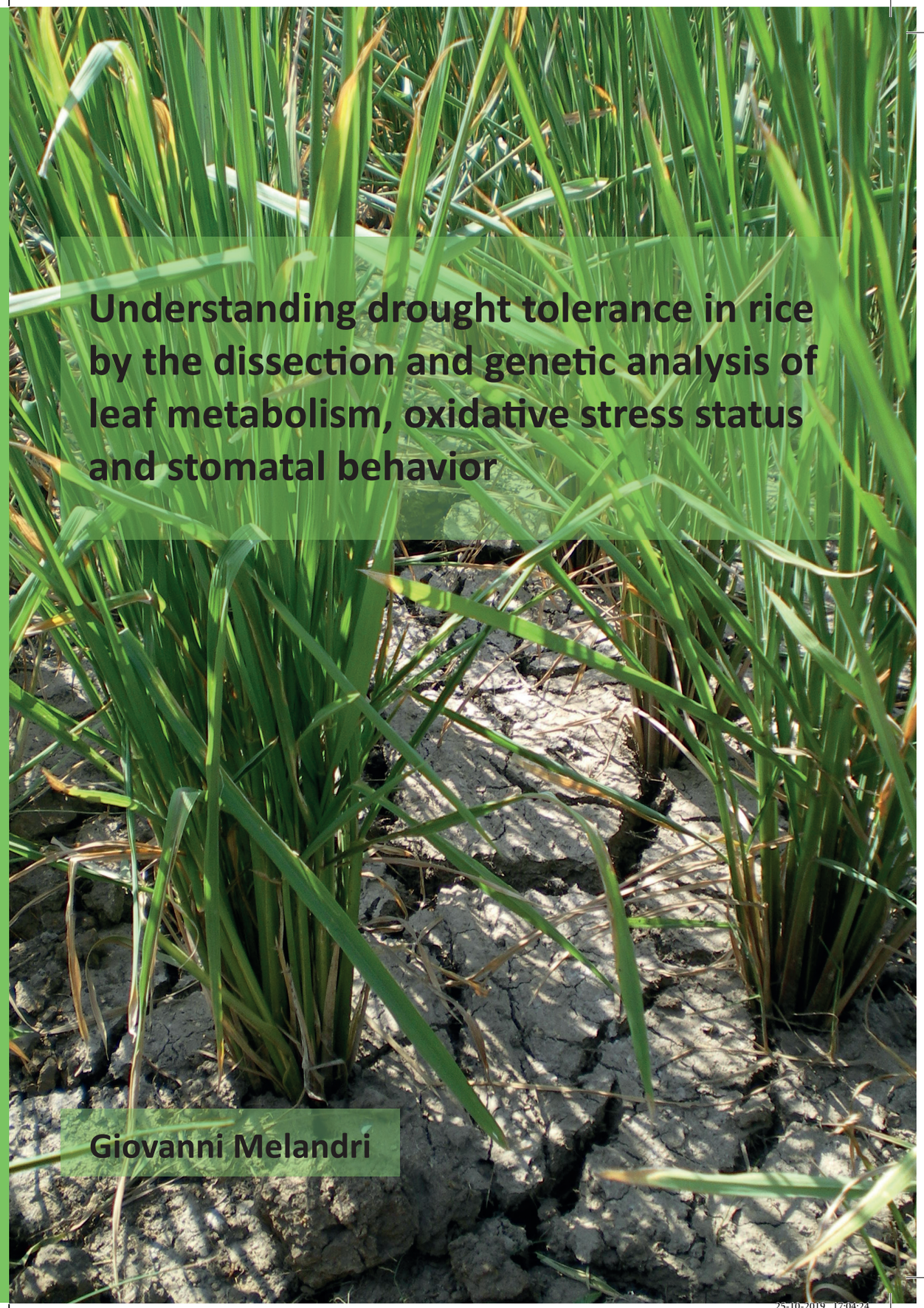
Understanding drought tolerance in rice by the dissection and genetic analysis of leaf metabolism, oxidative stress status and stomatal behavior

Giovanni Melandri 2019



Understanding drought tolerance in rice by the dissection and genetic analysis of leaf metabolism, oxidative stress status and stomatal behavior

Giovanni Melandri

Propositions

1. Developmental stage-related metabolic adaptations induce drought tolerance in rice. (this thesis)

2. Missing heritability is hidden under environmentally induced metabolic changes. (this thesis)

3. Although re-engineering of photorespiration increased tobacco growth under well-watered conditions (South *et al.*, 2019. Science, 363(6422)) it may result in suboptimal metabolism-related growth reduction under abiotic stresses.

