MIKC_A_type-box_C_expressed
VPD	Chr4	11930383	7.59E-06	LOC_Os04g21170	11300	hypothetical_protein
VPD	Chr4	11930383	7.59E-06	LOC_Os04g21160	12328	triacylglycerol_lipase_1_precursor_C_putative_C_expressed
VPD	Chr7	21075730	8.30E-06	LOC_Os07g35150	23779	hypothetical_protein
VPD	Chr7	21075730	8.30E-06	LOC_Os07g35160	20936	expressed_protein

VPD	Chr7	21075730	8.30E-06	LOC_Os07g35180	12762	hypothetical_protein
VPD	Chr7	21075730	8.30E-06	LOC_Os07g35190	2010	expressed_protein
VPD	Chr7	21075730	8.30E-06	LOC_Os07g35200	261	expressed_protein
VPD	Chr7	21075730	8.30E-06	LOC_Os07g35220	-7972	hypothetical_protein
VPD	Chr4	11879396	8.34E-06	LOC_Os04g21120	-8408	expressed_protein
VPD	Chr4	11879396	8.34E-06	LOC_Os04g21110	0	phosphoribulokinase_FUridine_kinase_family_protein_C_expressed
VPD	Chr4	11879396	8.34E-06	LOC_Os04g21130	-15286	F-box_protein_PP2-B1_C_putative_C_expressed
VPD	Chr4	7292815	8.76E-06	LOC_Os04g13190	3757	expressed_protein
VPD	Chr4	7292815	8.76E-06	LOC_Os04g13200	-4522	expressed_protein
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35290	10427	TKL_IRAK_DUF26-lc.10_-_DUF26_kinases_have_homology_to_DUF26_containi
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35300	3192	TKL_IRAK_DUF26-lc.11_-_DUF26_kinases_have_homology_to_DUF26_containi
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35330	-5219	TKL_IRAK_DUF26-lc.13_-_DUF26_kinases_have_homology_to_DUF26_containi
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35335	-9067	expressed_protein
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35340	-10757	TKL_IRAK_DUF26-lc.14_-_DUF26_kinases_have_homology_to_DUF26_containi
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35360	-18363	expressed_protein
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35370	-19628	TKL_IRAK_DUF26-lc.15_-_DUF26_kinases_have_homology_to_DUF26_containi
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35280	14720	TKL_IRAK_DUF26-lc.1_-_DUF26_kinases_have_homology_to_DUF26_containir
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35310	0	TKL_IRAK_DUF26-lc.12_-_DUF26_kinases_have_homology_to_DUF26_containi
VPD	Chr7	21135722	8.89E-06	LOC_Os07g35350	-14289	glucan_endo-1_C3-beta-glucosidase_precursor_C_putative_C_expressed
VPD	Chr4	12011935	9.31E-06	LOC_Os04g21320	-14983	membrane_associated_DUF588_domain_containing_protein_C_putative_C_expre
VPD	Chr4	7323981	9.66E-06	LOC_Os04g13230	-3210	expressed_protein
VPD	Chr4	7323981	9.66E-06	LOC_Os04g13210	11686	multidrug_resistance-associated_protein_C_putative_C_expressed
VPD	Chr4	7323981	9.66E-06	LOC_Os04g13220	0	ABC_transporter_family_protein_C_putative_C_expressed
VPD	Chr10	3748132	9.68E-06	LOC_Os10g07080	17767	expressed_protein
VPD	Chr10	3748132	9.68E-06	LOC_Os10g07114	4222	expressed_protein
VPD	Chr10	3748132	9.68E-06	LOC_Os10g07120	3606	expressed_protein
VPD	Chr10	3748132	9.68E-06	LOC_Os10g07130	1983	expressed_protein
VPD	Chr1	40291506	9.75E-06	LOC_Os01g69279	13344	expressed_protein
VPD	Chr1	40291506	9.75E-06	LOC_Os01g69290	4328	antifreeze_glycoprotein_C_putative_C_expressed
WUE	Chr6	10535084	1.32E-06	LOC_Os06g18080	5130	collagen_adhesion_protein_C_putative_C_expressed
WUE	Chr6	10535084	1.32E-06	LOC_Os06g18120	-14814	expressed_protein
WUE	Chr6	10535084	1.32E-06	LOC_Os06g18140	-18225	UDP-glucoronosyl_and_UDP-glucosyl_transferase_domain_containing_protein_C
WUE	Chr6	10485965	1.90E-06	LOC_Os06g17990	23003	expressed_protein
WUE	Chr6	10485965	1.90E-06	LOC_Os06g18010	5063	UDP-glucoronosyl_and_UDP-glucosyl_transferase_domain_containing_protein_C
WUE	Chr6	10485965	1.90E-06	LOC_Os06g18020	-2483	expressed_protein
WUE	Chr6	10485965	1.90E-06	LOC_Os06g18040	-17626	hypothetical_protein
WUE	Chr6	10485965	1.90E-06	LOC_Os06g18000	14737	protein_kinase_domain_containing_protein_C_expressed
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59020	23966	cytochrome_P450_C_putative_C_expressed
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59030	22856	expressed_protein
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59040	15402	expressed_protein
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59050	7346	cytochrome_P450_C_putative_C_expressed
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59060	5081	50S_ribosomal_protein_C_putative_C_expressed
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59070	3143	expressed_protein
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59080	1018	expressed_protein
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59090	0	thylakoid_lumenal_20_kDa_protein_C_putative_C_expressed
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59110	-5434	indole-3-acetate_beta-glucosyltransferase_C_putative_C_expressed
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59120	-8704	cyclin_C_putative_C_expressed
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59130	-12603	proline-rich_family_protein_C_putative_C_expressed
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59140	-22792	DNA-directed_RNA_polymerases_I_C_II_C_and_III_subunit_RPABC1_C_putativ
WUE	Chr1	34140727	2.69E-06	LOC_Os01g59100	-676	cytokinin-N-glucosyltransferase_1_C_putative_C_expressed
WUE	Chr2	6432419	2.71E-06	LOC_Os02g12330	13041	expressed_protein
WUE	Chr2	6432419	2.71E-06	LOC_Os02g12350	1799	histone_deacetylase_C_putative_C_expressed
WUE	Chr2	6432419	2.71E-06	LOC_Os02g12370	-10712	expressed_protein
WUE	Chr2	6432419	2.71E-06	LOC_Os02g12380	-14659	histone_deacetylase_C_putative_C_expressed
WUE	Chr2	6432419	2.71E-06	LOC_Os02g12360	-801	nuclear_protein_ZAP-related_C_putative_C_expressed
WUE	Chr2	6432419	2.71E-06	LOC_Os02g12340	8415	conserved_hypothetical_protein
WUE	Chr2	6445032	3.35E-06	LOC_Os02g12400	-23535	receptor-like_protein_kinase_precursor_C_putative_C_expressed
WUE	Chr1	34144047	5.81E-06	LOC_Os01g59150	-23886	tubulin_FFtsZ_domain_containing_protein_C_putative_C_expressed
WUE	Chr8	372529	7.39E-06	LOC_Os08g01590	18697	expressed_protein
WUE	Chr8	372529	7.39E-06	LOC_Os08g01610	5965	DUF250_domain_containing_protein_C_putative_C_expressed