4. The increasing global vapor pressure deficit (Yuan *et al.*, 2019. Science Advances, 5(8)) negates any positive effects of increased CO_2 levels and will require to re-design the future of agricultural crop productivity.

5. Although they seem highly ethical, vegans are quite unethical as they consider plants as inferior living organisms.

6. We are currently witnessing the shift from technological alienation to technological identification.

Propositions belonging to the thesis, entitled

Understanding drought tolerance in rice by the dissection and genetic analysis of leaf metabolism, oxidative stress status and stomatal behavior

Giovanni Melandri

Wageningen, 25 November 2019

Understanding drought tolerance in rice by the dissection and genetic analysis of leaf metabolism, oxidative stress status and stomatal behavior

Giovanni Melandri

Thesis committee

Promotor
Prof. Dr Harro J. Bouwmeester
Professor of Plant Physiology
Wageningen University & Research
Chair of Plant Hormone Biology
University of Amsterdam

Co-Promotor
Dr Carolien Ruyter-Spira
Researcher, Laboratory of Plant Physiology
Wageningen University & Research

Other members
Prof. Dr Paul C. Struik, Wageningen University & Research
Prof. Dr Yuling Bai, Wageningen University & Research
Dr Willem T. Kruijer, Wageningen University & Research
Dr Ric C.H. de Vos, Wageningen University & Research

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences (EPS)

Understanding drought tolerance in rice by the dissection and genetic analysis of leaf metabolism, oxidative stress status and stomatal behavior

Giovanni Melandri

Thesis

submitted in fulfillment of the requirements for the degree of doctor

at Wageningen University

by the authority of the Rector Magnificus,

Prof. Dr A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public on

Monday 25 November 2019

at 1:30 p.m. in the Aula.

Giovanni Melandri

Understanding drought tolerance in rice by the dissection and genetic analysis of leaf metabolism, oxidative stress status and stomatal behavior,
182 pages

PhD thesis Wageningen University, Wageningen, the Netherlands (2019)
With references, with summary in English

ISBN 978-94-6380-602-2

ISBN: 978-94-6395-217-0
DOI: https://doi.org/10.18174/507231

Table of contents

Chapter 1 **p. 7**
General introduction

Chapter 2 **p. 21**
Drought stress induces distinct physiological and biochemical responses
in lowland, aerobic and upland rice

Chapter 3 **p. 49**
Biomarkers for grain yield stability in rice under drought stress

Chapter 4 **p. 79**
Association mapping and genetic dissection of drought-induced
canopy temperature differences in rice

Chapter 5 **p. 109**
Genome wide association mapping reveals the genetic basis
of metabolic and enzymatic biomarkers for rice grain yield stability under drought