WUE	Chr8	372529	7.39E-06	LOC_Os08g01620	1543	shwachman-Bodian-Diamond_syndrome_protein_C_putative_C_expressed
WUE	Chr8	372529	7.39E-06	LOC_Os08g01630	0	DNA_repair_metallo-beta-lactamase_C_putative_C_expressed
WUE	Chr8	372529	7.39E-06	LOC_Os08g01640	-1562	Rf1_C_mitochondrial_precursor_C_putative_C_expressed
WUE	Chr8	372529	7.39E-06	LOC_Os08g01650	-8778	Rf1_C_mitochondrial_precursor_C_putative_C_expressed
WUE	Chr8	372529	7.39E-06	LOC_Os08g01660	-11559	mago_nashi_C_putative_C_expressed
WUE	Chr8	372529	7.39E-06	LOC_Os08g01670	-17133	invertase_Fpectin_methylesterase_inhibitor_family_protein_C_putative_C_expre
WUE	Chr8	372529	7.39E-06	LOC_Os08g01680	-22699	WD_domain_C_G-beta_repeat_domain_containing_protein_C_expressed
WUE	Chr8	372529	7.39E-06	LOC_Os08g01600	11734	polygalacturonase_C_putative_C_expressed
WUE	Chr1	34185829	7.93E-06	LOC_Os01g59200	-15690	DUF617_domain_containing_protein_C_expressed
WUE	Chr1	34168462	8.16E-06	LOC_Os01g59170	-12221	hypothetical_protein
WUE	Chr1	34168462	8.16E-06	LOC_Os01g59180	-13829	OsFBX27_-_F-box_domain_containing_protein_C_expressed
WUE	Chr1	34168462	8.16E-06	LOC_Os01g59160	-7459	UBA_FTS-N_domain_containing_protein_C_expressed
WUE	Chr1	34168462	8.16E-06	LOC_Os01g59190	-24518	expressed_protein
WUE	Chr6	10584075	8.50E-06	LOC_Os06g18150	20216	expressed_protein
WUE	Chr6	10584075	8.50E-06	LOC_Os06g18160	17894	expressed_protein
WUE	Chr6	10584075	8.50E-06	LOC_Os06g18164	2556	expressed_protein
WUE	Chr6	10584075	8.50E-06	LOC_Os06g18670	-4070	anthocyanidin_3-O-glucosyltransferase_C_putative_C_expressed
WUE	Chr6	10584075	8.50E-06	LOC_Os06g18680	-11266	expressed_protein
WUE	Chr6	10584075	8.50E-06	LOC_Os06g18690	-15129	expressed_protein
WUE	Chr6	3933935	8.92E-06	LOC_Os06g08080	9526	inorganic_H_B_pyrophosphatase_C_putative_C_expressed
WUE	Chr6	3933935	8.92E-06	LOC_Os06g08090	5990	heparanase-like_protein_precursor_C_putative_C_expressed
WUE	Chr6	3933935	8.92E-06	LOC_Os06g08110	-92	nodulin_C_putative_C_expressed
WUE	Chr6	3933935	8.92E-06	LOC_Os06g08120	-7752	plant_protein_of_unknown_function_domain_containing_protein_C_expressed
WUE	Chr6	3933935	8.92E-06	LOC_Os06g08130	-12890	expressed_protein
WUE	Chr6	3933935	8.92E-06	LOC_Os06g08154	-22978	receptor-like_protein_kinase_5_precursor_C_putative_C_expressed
WUE	Chr6	3933935	8.92E-06	LOC_Os06g08140	-13991	protein_phosphatase_2C_C_putative_C_expressed
WUE	Chr6	3933935	8.92E-06	LOC_Os06g08100	1133	jp18_C_putative_C_expressed
WUE	Chr1	10674512	9.15E-06	LOC_Os01g18850	17733	OsSPL1_-_SBP-box_gene_family_member_C_expressed
WUE	Chr1	10674512	9.15E-06	LOC_Os01g18860	14375	S-adenosylmethionine_synthetase_C_putative_C_expressed
WUE	Chr1	10674512	9.15E-06	LOC_Os01g18880	-655	ubiquitin_carboxyl-terminal_hydrolase_C_family_1_C_putative_C_expressed
WUE	Chr1	10674512	9.15E-06	LOC_Os01g18890	-4976	peroxidase_precursor_C_putative_C_expressed
WUE	Chr1	10674512	9.15E-06	LOC_Os01g18900	-12681	stigma_specific_peroxidase_precursor_C_putative_C_expressed
WUE	Chr1	10674512	9.15E-06	LOC_Os01g18910	-16886	peroxidase_precursor_C_putative_C_expressed
WUE	Chr1	10674512	9.15E-06	LOC_Os01g18920	-18815	hAT_dimerisation_domain-containing_protein_C_putative
WUE	Chr1	10674512	9.15E-06	LOC_Os01g18870	7609	helix-loop-helix_DNA-binding_domain_containing_protein_C_expressed

REFERENCES

Acevedo-Siaca, L. G., Coe, R., Quick, W. P., & Long, S. P. (2021). Variation between rice accessions in photosynthetic induction in flag leaves and underlying mechanisms. *Journal of Experimental Botany*, *72*(4), 1282–1294. https://doi.org/10.1093/jxb/eraa520

Acevedo-Siaca, L. G., Coe, R., Wang, Y., Kromdijk, J., Quick, W. P., & Long, S. P. (2020). Variation in photosynthetic induction between rice accessions and its potential for improving productivity. *New Phytologist*, *227*(4), 1097–1108. https://doi.org/10.1111/nph.16454

Adams, M. A., Buckley, T. N. and Turbull, T. L. (2020). Diminishing CO_2 driven gains in water use efficiency of global forests. *Nature Climate Change, 10*. 466-471. https://doi.org/10.1038/s41558-020-0747-7

Agnoun, Y., Biaou, S. S. H., Sié, M., Vodouhè, R. S., & Ahanchédé, A. (2012). The African Rice *Oryza glaberrima* Steud: Knowledge Distribution and Prospects. *International Journal of Biology*, *4*(3), p158. https://doi.org/10.5539/ijb.v4n3p158

Ainsworth, E. A., & Long, S. P. (2005). What have we learned from 15 years of free-air CO_2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO_2. *New Phytologist*, *165*(2), 351–372. https://doi.org/10.1111/j.1469-8137.2004.01224.x

Albar, L., Ndjiondjop, M.-N., Esshak, Z., Berger, A., Pinel, A., Jones, M., … Ghesquire, A. (2003). Fine genetic mapping of a gene required for Rice yellow mottle virus cell-to-cell movement. *TAG Theoretical and Applied Genetics*, *107*(2), 371–378. https://doi.org/10.1007/s00122-003-1258-4

Ashraf, M., & Harris, P. J. C. (2013). Photosynthesis under stressful environments: An overview. *Photosynthetica*, *51*(2), 163–190. https://doi.org/10.1007/S11099-013-0021-6

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

Bailey-Serres, J., Lee, S. C., & Brinton, E. (2012). Waterproofing crops: Effective flooding survival strategies. *Plant Physiology*, *160*(4), 1698–1709. https://doi.org/10.1104/PP.112.208173

Bailey-Serres, J., Parker, J. E., Ainsworth, E. A., Oldroyd, G. E. D., & Schroeder, J. I. (2019). Genetic strategies for improving crop yields. *Nature*, *575*. https://doi.org/10.1038/s41586-019-1679-0

Bank, W. (2009). *World Development Report 2010*. The World Bank. https://doi.org/10.1596/978-0-8213-7987-5

Barratt, G. E., Sparkes, D. L., Mcausland, L., & Murchie, E. H. (2020). Anisohydric sugar beet rapidly responds to light to optimize leaf water use efficiency utilizing numerous small stomata. *Annals of Botany*, *13*(1). https://doi.org/10.1093/aobpla/plaa067

Bednarczyk, D., Aviv-Sharon, E., Savidor, A., Levin, Y., & Charuvi, D. (2020). Influence of short-term exposure to high light on photosynthesis and proteins involved in photo-protective processes in tomato leaves. *Environmental and Experimental Botany*, *179*, 104198. https://doi.org/10.1016/J.ENVEXPBOT.2020.104198

Bimpong, I. K., Serraj, R., Chin, J. H., Ramos, J., Mendoza, E. M. T., Hernandez, J. E., … Brar, D. S. (2011). Identification of QTLs for drought-related traits in alien introgression lines derived from crosses of rice (*Oryza sativa* cv. IR64) × O. glaberrima under lowland moisture stress. *Journal of Plant Biology*, *54*(4), 237–250. https://doi.org/10.1007/s12374-011-9161-z

Bita, C. E., & 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*(273). https://doi.org/10.3389/FPLS.2013.00273

Black, R., Allen, L., Bhutta, & Za. (2008). Climate Change, Food Insecurity and Hunger Key Messages for UNFCCC Negotiators. *Lancet*, *371*(9608), 243–260. Retrieved from http://www.careclimatechange.org/files/reports/IASC_CC_FS.pdf

Blum, A. (2009). Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. *Field Crops Research, 112*(2).119-123. https://doi.org/10.1016/j.fcr.2009.03.009

Bocco, R., Lorieux, M., Seck, P. A., Futakuchi, K., Manneh, B., Baimey, H., & Ndjiondjop, M. N. (2012). Agro-morphological characterization of a population of introgression lines derived from crosses between IR 64 (*Oryza sativa* indica) and TOG 5681 (*Oryza glaberrima*) for drought tolerance. *Plant Science*, *183*, 65–76. https://doi.org/10.1016/j.plantsci.2011.09.010

Brown, & Escombe. (1901). VIII. Static diffusion of gases and liquids in relation to the assimilation of carbon and translocation in plants. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, *193*(185–193), 223–291. https://doi.org/10.1098/RSTB.1900.0014

Buckley, C. R., Caine, R. S., & Gray, J. E. (2020). Pores for thought: can genetic manipulation of stomatal density protect future rice yields? *Frontiers in Plant Science*,

10, 1. https://doi.org/10.3389/fpls.2019.01783

Burgess, A. J., Retkute, R., Preston, S. P., Jensen, O. E., Pound, M. P., Pridmore, T. P., & Murchie, E. H. (2016). The 4-dimensional plant: Effects of wind-induced canopy movement on light fluctuations and photosynthesis. *Frontiers in Plant Science*, *7*, 1392. https://doi.org/10.3389/fpls.2016.01392

Caine, R. S., Yin, X., Sloan, J., Harrison, E. L., Mohammed, U., Fulton, T., … Gray, J. E. (2019). Rice with reduced stomatal density conserves water and has improved drought tolerance under future climate conditions. *New Phytologist*, *221*(1), 371–384. https://doi.org/10.1111/nph.15344

Campany, C. E., Tjoelker, M. G., von Caemmerer, S., & Duursma, R. A. (2016). Coupled response of stomatal and mesophyll conductance to light enhances photosynthesis of shade leaves under sunflecks. *Plant Cell and Environment*, *39*(12), 2762–2773. https://doi.org/10.1111/pce.12841

Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M., & Keller, S. R. (2020). Genomic prediction of (mal)adaptation across current and future climatic landscapes, *Annual Review of Ecology, Evolution and Synthesis, 51*, 245–269. https://doi.org/10.1146/ANNUREV-ECOLSYS-020720-042553

Carmo-Silva, E., Andralojc, P. J., Scales, J. C., Driever, S. M., Mead, A., Lawson, T., … Parry, M. A. J. (2017). Phenotyping of field-grown wheat in the UK highlights contribution of light response of photosynthesis and flag leaf longevity to grain yield. *Journal of Experimental Botany*, *68*(13), 3473–3486. https://doi.org/10.1093/JXB/ERX169

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R., & Chhetri, N. (2014). A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*, *4*(4), 287–291. https://doi.org/10.1038/nclimate2153

Charles, H., Godfray, J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., … Toulmin, C. (2010). Food security: The challenge of feeding 9 billion people. *Science*, *12*(5967), 812–818.

Chatterjee, J., Thakur, V., Nepomuceno, R., Coe, R. A., Dionora, J., Elmido-Mabilangan, A., … Quick, W. P. (2020). Natural Diversity in Stomatal Features of Cultivated and Wild Oryza Species. *Rice*, *13*(1), 58. https://doi.org/10.1186/s12284-020-00417-0

Chaves, M. M., Pereira, J. S., Maroco, J., Rodrigues, M. L., Ricardo, C. P. P., Osório, M. L., … Pinheiro, C. (2002). How plants cope with water stress in the field. Photosynthesis and growth. *Annals of Botany*, *89*(SPEC. ISS.), 907–916.

https://doi.org/10.1093/aob/mcf105

Chen, Han, Lai, Z., Shi, J., Xiao, Y., Chen, Z., & Xu, X. (2010). Roles of arabidopsis WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC Plant Biology*, *10*(1), 1–15. https://doi.org/10.1186/1471-2229-10-281/FIGURES/10

Chen, Hongwei, Zhao, X., Zhai, L., Shao, K., Jiang, K., Shen, C., … Xu, J. (2020). Genetic Bases of the Stomata-Related Traits Revealed by a Genome-Wide Association Analysis in Rice (Oryza sativa L.). *Frontiers in Genetics*, *11*, 611. https://doi.org/10.3389/FGENE.2020.00611/BIBTEX

Chen, Q., Wang, B., Ding, H., Zhang, J., & Li, S. (2019). Review: The role of NADP-malic enzyme in plants under stress. *Plant Science, 281,* 216-282. https://doi.org/10.1016/j.plantsci.2019.01.010

Cheng, S. H., Willmann, M. R., Chen, H. C., & Sheen, J. (2002). Calcium Signaling through Protein Kinases. The Arabidopsis Calcium-Dependent Protein Kinase Gene Family. *Plant Physiology*, *129*(2), 469–485. https://doi.org/10.1104/PP.005645

Choquette, N. E., Ogut, F., Wertin, T. M., Montes, C. M., Sorgini, C. A., Morse, A. M., … Ainsworth, E. A. (2019). Uncovering hidden genetic variation in photosynthesis of field-grown maize under ozone pollution. *Global Change Biology*, *25*(12), 4327–4338. https://doi.org/10.1111/gcb.14794

Choudhury, F. K., Rivero, R. M., Blumwald, E., & Mittler, R. (2017). Reactive oxygen species, abiotic stress and stress combination. *The Plant Journal*, *90*(5), 856–867. https://doi.org/10.1111/TPJ.13299

Chowdhury, M. R., Ahamed, M. S., Mas-ud, M. A., Islam, H., Fatamatuzzohora, M., Hossain, M. F., … Matin, M. N. (2021). Stomatal development and genetic expression in *Arabidopsis thaliana* L. *Heliyon*, *7*(8), e07889. https://doi.org/10.1016/J.HELIYON.2021.E07889

Costa, M. V. J. Da, Ramegowda, Y., Ramegowda, V., Karaba, N. N., Sreeman, S. M., & Udayakumar, M. (2021). Combined Drought and Heat Stress in Rice: Responses, Phenotyping and Strategies to Improve Tolerance. *Rice Science*, *28*(3), 233–242. https://doi.org/10.1016/J.RSCI.2021.04.003

Cowling, S. B., Treeintong, P., Ferguson, J., Soltani, H., Swarup, R., Mayes, S., & Murchie, E. H. (2021). Out of Africa: characterising the natural variation in dynamic photosynthetic traits in a diverse population of African rice (*Oryza glaberrima*). *Journal of Experimental Botany,* erab459. https://doi.org/10.1093/JXB/ERAB459

Cubry, P., Pidon, H., Ta, K. N., Tranchant-Dubreuil, C., Thuillet, A. C., Holzinger, M., … Jouannic, S. (2020). Genome Wide Association Study Pinpoints Key Agronomic QTLs in African Rice *Oryza glaberrima*. *Rice*, *13*(1), 66. https://doi.org/10.1186/s12284-020-00424-1

Cubry, P., Tranchant-Dubreuil, C., Thuillet, A.-C., Monat, C., Ndjiondjop, M.-N., Labadie, K., … Vigouroux, Y. (2018). The Rise and Fall of African Rice Cultivation Revealed by Analysis of 246 New Genomes. *Current Biology*, *28*(14), 2274-2282.e6. https://doi.org/10.1016/J.CUB.2018.05.066

De Souza, A. P., Wang, Y., Orr, D. J., Carmo-Silva, E., & Long, S. P. (2020). Photosynthesis across African cassava germplasm is limited by Rubisco and mesophyll conductance at steady state, but by stomatal conductance in fluctuating light. *New Phytologist*, *225*(6), 2498–2512. https://doi.org/10.1111/nph.16142

Deans, R. M., Brodribb, T. J., Busch, F. A., & Farquhar, G. D. (2019). Plant water-use strategy mediates stomatal effects on the light induction of photosynthesis. *New Phytologist*, *222*(1). 385-392. https://doi.org/10.1111/nph.15572

Dittberner, H., Korte, A., Mettler-Altmann, T., Weber, A. P. M., Monroe, G., & de Meaux, J. (2018). Natural variation in stomata size contributes to the local adaptation of water-use efficiency in *Arabidopsis thaliana*. *Molecular Ecology*, *27*(20), 4052–4065. https://doi.org/10.1111/mec.14838

Dow, G. J., Bergmann, D. C., & Berry, J. A. (2014). An integrated model of stomatal development and leaf physiology. *New Phytologist*, *201*(4), 1218–1226. https://doi.org/10.1111/NPH.12608

Draic, P., Flood, J., Harbinson, J., & Aarts, M. G. M. (2011). Natural genetic variation in plant photosynthesis. *Trends in Plant Science*, *16*(6), 327–335. https://doi.org/10.1016/j.tplants.2011.02.005

Drake, P. L., Froend, R. H., & Franks, P. J. (2013a). Smaller, faster stomata: scaling of stomatal size, rate of response, and stomatal conductance. *Journal of Experimental Botany*, *64*(2), 495–505. https://doi.org/10.1093/jxb/ers347

Duartě, K. T. N., De Carvalho, M. A. G., & Martins, P. S. (2017). Segmenting high-quality digital images of stomata using the wavelet spot detection and the watershed transform. *VISIGRAPP 2017 - Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, *4*, 540–547. https://doi.org/10.5220/0006168105400547

Dufey, I., Draye, X., Lutts, S., Lorieux, M., Martinez, C., & Bertin, P. (2015). Novel QTLs in

an interspecific backcross *Oryza sativa* × *Oryza glaberrima* for resistance to iron toxicity in rice. *Euphytica*, *204*(3), 609–625. https://doi.org/10.1007/s10681-014-1342-7

Ellis, E., Maslin, M., Boivin, N., & Bauer, A. (2016). Involve social scientists in defining the Anthropocene. *Nature 2016 540:7632*, *540*(7632), 192–193. https://doi.org/10.1038/540192a