Chapter 6 **p. 145**
General discussion

References **p. 157**

Summary **p. 177**

Aknowledgements **p. 179**

CHAPTER

General introduction

Rice and water

As a staple food for more than half of the world's population, Asian rice (*Oryza sativa*) is considered the most important cereal crop worldwide and it is by far the most consumed food crop by people living in low- and low to middle-income countries (www.ricepedia.org, CGIAR). Molecular and archeological evidence dates the first domestication of Asian rice from its wild ancestor *Oryza rufipogon* to ~9,000-10,000 years ago, in the lower Yangtze river region of China (Fuller *et al.*, 2010; Molina *et al.*, 2011; Wing *et al.*, 2018). Today, rice is cultivated on more than 140 million hectares globally, with more than 90% produced in Asia. Approximately 60% is grown in irrigated lowland, 34% in rainfed lowland, and 6% in rainfed upland agroecosystems (FAO, 2014). The predominance of irrigated lowland systems is due to the higher grain yields that can be achieved when it is possible to control access to water. The grain yield penalty that occurs when rice is grown in rainfed cultivation systems is related to the variability and uncertainty of rainfall patterns, and to the crop's sensitivity to drought, disease and micronutrient disorders when water is limited, particularly during the reproductive stage (Venuprasad *et al.*, 2007). The Green Revolution (1965-1990) promoted the widespread use of irrigation infrastructure and rice varieties adapted to irrigated conditions as part of a strategy to enhance yield and productivity of rice worldwide. Indirectly, this led to an overall increase in the sensitivity of high-yielding varieties (HYV) of rice to water limitation. While striving to improve yield and food security, new rice varieties were selected in irrigated ecosystems without considering water as a limiting factor (Pingali, 2012; Kumar *et al.*, 2014). As a consequence, the widely adopted HYVs associated with the Green Revolution show large yield losses when exposed to drought stress (Lafitte *et al.*, 2006; Vikram *et al.*, 2015). In more recent years, there has been increasing awareness that, in the coming decades, water scarcity will represent a major constraint for agriculture, necessitating a focus on increasing water-use-efficiency (WUE) in agricultural systems (Rijsberman, 2006). Several management practices, such as alternate wetting and drying and saturated soil culture, have been developed through the years to enhance WUE in rice cultivation (Borrell *et al.*, 1997; Tabbal *et al.*, 2002). Despite these developments, there is an unavoidable trade-off when water-saving practices are applied to high-yielding rice production systems in irrigated lowland environments, leading to a severe reduction in grain yield (Bouman and Tuong, 2001; Peng *et al.*, 2006). It is therefore necessary to develop new rice varieties better adapted to water-limited production systems. Prior to the Green Revolution, traditional rice varieties were

cultivated for centuries in rainfed lowland and upland regions of Asia. Their selection in water-limited rainfed ecosystems resulted in adaptive drought-resistant characteristics, accompanied by a genetic differentiation between lowland and upland rice varieties (Xia *et al.*, 2014). This genetic diversity represents a valuable reservoir of potentially useful genes and genetic mechanisms associated with field-level drought resistance of use to rice breeders. The possibility to move favorable alleles from drought-tolerant donors to high-yielding but drought-susceptible rice varieties could help to improve the yield stability of HYVs under water-limited conditions (Luo, 2010).

Linking drought tolerance traits to rice productivity

Before describing the meaning of 'tolerance' or 'resistance' to drought considered in my thesis, it is necessary to provide a definition of 'drought' from the perspective of crop physiologists and breeders. Among the many possible definitions presented in the literature, I prefer the one used by Jones (2013) that links water supply to plant productivity. In this context drought refers to "*any combination of restricted water supply (e.g. as a result of low rainfall or poor soil water storage) and/or enhanced rate of water loss (resulting from high evaporative demand) that tends to reduce plant productivity*".

Given this definition, the productivity of plants under drought is affected by two main factors: 1- the length and intensity of the drought period; 2- the plant phenological stage (Blum, 2005). The presence of extended and severe drought periods is typical of arid and semiarid environments to which plants have adapted by developing 'drought survival' mechanisms (e.g. desiccation tolerance in resurrection plants) which are not compatible with acceptable yields in modern agricultural systems and, thus with the production of grain crops (Sinclair, 2011). This is particularly true for a semi-aquatic crop like rice that, among the cereals, shows greatest sensitivity to water limitation (Venuprasad *et al.*, 2007). The two production ecosystems where a rice crop typically experiences periods of drought stress are rainfed uplands and rainfed lowlands. In rainfed upland fields where standing water is usually rarely present, conditions of mild drought stress affect the crop during its entire life cycle and severe drought can be commonly experienced between major rainfall events (Kamoshita *et al.*, 2008). Under these conditions, genotypes characterized by 'drought avoidance' mechanisms are associated with better yields. The avoidance strategy aims to maintain a high plant water status under