Faralli, M., Matthews, J., & Lawson, T. (2019). Exploiting natural variation and genetic manipulation of stomatal conductance for crop improvement. *Current Opinion in Plant Biology*, *49*, 1–7. https://doi.org/10.1016/j.pbi.2019.01.003

Farquhar, G. D., & Sharkey, T. D. (1982). Stomatal Conductance and Photosynthesis. *Annual Review of Plant Physiology*, *33*(1), 317–345. https://doi.org/10.1146/annurev.pp.33.060182.001533

Fetter, K. C., Eberhardt, S., Barclay, R. S., Wing, S., & Keller, S. R. (2019). StomataCounter: a neural network for automatic stomata identification and counting. *New Phytologist*, *223*(3), 1671–1681. https://doi.org/10.1111/nph.15892

FAO, IFAD, UNICEF, WFP and WHO. 2021. The State of Food Security and Nutrition in the World 2021. Transforming food systems for food security, improved nutrition and affordable healthy diets for all. Rome, FAO. https://doi.org/10.4060/cb4474en

Foyer, C. H., & Noctor, G. (2020). Redox Homeostasis and Signaling in a Higher-CO2 World. *Annual Review of Plant Biology, 5*(*71)*, 157–182. https://doi.org/10.1146/ANNUREV-ARPLANT-050718-095955

Franks, P. J., & Farquhar, G. D. (2007). The mechanical diversity of stomata and its significance in gas-exchange control. *Plant Physiology*, *143*(1), 78–87. https://doi.org/10.1104/pp.106.089367

Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *Molecular Biology and Evolution*, *30*(7), 1687–1699. https://doi.org/10.1093/MOLBEV/MST063

Fujita, M., Fujita, Y., Noutoshi, Y., Takahashi, F., Narusaka, Y., Yamaguchi-Shinozaki, K., & Shinozaki, K. (2006). Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Current Opinion in Plant Biology*, *9*(4), 436–442. https://doi.org/10.1016/J.PBI.2006.05.014

Gaion, L. A., & Carvalho, R. F. (2021). Stomatal response to drought is modulated by gibberellin in tomato. *Acta Physiologiae Plantarum*, *43*(9), 1–6. https://doi.org/10.1007/S11738-021-03286-2/FIGURES/2

Gao, C. (2021). Leading Edge Genome engineering for crop improvement and future agriculture. *Cell*, *184*(6), 1621–1635. https://doi.org/10.1016/j.cell.2021.01.005

Georgieva, K., Sárvári, É., & Keresztes, Á. (2010). Protection of thylakoids against combined light and drought by a lumenal substance in the resurrection plant *Haberlea rhodopensis*. *Annals of Botany*, *105*(1), 117. https://doi.org/10.1093/AOB/MCP274

Ghassemi, F., Jakeman, A. J., & Nix, H. A. (1995). *Salinisation of land and water resources: human causes, extent, management and case studies*. CAB international.

Ghesquière, A., Séquier, J., Second, G., & Lorieux, M. (1997). First steps towards a rational use of African rice, *Oryza glaberrima*, in rice breeding through a "contig line" concept. *Euphytica*, *96*(1), 31–39. https://doi.org/10.1023/A:1003045518236

Giuliani, R., Koteyeva, N., Voznesenskaya, E., Evans, M. A., Cousins, A. B., & Edwards, G. E. (2013). Coordination of Leaf Photosynthesis, Transpiration, and Structural Traits in Rice and Wild Relatives (Genus Oryza). *Plant Physiology, 162*(3), 1632–1651. https://doi.org/10.1104/pp.113.217497

Głowacka, K., Kromdijk, J., Kucera, K., Xie, J., Cavanagh, A. P., Leonelli, L., … Long, S. P. (2018). Photosystem II Subunit S overexpression increases the efficiency of water use in a field-grown crop. *Nature Communications*, *9*(1), 1–9. https://doi.org/10.1038/s41467-018-03231-x

González, D., Fuentes, S., & Serna, L. (2017). Interactions among gibberellins, brassinosteroids and genes regulate stomatal development in the Arabidopsis hypocotyl. *International Journal of Developmental Biology*, *61*(6–7), 383–387. https://doi.org/10.1387/IJDB.170021LS

Griffiths, J., Murase, K., Rieu, I., Zentella, R., Zhang, Z. L., Powers, S. J., … Thomas, S. G. (2006). Genetic Characterization and Functional Analysis of the GID1 Gibberellin Receptors in *Arabidopsis*. *The Plant Cell*, *18*(12), 3399. https://doi.org/10.1105/TPC.106.047415

Hara, K., Yokoo, T., Kajita, R., Onishi, T., Yahata, S., Peterson, K. M., … Kakimoto, T. (2009). Epidermal cell density is autoregulated via a secretory peptide, EPIDERMAL PATTERNING FACTOR 2 in *Arabidopsis* leaves. *Plant and Cell Physiology*, *50*(6), 1019–1031. https://doi.org/10.1093/PCP/PCP068

He, M., He, C. Q., & Ding, N. Z. (2018). Abiotic stresses: General defenses of land plants and chances for engineering multistress tolerance. *Frontiers in Plant Science*, *871*, 1771. https://doi.org/10.3389/fpls.2018.01771

Hetherington, A. M., & Woodward, F. I. (2003). The role of stomata in sensing and driving

environmental change. *Nature*. https://doi.org/10.1038/nature01843

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*(15), 1965–1978. https://doi.org/10.1002/joc.1276

Hippler, M., Minagawa, J., & Takahashi, Y. (2021). Photosynthesis and Chloroplast Regulation—Balancing Photosynthesis and Photoprotection under Changing Environments. *Plant and Cell Physiology*, *62*(7), 1059–1062. https://doi.org/10.1093/PCP/PCAB139

Hong, E. P., & Park, J. W. (2012). Sample Size and Statistical Power Calculation in Genetic Association Studies. *Genomics & Informatics*, *10*(2), 117. https://doi.org/10.5808/GI.2012.10.2.117

Horie, T., Karahara, I., & Katsuhara, M. (2012). Salinity tolerance mechanisms in glycophytes: An overview with the central focus on rice plants. *Rice*, *5*(1), 1–18. https://doi.org/10.1186/1939-8433-5-11/FIGURES/6

Huang, X., Zhao, Y., Wei, X. Li, C., Wang, A., Zhao, Q…Han, B. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nature Genetics, 44***.** 32–39. https://doi.org/10.1038/ng.1018

Huang, X., Zhao, Q., & Han, B. (2015). Comparative Population Genomics Reveals Strong Divergence and Infrequent Introgression between Asian and African Rice. *Molecular Plant*, *8*(6), 958–960. https://doi.org/10.1016/J.MOLP.2015.01.010

Hubbart, S., Smillie, I. R. A., Heatley, M., Swarup, R., Foo, C. C., Zhao, L., & Murchie, E. H. (2018). Enhanced thylakoid photoprotection can increase yield and canopy radiation use efficiency in rice. *Communications Biology*, *1*(1), 1–12. https://doi.org/10.1038/s42003-018-0026-6

Hunt, L., & Gray, J. E. (2009). The Signaling Peptide EPF2 Controls Asymmetric Cell Divisions during Stomatal Development. *Current Biology*, *19*(10), 864–869. https://doi.org/10.1016/J.CUB.2009.03.069

Hyun, M. K., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., & Eskin, E. (2008). Efficient control of population structure in model organism association mapping. *Genetics*, *178*(3), 1709–1723. https://doi.org/10.1534/genetics.107.080101

Iizumi, T., Shiogama, H., Imada, Y., Hanasaki, N., Hiroki Takikawa, & Nishimori, | Motoki. (2018). Crop production losses associated with anthropogenic climate change for 1981-2010 compared with preindustrial levels. *International Journal of Climatology*, *38*(14), 5405-5417. https://doi.org/10.1002/joc.5818

Ikeda, R., Sokei, Y., & Akintayo, I. (2009). Seed fertility of F1 hybrids between upland rice NERICA cultivars. *Breeding Science*, *59*, 27–35.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland.

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 [P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.-O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R.van. Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, M. Belkacemi, J. Malley, (eds.)]. In press

Nadeau, J. A., & Sack, F. D (2002). Stomatal development in Arabidopsis. *The Arabidopsis Book*, *1*, e0066. https://doi.org/10.1199/TAB.0066

Jägermeyr, J., Müller, C., Ruane, A. C., Elliott, J., Balkovic, J., Castillo, O., … Rosenzweig, C. (2021). Climate impacts on global agriculture emerge earlier in new generation of climate and crop models. *Nature Food*, *2*(11), 873–885. https://doi.org/10.1038/s43016-021-00400-y

Jain, M., Tyagi, A. K., & Khurana, J. P. (2006). Genome-wide analysis, evolutionary expansion, and expression of early auxin-responsive SAUR gene family in rice (*Oryza sativa*). *Genomics*, *88*(3), 360–371. https://doi.org/10.1016/J.YGENO.2006.04.008

Jarvit, S., Gollanf, P. J., & Aro, E. M. (2013). Understanding the roles of the thylakoid lumen in photosynthesis regulation. *Frontiers in Plant Science*, *4(*434). https://doi.org/10.3389/FPLS.2013.00434/BIBTEX

Jayakody, H., Liu, S., Whitty, M., & Petrie, P. (2017). Microscope image based fully automated stomata detection and pore measurement method for grapevines. *Plant Methods*, *13*(1), 94. https://doi.org/10.1186/s13007-017-0244-9

Jayakody, H., Petrie, P., Boer, H. J. de, & Whitty, M. (2021). A generalised approach for high-throughput instance segmentation of stomata in microscope images. *Plant Methods*, *17*(1), 27. https://doi.org/10.1186/s13007-021-00727-4

Johnson, M. P. (2016). Photosynthesis. *Essays in Biochemistry*, *60*(3), 255–273. https://doi.org/10.1042/EBC20160016

Jung, H., Lee, D. K., Choi, Y. Do, & Kim, J. K. (2015). OsIAA6, a member of the rice Aux/IAA gene family, is involved in drought tolerance and tiller outgrowth. *Plant Science*, *236*, 304–312. https://doi.org/10.1016/J.PLANTSCI.2015.04.018

K, H., R, K., KU, T., DC, B., & T, K. (2007). The secretory peptide gene EPF1 enforces the stomatal one-cell-spacing rule. *Genes & Development*, *21*(14), 1720–1725. https://doi.org/10.1101/GAD.1550707

Kaiser, E., Morales, A., & Harbinson, J. (2018). Fluctuating light takes crop photosynthesis on a rollercoaster ride. *Plant Physiology, 176*(2). 977-989 . https://doi.org/10.1104/pp.17.01250

Kang, H., Sridhar, V., Mainuddin, M., & Trung, L. D. (2021). Future rice farming threatened by drought in the Lower Mekong Basin. *Scientific Reports*, *11*(1), 9383. https://doi.org/10.1038/S41598-021-88405-2

Kang, Y., Khan, S., & Ma, X. (2009). Climate change impacts on crop yield, crop water productivity and food security - A review. *Progress in Natural Science*, *19*(12). 1665-1674. https://doi.org/10.1016/j.pnsc.2009.08.001

Kijima, Y., Otsuka, K., & Sserunkuuma, D. (2011). An Inquiry into Constraints on a Green Revolution in Sub-Saharan Africa: The Case of NERICA Rice in Uganda. *World Development*, *39*(1), 77–86. https://doi.org/10.1016/j.worlddev.2010.06.010

Kikuta, M., Makihara, D., Arita, N., Miyazaki, A., & Yamamoto, Y. (2017). Growth and yield responses of upland NERICAs to variable water management under field conditions. *Plant Production Science*, *20*(1), 36–46, https://doi.org/10.1080/1343943X.2016.1245102

Kimura, H., Hashimoto-Sugimoto, M., Iba, K., Terashima, I., & Yamori, W. (2020). Improved stomatal opening enhances photosynthetic rate and biomass production in fluctuating light. *Journal of Experimental Botany*, *71*(7), 2339–2350, http://www.ncbi.nlm.nih.gov/pubmed/32095822

Kissoudis, C., Van De Wiel, C., Gf Visser, R., & Van Der Linden, G. (2016). Future-proof crops: challenges and strategies for climate resilience improvement. *Current Opinion in Plant Biology*, *30*, 47–56. https://doi.org/10.1016/j.pbi.2016.01.005

Korte, A. and Farlow, A. (2013). The advantages and limitations of traits analysis with GWAS: a review. Plant Methods, 9(29). 1-9. https://doi.org/ 10.1186/1746-4811-9-29

Kostaki, K. I., Coupel-Ledru, A., Bonnell, V. C., Gustavsson, M., Sun, P., McLaughlin, F. J., … Franklin, K. A. (2020). Guard cells integrate light and temperature signals to control stomatal aperture. *Plant Physiology*, *182*(3), 1404–1419. https://doi.org/10.1104/PP.19.01528

Kromdijk, J., Głowacka, K., Leonelli, L., Gabilly, S. T., Iwai, M., Niyogi, K. K., & Long, S. P. (2016). Improving photosynthesis and crop productivity by accelerating recovery

from photoprotection. *Science*, *354*(6314), 857–861. https://doi.org/10.1126/science.aai8878

Kumar Mohanta, T., Tufail Bashir, ·, Hashem, · Abeer, Elsayed, ·, Abd_Allah, F., Abdul, ·, … Al-Harrasi, S. (2018). Early Events in Plant Abiotic Stress Signaling: Interplay Between Calcium, Reactive Oxygen Species and Phytohormones. *Journal of Plant Growth Regulation*, *37*, 1033–1049. https://doi.org/10.1007/s00344-018-9833-8\

Mohanta, T. K., Bashir, T., Hashem, A., Abd_Allah, E. F., Khan, A. L., & Al-Harrasi, A. S. (2018). Early events in plant abiotic stress signaling: interplay between calcium, reactive oxygen species and phytohormones. *Journal of Plant Growth Regulation*, *37*(4), 1033-1049.

Laga, H., Shahinnia, F., & Fleury, D. (2014). Image-based plant stornata phenotyping. *2014 13th International Conference on Control Automation Robotics and Vision, ICARCV 2014*, 217–222. https://doi.org/10.1109/ICARCV.2014.7064307

Lawson, T., & Blatt, M. R. (2014). Stomatal Size, Speed, and Responsiveness Impact on Photosynthesis and Water Use Efficiency. *Plant Physiology, 164*(4). 1556-1570. https://doi.org/10.1104/pp.114.237107

Lawson, T., & Blatt, M. R. (2014). Topical Review on Stomata and Water Use Efficiency Stomatal Size, Speed, and Responsiveness Impact on Photosynthesis and Water Use Efficiency. *Plant Physiology*, *164*, 1556–1570. https://doi.org/10.1104/pp.114.237107

Lawson, T., & Vialet-Chabrand, S. (2019). Speedy stomata, photosynthesis and plant water use efficiency. *New Phytologist*, *221*(1). 93-98. https://doi.org/10.1111/nph.15330

Lawson, T., von Caemmerer, S., & Baroli, I. (2010). Photosynthesis and stomatal behaviour. *Progress in Botany, 72* (pp. 265-304). Springer, Berlin, Heidelberg.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 1–18. https://doi.org/10.18637/jss.v025.i01

Lewis, S. L., & Maslin, M. A. (2015). Defining the Anthropocene. *Nature. 2015, 519:7542*, *519*(7542), 171–180. https://doi.org/10.1038/nature14258

Li, J., Zhou, J., Zhang, Y., Yang, Y., Pu, Q., & Tao, D. (2020). New Insights Into the Nature of Interspecific Hybrid Sterility in Rice. *Frontiers in Plant Science*, *11*, 1445. https://doi.org/10.3389/fpls.2020.555572

Li, L., Aro, E. M., & Millar, A. H. (2018). Mechanisms of Photodamage and Protein Turnover in Photoinhibition. *Trends in Plant Science*, *23*(8), 667–676. https://doi.org/10.1016/J.TPLANTS.2018.05.004

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., … Zitnick, C. L.

(2014). Microsoft COCO: Common Objects in Context. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8693 LNCS*(PART 5), 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

Lin, C. and Sauter M. (2018) Control of Adventitious Root Architecture in Rice by Darkness, Light, and Gravity. *Plant Physiology, 176*(2):1352-1364. https://doi.org/10.1104/pp.17.01540.

Linares, O. F. (2002). African rice (*Oryza glaberrima*): history and future potential. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(25), 16360–16365. https://doi.org/10.1073/pnas.252604599

Lintala, M., Lehtimäki, N., Benz, J. P., Jungfer, A., Soll, J., Aro, E. M., … Mulo, P. (2012). Depletion of leaf-type ferredoxin-NADP + oxidoreductase results in the permanent induction of photoprotective mechanisms in Arabidopsis chloroplasts. *Plant Journal*, *70*(5), 809–817. https://doi.org/10.1111/J.1365-313X.2012.04930.X

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., … Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397–2399. https://doi.org/10.1093/BIOINFORMATICS/BTS444

Liu, J., & Last, R. L. (2017). A chloroplast thylakoid lumen protein is required for proper photosynthetic acclimation of plants under fluctuating light environments. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(38), E8110–E8117. https://doi.org/10.1073/PNAS.1712206114/-/DCSUPPLEMENTAL

Liu, X., Huang, M., Fan, B., Buckler, E. S., & Zhang, Z. (2016). Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. *PLOS Genetics*, *12*(2), e1005767. https://doi.org/10.1371/JOURNAL.PGEN.1005767

Qu, M., Hamdani, S., Li, W., Wang, S., Tang, J., Chen, Z., ... & Zhu, X. (2016). Rapid stomatal response to fluctuating light: an under-explored mechanism to improve drought tolerance in rice. *Functional Plant Biology*, *43*(8), 727-738.