drought, which can be achieved by reducing water use and/or enhancing water uptake and, therefore, retaining high cellular hydration (Blum, 2005). Typical traits contributing to drought avoidance are early flowering (also known as 'drought escape'), reduced leaf growth, limited tillering, reduced stomatal conductance and deep rooting (Bernier *et al.*, 2008). All these drought avoidance-related traits, which are beneficial for improved yield performance in rainfed upland fields, may have an opposite effect in production ecosystems characterized by higher water availability (Tardieu, 2012). Rainfed lowland rice fields commonly have standing water in the paddies and drought stress may develop only with the occurrence of rainless periods. The timing of these periods generates three possible drought patterns: terminal drought (during the reproductive stage and grain filling), early drought (during the vegetative stage before maximum tillering) and intermittent drought (Kamoshita *et al.*, 2008). Under these conditions, drought avoidance mechanisms, by promoting a constitutive moderation of water use throughout the crop life cycle, would hamper maximized plant productivity during the extended periods of water availability, thus reducing the potentially high crop yield (Blum, 2005). Also a trait like deep rooting that, by increasing water uptake, helps to stabilize yield during periods of drought in upland soils (Kamoshita *et al.*, 2008), is not necessarily a favorable trait in rainfed lowland ecosystems. Indeed, a deep root system in rice is almost always associated with a reduced number of tillers (typical of upland drought-resistant varieties) and, therefore, contrasts with one of the most important traits favoring high yield potential (Blum, 2005). Nevertheless, Uga *et al.* (2013) discovered that a single quantitative trait locus (*DRO1*), by altering the root growth angle, promotes deeper rooting in rice without changing shoot and root biomass. The introduction of *DRO1* into the genetic background of IR64, a widely grown lowland HYV, increased grain yield under moderate and severe drought stress without a yield penalty under irrigated conditions (Uga *et al.*, 2013; Arai-Sanoh *et al.*, 2014). The *DRO1* case is a perfect example of how 'drought tolerance' in high-yielding varieties can be improved by identifying useful traits that have no detrimental effects on yield potential, and by manipulating the corresponding genes (e.g. by marker assisted selection and/or gene transformation/editing) (Cattivelli *et al.*, 2008). A trait with a similar positive potential effect on yield is osmotic adjustment that, through the stress-induced accumulation of specific metabolites, enables plants to maintain water absorption and cell turgor pressure under drought stress (low water potential), thus maintaining active photosynthesis and expansion of growth (Cattivelli *et al.*, 2008; Blum, 2017). Other similar stress tolerance mechanisms/traits

aimed at protecting against drought-induced cellular damage, and corresponding negative effects on yield, are the detoxification of reactive oxygen species (ROS) and the accumulation of proteins (e.g. LEAs and dehydrins) and metabolites (e.g. proline, sugars and glycine betaine) with a protective function (Claeys and Inze, 2013). All these drought tolerance mechanisms are controlled by stress-induced gene regulatory networks (Nakashima *et al.*, 2009) and, therefore, by being activated only under water-limited conditions, they do not result in yield penalty under good water supply. Inducible traits can be particularly important in rainfed lowland ecosystems because they are activated only when the paddy dries out as rains stop, and can be de-activated when water again becomes available. A better understanding of the effect of these mechanisms in rice and the manipulation of their underlying genetic control could help to reduce drought-induced yield loss in rainfed lowland ecosystems as well as in wetting-and-drying management systems in irrigated paddies.

The effect of drought on plant primary metabolism

Drought stress impacts the physiology of plants and alters many cellular functions causing a disruption of metabolic homeostasis that requires reprogramming to adapt to the stress. This metabolic reprogramming is determined by the necessity of maintaining essential metabolism while adapting to the new stressful conditions and, simultaneously, by the need to produce anti-stress agents to mitigate or tolerate the possible damages caused by low water potential (Claeys and Inze, 2013). In the last 10-15 years, studies based on targeted and un-targeted metabolic profiling of plant tissues have generated data on the metabolic changes induced by abiotic stresses, including drought, and have been the focus of many reviews (Shulaev *et al.*, 2008; Obata and Fernie, 2012; Krasensky and Jonak, 2012; Nakabayashi and Saito, 2015). Recently, and more relevant for the scope of this thesis, a comparison of the metabolic responses to drought between *Arabidopsis* (based on five independent studies) and rice (based on a single study) was reviewed by Fàbregas and Fernie (2019).