McAdam, S. A., Chater, C. C., Shpak, E. D., Raissig, M. T., & Dow, G. J. (2021). Linking Stomatal Development and Physiology: From Stomatal Models to Non-model Species and Crops. *Frontiers in Plant Science*, *12*. 1-5. https://doi.org/10.3389/fpls.2021.743964

McAusland, L., & Murchie, E. H. (2020). Start me up; harnessing natural variation in photosynthetic induction to improve crop yields. *New Phytologist*, *227*(4), 989–991. https://doi.org/10.1111/nph.16634

McAusland, L., Vialet-Chabrand, S., Davey, P., Baker, N. R., Brendel, O., & Lawson, T.

(2016). Effects of kinetics of light-induced stomatal responses on photosynthesis and water-use efficiency. *New Phytologist*, *211*(4), 1209–1220. https://doi.org/10.1111/nph.14000

McKown, A. D., Guy, R. D., Quamme, L., Klápště, J., La Mantia, J., Constabel, C. P., … Azam, M. S. (2014). Association genetics, geography and ecophysiology link stomatal patterning in *Populus trichocarpa* with carbon gain and disease resistance trade-offs. *Molecular Ecology*, *23*(23), 5771–5790. https://doi.org/10.1111/mec.12969

Medlyn, B. E., Kauwe, M. G., Lin, Y., Knauer, J., Durrsma, R. A., Williams, C. A.,…Wingate, L. (2017). How do leaf and ecosystem measures of water use efficiency compare? *New Phytologist, 216*(2017). https://doi/org/10.111.nph.14626

Medrano, H., Tomás, M., Martorell, S., Flexas, J., Hernández, E.…Bota, J. (2015). From leaf to whole-plant water use efficiency (WUE) in complex canopies: Limitations of leaf WUE as a selection target. The Crop Journal, 3(3). 200-228. https://doi.org/10.1016/j.cj.2015.04.002

Melotto, M., Underwood, W., & He, S. Y. (2008). Role of Stomata in Plant Innate Immunity and Foliar Bacterial Diseases. *Annual Review of Phytopathology*, *46*, 101–122. https://doi.org/10.1146/annurev.phyto.121107.104959

Merk, H. L., Yarnes, S. C., Deynze, A. Van, Tong, N., Menda, N., Mueller, L. A., … Francis, D. M. (2012). Trait Diversity and Potential for Selection Indices Based on Variation Among Regionally Adapted Processing Tomato Germplasm. *Journal of the American Society for Horticultural Science*, *137*(6), 427–437. https://doi.org/10.21273/JASHS.137.6.427

Meyer, R. S., Choi, J. Y., Sanches, M., Plessis, A., Flowers, J. M., Amas, J., … Purugganan, M. D. (2016). Domestication history and geographical adaptation inferred from a SNP map of African rice. *Nature Genetics 2016 48:9*, *48*(9), 1083–1088. https://doi.org/10.1038/NG.3633

Meyer, R. S., & Purugganan, M. D. (2013). Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics 2013 14:12*, *14*(12), 840–852. https://doi.org/10.1038/nrg3605

Mir, R. R., Reynolds, M., Pinto, F., Khan, M. A., & Bhat, M. A. (2019). High-throughput phenotyping for crop improvement in the genomics era. *Plant Science*, *282*, 60–72. https://doi.org/10.1016/J.PLANTSCI.2019.01.007

Mohammed, U., Caine, R. S., Atkinson, J. A., Harrison, E. L., Wells, D., Chater, C. C., … Murchie, E. H. (2019). Rice plants overexpressing OsEPF1 show reduced stomatal

density and increased root cortical aerenchyma formation. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-41922-7

Monat, C., Pera, B., Ndjiondjop, M.-N., Sow, M., Tranchant-Dubreuil, C., Bastianelli, L., … Sabot, F. (2017). De Novo Assemblies of Three *Oryza glaberrima* Accessions Provide First Insights about Pan-Genome of African Rices. *Genome Biology and Evolution*, *9*(1), 1–6. https://doi.org/10.1093/gbe/evw253

Moore, C. E., Meacham-Hensold, K., Lemonnier, P., Slattery, R. A., Benjamin, C., Bernacchi, C. J., … Cavanagh, A. P. (2021). The effect of increasing temperature on crop photosynthesis: From enzymes to ecosystems. *Journal of Experimental Botany*, *72*(8), 2822–2844. https://doi.org/10.1093/jxb/erab090

Muir, C. D. (2020). A Stomatal Model of Anatomical Tradeoffs Between Gas Exchange and Pathogen Colonization. *Frontiers in Plant Science*, *11*, 518991. https://doi.org/10.3389/fpls.2020.518991

Müller, P., Li, X. P., & Niyogi, K. K. (2001). Non-Photochemical Quenching. A Response to Excess Light Energy. *Plant Physiology*, *125*(4), 1558–1566. https://doi.org/10.1104/PP.125.4.1558

Murchie, E H, & Lawson, T. (2013). Chlorophyll fluorescence analysis: a guide to good practice and understanding some new applications. *Journal of Experimental Botany*, *64*(13), 3983–3998. https://doi.org/10.1093/jxb/ert208

Murchie, E. H., Ali, A., & Herman, T. (2015). Photoprotection as a Trait for Rice Yield Improvement: Status and Prospects. *Rice*, *8*(1). https://doi.org/10.1186/s12284-015-0065-2

Murchie, E. H., Kefauver, S., Araus, J. L., Muller, O., Rascher, U., Flood, P. J., & Lawson, T. (2018). Measuring the dynamic photosynthome. *Annals of Botany*, *122*(2). 207-220. https://doi.org/10.1093/aob/mcy087

Murchie, E. H., & Ruban, A. V. (2020). Dynamic non-photochemical quenching in plants: from molecular mechanism to productivity. *Plant Journal*, *101*(4), 885–896. https://doi.org/10.1111/tpj.14601

Mustroph, A. (2018). Improving flooding tolerance of crop plants. *Agronomy, 8*(9). 160- 185. https://doi.org/10.3390/agronomy8090160

Muthayya, S., Sugimoto, J. D., Montgomery, S., & Maberly, G. F. (2014). An overview of global rice production, supply, trade, and consumption. *Annals of the New York Academy of Sciences*, *1324*(1), 7–14. https://doi.org/10.1111/nyas.12540

Nabholz, B., Sarah, G., Sabot, F., Ruiz, M., Adam, H., Nidelet, S., … Glémin, S. (2014).

Transcriptome population genomics reveals severe bottleneck and domestication cost in the African rice (*Oryza glaberrima*). *Molecular Ecology*, *23*(9), 2210–2227. https://doi.org/10.1111/mec.12738

Nabi, R. B. S., Tayade, R., Hussain, A., Adhikari, A., Lee, I. J., Loake, G. J., & Yun, B. W. (2021). A novel duf569 gene is a positive regulator of the drought stress response in arabidopsis. *International Journal of Molecular Sciences*, *22*(10), 5316. https://doi.org/10.3390/IJMS22105316/S1

Nassirou, T. Y., & He, Y. (2011). *Molecular Plant Breeding. Molecular Plant Breeding* (Vol. 2). Retrieved from http://biopublisher.ca/index.php/mpb/article/view/163/250

Ohsumi, A., Kanemura, T., Homma, K., Horie, T., & Shiraiwa, T. (2007). Genotypic variation of stomatal conductance in relation to stomatal density and length in rice (*Oryza sativa* L.). *Plant Production Science*, *10*(3), 322–328. https://doi.org/10.1626/pps.10.322

Oort, P. A. J. (2018). Mapping abiotic stresses for rice in Africa: Drought, cold, iron toxicity, salinity and sodicity. *Field Crops Research, 219*(2018). 55-75. https://doi.org/10.1016/j.fcr.2018.01.016

Orjuela, J., Sabot, F., Chéron, S., Vigouroux, Y., Adam, H., Chrestin, H., … Ghesquière, A. (2014). An extensive analysis of the African rice genetic diversity through a global genotyping. *Theoretical and Applied Genetics*, *127*(10), 2211–2223. https://doi.org/10.1007/s00122-014-2374-z

Orth, A., Wilson, E. R., Thompson, J. G., & Gibson, B. C. (2018). A dual-mode mobile phone microscope using the onboard camera flash and ambient light. *Scientific Reports 2018 8:1*, *8*(1), 1–8. https://doi.org/10.1038/s41598-018-21543-2

Outlaw, W. H. (2003). Integration of cellular and physiological functions of guard cells. *Critical Reviews in Plant Sciences*, *22*(6), 503–529. https://doi.org/10.1080/713608316

Panda, D., & Barik, J. (2021). Flooding Tolerance in Rice: Focus on Mechanisms and Approaches. *Rice Science*, *28*(1), 43–57. https://doi.org/10.1016/J.RSCI.2020.11.006

Panda, D., Mishra, S. S., & Behera, P. K. (2021). Drought Tolerance in Rice: Focus on Recent Mechanisms and Approaches. *Rice Science*, *28*(2), 119–132. https://doi.org/10.1016/J.RSCI.2021.01.002

Pandey, V., & Shukla, A. (2015). Acclimation and Tolerance Strategies of Rice under Drought Stress. *Rice Science,* 22(4). 147-161. https://doi.org/10.1016/j.rsci.2015.04.001

Panetta, A. M., Stanton, M. L., & Harte, J. (2018). Climate warming drives local extinction: Evidence from observation and experimentation. *Science Advances*, *4*(2), eaaq1819–

eaaq1819. https://doi.org/10.1126/sciadv.aaq1819

Parry, M., Rosenzweig, C., & Livermore, M. (2005). Climate change, global food supply and risk of hunger. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1463), 2125-2138.

Pidon, H., Chéron, S., Ghesquière, A., & Albar, L. (2020). Allele mining unlocks the identification of RYMV resistance genes and alleles in African cultivated rice. *BMC Plant Biology*, *20*(222), 1-14. https://doi.org/10.1186/s12870-020-02433-0

Purugganan, M. D. (2014). An evolutionary genomic tale of two rice species. *Nature Genetics*, *46*(9). 931-932. https://doi.org/10.1038/ng.3071

Qu, M., Essemine, J., Xu, J., Ablat, G., Perveen, S., Wang, H., … Zhu, X. (2020). Alterations in stomatal response to fluctuating light increase biomass and yield of rice under drought conditions. *Plant Journal*, *104*(5), 1334–1347. https://doi.org/10.1111/tpj.15004

Qu, M., Zheng, G., Hamdani, S., Essemine, J., Song, Q., Wang, H., … Zhu, X. G. (2017). Leaf photosynthetic parameters related to biomass accumulation in a global rice diversity survey. *Plant Physiology*, *175*(1), 248–258. https://doi.org/10.1104/pp.17.00332

Ramírez, J. I. S., & Maiti, R. (2016). Research Trends in Abiotic Stress Resistance of Crops. *Bioresource and Stress Management*, 131–163. https://doi.org/10.1007/978-981-10-0995-2_9

Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., & Chatterjee, S. (2019). Climate change has likely already affected global food production. *Plos One, 14*(5), e0217148. https://doi.org/10.1371/journal.pone.0217148

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

Ritz, C., Baty, F., Streibig, J. C., & Gerhard, D. (2015). Dose-Response Analysis Using R. *Plos One*, *10*(12), e0146021. https://doi.org/10.1371/journal.pone.0146021

Robinson, G. K. (1991). That BLUP is a good thing: the estimation of random effects. *Statistical Science*, *6*(1). 15-32.

Román-Palacios, C., & Wiens, J. J. (2020). Recent responses to climate change reveal the drivers of species extinction and survival. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(8), 4211–4217. https://doi.org/10.1073/PNAS.1913007117/-/DCSUPPLEMENTAL

Ruban, A. V., & Wilson, S. (2021). The mechanism of non-photochemical quenching in

plants: Localization and driving forces. *Plant and Cell Physiology*, *62*(7), 1063–1072. https://doi.org/10.1093/pcp/pcaa155

Sarla, N., & Mallikarjuna Swamy, B. P. (2005). *Oryza glaberrima*: a source for the improvment of *Oryza sativa*. *Current Science*, *89*(6), 955–963. https://doi.org/10.2307/24110748

Schröder, R., Atkinson, R. G., & Redgwell, R. J. (2009). Re-interpreting the role of endo-beta-mannanases as mannan endotransglycosylase/hydrolases in the plant cell wall. *Annals of Botany*, *104*(2), 197–204. https://doi.org/10.1093/AOB/MCP120

Seck, P. A., Diagne, A., Mohanty, S., & Wopereis, M. C. S. (2012). Crops that feed the world 7: Rice. *Food Security*, *4*(1), 7–24. https://doi.org/10.1007/S12571-012-0168-1

Sempéré, G., Philippe, F., Dereeper, A., Ruiz, M., Sarah, G., & Larmande, P. (2016). Gigwa—Genotype investigator for genome-wide analyses. *GigaScience*, *5*(1). https://doi.org/10.1186/S13742-016-0131-8

Shen, Y., Zhao, Z., Ma, H., Bian, X., Yu, Y., Yu, X., … Wan, J. (2015). Fine mapping of S37, a locus responsible for pollen and embryo sac sterility in hybrids between *Oryza sativa* L. and *O. glaberrima* Steud. *Plant Cell Reports*, *34*(11), 1885–1897. https://doi.org/10.1007/s00299-015-1835-4

Singh, R. K., Kota, S., & Flowers, T. J. (2021). Salt tolerance in rice: seedling and reproductive stage QTL mapping come of age. *Theoretical and Applied Genetics*, *134*(11), 3495. https://doi.org/10.1007/s00122-021-03890-3

Singh, R., Singh, Y., Xalaxo, S., Verulkar, S., Yadav, N., Singh, S., … Singh, N. K. (2016). From QTL to variety-harnessing the benefits of QTLs for drought, flood and salt tolerance in mega rice varieties of India through a multi-institutional network. *Plant Science*, *242*, 278–287. https://doi.org/10.1016/j.plantsci.2015.08.008

Sjögren, L. L. E., Stanne, T. M., Zheng, B., Sutinen, S., & Clarke, A. K. (2006). Structural and Functional Insights into the Chloroplast ATP-Dependent Clp Protease in Arabidopsis. *The Plant Cell*, *18*(10), 2635. https://doi.org/10.1105/TPC.106.044594

Slattery, R. A., Walker, B. J., Weber, A. P. M., & Ort, D. R. (2018). The impacts of fluctuating light on crop performance. *Plant Physiology*, *176*(2), 990–1003. https://doi.org/10.1104/pp.17.01234

Song, G., Jia, M., Chen, K., Kong, X., Khattak, B., Xie, C., ... & Mao, L. (2016). CRISPR/Cas9: a powerful tool for crop genome editing. *The Crop Journal*, *4*(2), 75-82.

Sun, J., Ye, M., Peng, S., & Li, Y. (2016). Nitrogen can improve the rapid response of photosynthesis to changing irradiance in rice (Oryza sativa L.) plants. *Scientific*

Reports, *6*(1), 1-10.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-December*, 2818–2826. https://doi.org/10.1109/CVPR.2016.308

Tak, H., Negi, S., & Ganapathi, T. R. (2021). The 5′-upstream region of WRKY18 transcription factor from banana is a stress-inducible promoter with strong expression in guard cells. *Physiologia Plantarum*, *173*(4), 1335–1350. https://doi.org/10.1111/PPL.13326

Takezaki, N., & Nei, M. (1996). Genetic Distances and Reconstruction of Phylogenetic Trees From Microsatellite DNA. *Genetics*, *144*(1). 389-399.

Tam, V., Patel, N., Turcotte, M., Bossé, Paré, G. and Meyre, D. (2019). Benefits and limitations of genome wide association studies. *Nature Reviews Genetics, 20*(2019). 467-484. https://doi.org/10.1038/s41576-019-0127-1

Tanaka, Y., Sugano, S. S., Shimada, T., & Hara-Nishimura, I. (2013). Enhancement of leaf photosynthetic capacity through increased stomatal density in Arabidopsis. *New Phytologist*, *198*(3), 757–764. https://doi.org/10.1111/nph.12186

Taylor, S. H., & Long, S. P. (2017). Slow induction of photosynthesis on shade to sun transitions in wheat may cost at least 21% of productivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1730). https://doi.org/10.1098/rstb.2016.0543

Phung, N. T. P., Mai, C. D., Hoang, G. T., Truong, H. T. M., Lavarenne, J., Gonin, M., ... & Courtois, B. (2016). Genome-wide association mapping for root traits in a panel of rice accessions from Vietnam. *BMC Plant Biology*, *16*(1), 1-19.