Overall, the most salient pattern observed across all these studies is a drought-induced accumulation of many primary metabolites. Among them, increased levels of raffinose family oligosaccharides (RFOs), such as raffinose and galactinol, are described as an early drought response in many plant species (Obata and Fernie, 2012), including rice (Todaka *et al.*, 2017). The early accumulation of RFOs under stress has been attributed to the osmoprotective

function of these compounds and to their activity as scavengers of hydroxyl radicals (Nishizawa *et al.*, 2008; Ende, 2013). Other sugars that also show an early response to drought are glucose, fructose, sucrose and erythritol (Fàbregas and Fernie, 2019). Except for sucrose in rice, which shows a decrease (Todaka *et al.*, 2017), these sugars display an accumulation in *Arabidopsis* and rice plants exposed to early, moderate drought stress. Under drought, the accumulation of sugars has been associated with the carbon surplus caused by stress-induced growth reduction, the earliest response to the stress, without a decrease in photosynthetic rate (Hummel *et al.*, 2010; Muller *et al.*, 2011). Sugars accumulating in response to drought stress can serve as osmolytes to maintain cell turgor and protect cell membranes and proteins (Krasensky and Jonak, 2012). Prolonged and severe drought induces the closure of stomata to reduce water loss, but this, in turn, limits photosynthesis and alters all the associated metabolic pathways, including the metabolism of carbohydrates (Pinheiro and Chaves, 2011). Under these conditions, high levels of sugar likely originate from stress-induced starch degradation (Sulpice *et al.*, 2009; Pinheiro and Chaves, 2011).

A very common metabolic response, shared by many plant species under severe drought stress, is the accumulation of free amino acids (Obata and Fernie, 2012; Fàbregas and Fernie, 2019). The occurrence of high levels of free amino acids only under severe stress, when a limited photosynthetic rate reduces biosynthesis and the stress favors leaf senescence, suggests that their origin is mainly from protein catabolism (Krasensky and Jonak, 2012). This abundant amino acid pool feeds into alternative metabolic pathways which can serve to withstand the stress. Branched-chain amino acids (BCAAs), for example, such as isoleucine, leucine, valine, lysine and beta-alanine, are utilized as alternative substrates for mitochondrial respiration under stressful conditions (Araujo *et al.*, 2012; Pires *et al.*, 2016). Aromatic amino acids (AAAs), such as tryptophan, tyrosine and phenylalanine, are precursors of many biosynthetic pathways of secondary metabolites (Vogt, 2010; Galili *et al.*, 2015) important for adaptation to diverse forms of abiotic stress, including drought (Nakabayashi and Saito, 2015). Different from the BCAAs and AAAs , the high levels of two specific amino acids, proline and GABA, which are involved in ROS scavenging, osmoregulation and coordination of carbon-nitrogen balance (Szabados and Savouré, 2010; Verslues and Juenger, 2011; Michaeli and Fromm, 2015), could originate from active production under severe drought stress and not merely from the catabolism of protein. Similar to amino acids, the levels of many organic acids

of the tricarboxylic acid (TCA) cycle markedly change under moderate-to-severe drought stress, with an increase shown in *Arabidopsis* but a decrease in rice (Fàbregas and Fernie, 2019). This difference underlines an opposite response of the TCA cycle and, therefore, of mitochondrial respiratory metabolism, between these two species in response to drought stress. Even though the above-mentioned studies on changes in single metabolic pathways have greatly improved our understanding of the adaptations of plant primary metabolism to drought, future studies will have to integrate these independent pathways into more comprehensive modeling approaches (Sulpice and McKeown, 2015). Apart from the plant species, these models will also have to consider variables with a strong influence on plant metabolism, such as the specific plant tissue (seed, leaf, root), the developmental stage (young/mature leaf, vegetative/reproductive stage) and, primarily in the case of crop species, the difference in the metabolic response to drought between controlled and field environments (Obata and Fernie, 2012; Fàbregas and Fernie, 2019).

Drought-induced oxidative stress and damage

In plant cells under optimal conditions, ROS production is a physiological consequence of metabolic reactions during photosynthesis and respiration, and is mainly located in cell organelles such as chloroplasts, peroxisomes and mitochondria (Apel and Hirt, 2004). ROS are known as toxic compounds responsible for oxidative stress as well as important signalling molecules involved in many developmental, metabolic and defensive pathways (Mittler *et al.*, 2011; Mittler, 2017). In photosynthetic cells exposed to light, chloroplasts and peroxisomes are the main sources of ROS with an estimated production of hydrogen peroxide (H_2O_2) of 4,000 and 10,000 nmol m^{-2} s^{-1}, respectively (Foyer and Noctor, 2003). Mitochondria are also known to be responsible for ROS generation under non-stressed light conditions, but to a lower extent than chloroplasts and peroxisomes. In mitochondria, H_2O_2 formation was estimated to be 30 to 100 times lower than in the other two organelles with a production of less than 200 nmol m^{-2} s^{-1} (Foyer and Noctor, 2003).