Thiémélé, D., Boisnard, A., Ndjiondjop, M. N., Chéron, S., Séré, Y., Aké, S., … Albar, L. (2010). Identification of a second major resistance gene to Rice yellow mottle virus, RYMV2, in the African cultivated rice species, *O. glaberrima*. *Theoretical and Applied Genetics*, *121*(1), 169–179. https://doi.org/10.1007/S00122-010-1300-2/FIGURES/3

Tinoco-Ojanguren, C., & Pearcy, R. W. (1993). Stomatal dynamics and its importance to carbon gain in two rainforest Piper species - I. VPD effects on the transient stomatal response to lightflecks. *Oecologia*, *94*(3), 388–394. https://doi.org/10.1007/BF00317114

Todesco, M., Balasubramanian, S., Hu, T.T., Traw, M.B., Horton, M.,…Weigel, D. (2010). Natural allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature*, *465*(2010). 632-636. https://doi.org/10.1038/nature09083.

Tuberosa, R., & Salvi, S. (2006). Genomics-based approaches to improve drought tolerance of crops. *Trends in Plant Science*, *11*(8), 405–412. https://doi.org/10.1016/j.tplants.2006.06.003

Ueguchi-Tanaka, M., Ashikari, M., Nakajima, M., Itoh, H., Katoh, E., Kobayashi, M., … Matsuoka, M. (2005). GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature*, *437*(7059), 693–698. https://doi.org/10.1038/NATURE04028

Uga, Y., Sugimoto, K., Ogawa, S., Rane, J., Ishitani, M., Hara, N., … Yano, M. (2013). Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. *Nature Genetics*, *45*(9). https://doi.org/10.1038/ng.2725

Verslues, P. E., & Zhu, J. K. (2007). New developments in abscisic acid perception and metabolism. *Current Opinion in Plant Biology*, *10*(5), 447–452. https://doi.org/10.1016/J.PBI.2007.08.004

Vialet-Chabrand, S., Matthews, J. S. A., Simkin, A. J., Raines, C. A., & Lawson, T. (2017). Importance of fluctuations in light on plant photosynthetic acclimation. *Plant Physiology*, *173*(4), 2163–2179. https://doi.org/10.1104/pp.16.01767

Vialet-Chabrand, S. R. M., Matthews, J. S. A., McAusland, L., Blatt, M. R., Griffiths, H., & Lawson, T. (2017). Temporal dynamics of stomatal behavior: Modeling and implications for photosynthesis and water use. *Plant Physiology*, *174*(2), 603–613. https://doi.org/10.1104/pp.17.00125

Wahid, A., Gelani, S., Ashraf, M., & Foolad, M. R. (2007). Heat tolerance in plants: An overview. *Environmental and Experimental Botany*, *61*, 199–223. https://doi.org/10.1016/j.envexpbot.2007.05.011

Wambugu, P. W., Ndjiondjop, M. N., & Henry, R. (2019). Advances in molecular genetics and genomics of african rice (*Oryza glaberrima* steud). *Plants*, *8*(10). 376-393. https://doi.org/10.3390/plants8100376

Wang, H., Xu, X., Garrett Vieira, F., Xiao, Y., Li, Z., Wang, J., … Chu, C. (2016). The Power of Inbreeding: NGS-Based GWAS of Rice Reveals Convergent Evolution during Rice Domestication. *Molecular Plant*, *9*, 975–985. https://doi.org/10.1016/j.molp.2016.04.018

Wang, M., Yu, Y., Haberer, G., Marri, P. R., Fan, C., Goicoechea, J. L., … Wing, R. A. (2014). The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. *Nature Genetics*, *46*(9), 982–988.

https://doi.org/10.1038/ng.3044

Wang, P., Li, L., Wei, H., Sun, W., Zhou, P., Zhu, S., … Zhuge, Q. (2021). Genome-wide and comprehensive analysis of the multiple stress-related caf1 (Ccr4-associated factor 1) family and its expression in poplar. *Plants*, *10*(5), 981. https://doi.org/10.3390/PLANTS10050981/S1

Wang, X., Zhao, C., Müller, C., Wang, C., Ciais, P., Janssens, I., ... & Piao, S. (2020). Emergent constraint on crop yield response to warmer temperature from field experiments. *Nature Sustainability*, *3*(11), 908-916.

Wang, Y., Cai, S., Yin, L., Shi, K., Xia, X., Zhou, Y., … Zhou, J. (2015). Tomato HsfA1a plays a critical role in plant drought tolerance by activating ATG genes and inducing autophagy. *Autophagy*, *11*(11), 2033. https://doi.org/10.1080/15548627.2015.1098798

Wolfe, M. D., & Tonsor, S. J. (2014). Adaptation to spring heat and drought in northeastern Spanish *Arabidopsis thaliana*. *New Phytologist*, *201*(1), 323–334. https://doi.org/10.1111/nph.12485

Xu, P., Zhou, J., Li, J., Hu, F., Deng, X., Feng, S., … Tao, D. (2014). Mapping three new interspecific hybrid sterile loci between *Oryza sativa* and *O. glaberrima*. *Breeding Science*, *63*(5), 476–482. https://doi.org/10.1270/jsbbs.63.476

Xu, Z., Jiang, Y., Jia, B., & Zhou, G. (2016). Elevated-CO_2 response of stomata and its dependence on environmental factors. *Frontiers in Plant Science*, *7*(657), 657. https://doi.org/10.3389/FPLS.2016.00657/BIBTEX

Y, K., M, de la B., JP, H., H, K., WR, M., S, O., … T, M. (2013). Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice*, *6*(1), 4–4. https://doi.org/10.1186/1939-8433-6-4

Yang, J., Li, C., Kong, D., Guo, F., & Wei, H. (2020, December 4). Light-Mediated Signaling and Metabolic Changes Coordinate Stomatal Opening and Closure. *Frontiers in Plant Science*, *11*. 601478. https://doi.org/10.3389/fpls.2020.601478

Yerlikaya, B. A., Ömezli, S., & Aydoğan, N. (2020). Climate Change Forecasting and Modeling for the Year of 2050. *Environment, Climate, Plant and Vegetation Growth*, 109–122. https://doi.org/10.1007/978-3-030-49732-3_5

York, L. M., Slack, S., Bennett, M. J., & Foulkes, M. J. (2018). Wheat shovelomics I: A field phenotyping approach for characterising the structure and function of root systems in tillering species. *BioRxiv*. https://doi.org/10.1101/280875

Yoshida, K., Terashima, I., & Noguchi, K. (2006). Distinct Roles of the Cytochrome Pathway and Alternative Oxidase in Leaf Photosynthesis. *Plant and Cell Physiology*,

47(1), 22–31. https://doi.org/10.1093/PCP/PCI219

Yuan, J. S., Yang, X., Lai, J., Lin, H., Cheng, Z. M., Nonogaki, H., & Chen, F. (2007). The endo-beta-mannanase gene families in Arabidopsis, rice, and poplar. *Functional & Integrative Genomics*, *7*(1), 1–16. https://doi.org/10.1007/S10142-006-0034-3

Zaidi, P. H., Seetharam, K., Krishna, G., Krishnamurthy, L., Gajanan, S., Babu, R., … Doreen, H. W. E. (2016). Genomic Regions Associated with Root Traits under Drought Stress in Tropical Maize (*Zea mays* L.). *Plos One*, *11*(10), e0164340. https://doi.org/10.1371/journal.pone.0164340

Zendonadi dos Santos, N., Piepho, H. P., Condorelli, G. E., Licieri Groli, E., Newcomb, M., Ward, R., ... & Muller, O. (2021). High-throughput field phenotyping reveals genetic variation in photosynthetic traits in durum wheat under drought. *Plant, Cell & Environment*, *44*(9), 2858-2878. https://doi.org/10.1111/pce.14136

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., … Asseng, S. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(35), 9326–9331. https://doi.org/10.1073/pnas.1701762114

Zhao, C., Piao, S., Wang, X., Huang, Y., Ciais, P., Elliott, J., … Peñuelas, J. (2016). Plausible rice yield losses under future climate warming. *Nature Plants*, *3*(1), 1–5. https://doi.org/10.1038/nplants.2016.202

Zhou, L., Hao, Y., Lu, G., Wang, P., Guo, H., & Cheng, H. (2020). Cloning and functional analysis of AmDUF1517 promoter from *Ammopiptanthus mongolicus*. *Journal of Bioscience and Bioengineering*, *130*(3), 233–238. https://doi.org/10.1016/J.JBIOSC.2020.04.014

Zhu, H., & Zhou, X. (2020). Statistical methods for SNP heritability estimation and partition: A review. *Computational and Structural Biotechnology Journal*, *18*, 1557-1568. https://doi.org/10.1016/j.csbj.2020.06.011

Zoulias, N., Harrison, E. L., Casson, S. A., & Gray, J. E. (2018). Molecular control of stomatal development. *Biochemical Journal*, *475*(2), 441. https://doi.org/10.1042/BCJ20170413