In chloroplasts, the reaction centres photosystem I (PSI) and photosystem II (PSII) generate ROS through a mechanism called photoproduction. The overloading of the electron transport chain in PSI is responsible for the diversion of part of the electron flow from ferredoxin to O_2 with the generation of its reduced product, superoxide anion (O_2^-), *via* the Mehler reaction.

O_2^- is then converted to H_2O_2 by the enzyme superoxide dismutase and then eventually reduced to water by the enzyme ascorbate peroxidase (APX) (Asada, 2006; Das and Roychoudhury, 2014). The H_2O_2-scavenging activity of APX is exerted by oxidizing ascorbate (AsA) to monodehydroascorbate (MDHA) that can be directly regenerated to AsA by the enzyme monodehydroascorbate reductase (MDHAR). MDHA is a short-lived compound that spontaneously and rapidly converts to dehydroascorbate (DHA) which, in turn, can be reduced back to AsA by the action of the enzyme dehydroascorbate reductase (DHAR). In the regeneration of AsA, DHAR utilizes reduced glutathione as electron donor and it participates, together with the enzyme glutathione reductase, in the ascorbate-glutathione cycle, the main redox hub in plants (Foyer and Noctor, 2011). In PSII of chloroplasts, a lack of energy dissipation during photosynthesis promotes the formation of triplet state chlorophyll that can react with triplet state oxygen ($^3O_2$) generating the extremely reactive singlet oxygen ($^1O_2$). $^1O_2$ can be scavenged by α-tocopherol (vitamin E), a very effective lipid soluble antioxidant present in the thylakoid membranes of chloroplasts (Gill and Tuteja, 2010).

The effectiveness of the above-mentioned ROS scavenging mechanisms is drastically reduced by conditions that favour ROS overproduction (Asada, 2006). Under drought, plants have to balance the need for CO_2 for photosynthesis with reduced water loss, which is mainly achieved by closing the stomata. Stomata closure reduces the CO_2 supply for photosynthesis and hence indirectly generates a relative excess of photon intensity that causes enhanced ROS photoproduction (Noctor *et al.*, 2014; Das and Roychoudhury, 2014; You and Chan, 2015). Under severe and prolonged drought stress, chloroplast antioxidant defences are no longer able to guarantee the scavenging of ROS overproduction that is likely to be the prime cause for the inhibition of the repair of the photodamaged PSII. This damage, known as photoinhibition, decreases photosynthetic activity and therefore growth and productivity (Takahashi and Murata, 2008; Takahashi and Badger, 2011).

Besides chloroplasts, peroxisomes are the major source of ROS in plant cells. Under drought, ROS generation in the peroxisomes is tightly associated with reduction in photosynthetic activity at the level of the Calvin cycle. Under these conditions, the Rubisco enzyme, which has a dual activity as carboxylase and oxygenase, fixes more O_2 (oxygenase activity) because of the limitation in the CO_2 supply resulting from drought-induced stomata closure (Noctor, 2002). The oxygenase activity of Rubisco generates phosphoglycolate (2PGA), a toxic

compound, that is recycled to 3-phosphoglycerate (3PGA). This recycling pathway is known as photorespiration and allows the Calvin cycle to proceed into its following steps towards central metabolic pathways and, simultaneously, to regenerate the Rubisco substrate ribulose 1,5-bisphosphate (Peterhansel and Maurino, 2011). The photorespiration pathway is characterized by a high metabolic cost in terms of carbon and energy losses. In addition, in this pathway, 2PGA is initially converted to glycolate that is then transported from the chloroplasts to the peroxisomes where it is converted to glyoxylate by the enzyme glycolate oxidase (Peterhansel and Maurino, 2011). This conversion generates high amounts of H_2O_2 which can be directly converted to water and O_2 by the enzyme catalase (Das and Roychoudhury, 2014). Noctor (2002) estimated that photorespiration is the source of over 70% of the H_2O_2 generated in plants under drought.

As previously mentioned, drought, like other abiotic stresses, induces ROS accumulation by altering the balance between ROS production and the cellular scavenging capacity with a resulting state of "oxidative stress" that, in turn, can lead to "oxidative damage" (Halliwell, 2006). When ROS accumulation is particularly severe, and reaches phytotoxic levels, the resulting oxidative damage to proteins, lipids and DNA can lead to cell death (Van Breusegem and Dat, 2006). Polyunsaturated fatty acids (PUFAs) and, particularly, linoleic (18:2) and linolenic (18:3) acids, are the major fatty acids in the membranes of thylakoids and other cell membranes. These two fatty acids are the favourite targets of ROS-mediated peroxidation which generates lipid hydroperoxides that, in turn, can oxidize other neighbouring PUFAs. The establishment of this reaction chain results in decreased membrane fluidity, increased leakiness and causes secondary damage to membrane proteins (Møller *et al.*, 2007). In addition to this highly deleterious effect, aldehydic products of PUFA peroxidation can generate secondary damage to other molecules. This is the case for malondialdehyde, a product of the peroxidation of linolenic acid, that can form a conjugate with guanine and indirectly modify DNA (Møller *et al.*, 2007). Cell death is the extreme consequence of ROS-mediated oxidative stress generated by the metabolic imbalance due to stressful conditions. Indeed, the enhanced production of ROS in plant organelles and the consequent generation of many oxidized metabolites and carbonylated and/or dithiol-disulfide exchanged proteins can trigger signalling pathways (retrograde signalling to the nucleus) aimed to activate regulatory responsive genes. These genes are then able to buffer ROS accumulation by

increasing the above-mentioned ROS-scavenging enzymes and antioxidant molecules to acclimate plant tissues to the stressful conditions (Gill and Tuteja, 2010; Suzuki *et al.*, 2012).

Breeding for drought tolerance by selecting biochemical and physiological traits

Direct selection for high grain yield under drought is difficult because the trait is defined by low heritability, complex polygenic control and strong genotype-by-environment interactions (Cattivelli *et al.*, 2008). For these reasons, targeting biochemical and physiological traits that underlie the complex yield trait, represents an alternative breeding strategy for increasing yield under drought (Hu and Xiong, 2014; Pandey and Shukla, 2015; Reynolds and Langridge, 2016). In previous paragraphs, I discussed several of these traits and, among them, mentioned the stress-induced production of compatible solutes and antioxidants that can stabilize proteins, maintain cell turgor and remove the excess levels of ROS (Krasensky and Jonak, 2012). All these metabolites/antioxidants could be directly targeted by breeding programs, as biomarkers, once their positive association with yield performance under drought has been demonstrated (Fernandez *et al.*, 2016). Additionally, considering that these metabolite/antioxidant traits are characterized by less complex genetic control than yield (Reynolds and Langridge, 2016), they might be converted into genetic markers, more suitable for breeding purposes. Confirming this possibility, many metabolic and enzymatic quantitative trait loci (QTLs) were identified in experiments involving large-scale metabolite and enzyme profiling of bi-parental and mapping populations of different plant species (Keurentjes *et al.*, 2008; Zhang *et al.*, 2010; Carreno-Quintero *et al.*, 2013; Luo, 2015; Fernie and Tohge, 2017), including rice (Gong *et al.*, 2013; Chen *et al.*, 2014; Dong *et al.*, 2015; Matsuda *et al.*, 2015). However, all these genomic regions were identified under non-stress conditions and, therefore, there is a need for new studies to search for stress-induced metabolic and enzymatic QTLs associated with grain yield stability under favorable and drought stress conditions. A similar approach as for biochemical markers, could be adopted for the selection of physiological traits important for drought tolerance, such as regulation of photosynthesis, total transpiration, stomatal conductance, water use efficiency (WUE) and late senescence (Pandey and Shukla, 2015). Despite the fact that these physiological traits are characterized by higher genetic complexity than metabolites and enzymes, they are considered important phenotypes in breeding for drought tolerance (Reynolds and Langridge, 2016), and their translation into genetic markers is of great interest. Interestingly, QTLs associated with

differences in canopy temperature, a good proxy for stomatal conductance (Leinonen *et al.*, 2006; Munns *et al.*, 2010), have been detected in wheat and linked to drought and heat stress tolerance (Rebetzke *et al.*, 2012; Rutkoski *et al.*, 2016). Considering that genotypic variation in stomatal conductance has been described in rice (Ohsumi *et al.*, 2006; Ouyang *et al.*, 2017), the identification of the genetic control of this variation offers a potential target for improving drought tolerance in this important crop species.

Thesis outline

The aim of my thesis research is to investigate how drought-induced changes in rice physiology, central metabolism, and oxidative stress status impact crop growth and yield. My work should result in the identification of physiological traits and metabolic- and oxidative stress-related biomarkers that can potentially be used in breeding programs to stabilize rice grain yield under drought stress. I also used the genetic diversity of a rice genome wide association mapping panel, consisting of almost 300 *indica* rice accessions, to map genes and genomic regions associated with quantitative variation for these traits and biomarkers in an effort to identify useful genetic markers and to unravel underlying mechanisms.

This dissertation consists of six chapters including the present one that serves as a General introduction on the research topic (**Chapter 1**).

In **Chapter 2**, I describe the physiological and leaf biochemical responses of three *indica* rice varieties selected for their contrasting responses to increasing drought severity and recovery following re-watering at the vegetative stage in a controlled environment and at the reproductive stage in the field. The results of these experiments provided the basis for a comparison of the response of these three rice varieties at two different developmental stages and in two different environments.

In **Chapter 3**, I analyse the stress-induced changes in flag leaf central metabolism and oxidative stress status in ~300 *indica* rice accessions exposed to drought in the field at the reproductive stage. The levels of metabolites and oxidative stress markers/enzymes are then used to generate a multivariate model for the prediction of grain yield loss across the accessions of the panel and to identify robust biomarkers for response to drought.

In **Chapter 4**, I describe canopy temperature differences in ~300 *indica* rice accessions detected under well-watered and reproductive stage drought stress conditions in the field, and correlate canopy temperature with a set of agronomic traits, including grain yield. Canopy temperature variation is then used as a quantitative trait for association mapping to identify QTLs/genes controlling trait variation, and candidate genes for control of stomatal conductance are discussed.

In **Chapter 5**, I use the multivariate modeling approach of Chapter 3 to compare the predictive power of the metabolic/oxidative stress dataset and of genetic markers for the prediction of grain yield under well-watered and drought-stress conditions and for grain yield-loss. The best metabolic/oxidative stress predictors from these models are used for association mapping to identify QTLs/genes controlling these predictors, and candidate genes are discussed.

Finally, in **Chapter 6** I discuss the main highlights of this thesis and connect the results of the different experimental chapters. I describe the prospects for using the main outcomes of my work in agriculture and breeding, and discuss further research that is needed to fully exploit my findings.

CHAPTER

2

Drought stress induces distinct physiological and biochemical responses in lowland, aerobic and upland rice

Giovanni Melandri, Hamada AbdElgawad, Kristýna Floková,
Diaan C. Jamar, C. Gerard van der Linden, Han Asard,
Gerrit T. S. Beemster, Carolien Ruyter-Spira, Harro J. Bouwmeester

In preparation for submission

ABSTRACT

A clear understanding of physiological and biochemical adaptations to drought can help to identify the mechanisms underlying tolerance to this stress. In this study, three rice (*Oryza sativa*) varieties, characterized by contrasting levels of drought tolerance, were exposed to drought stress at the vegetative and reproductive stage. Under controlled conditions, changes in biomass, leaf metabolites, oxidative stress markers and enzyme activities were analysed in each genotype under increasing drought stress and after re-watering during vegetative development. The two drought tolerant genotypes, Apo and UPL Ri-7, displayed a conservative water use in contrast to the susceptible genotype IR64 that displayed high water consumption and consequent leaf dehydration. A sugar-mediated osmotic acclimation in UPL Ri-7 (upland rice) and a strong antioxidative response in Apo (aerobic rice) limited the drought-induced biomass loss in these two genotypes, while biomass loss was high in IR64 (lowland rice), also after recovery. In the reproductive stage in the field, sugar export from the flag leaves to the developing panicles excluded osmotic adjustment as a mechanism to withstand drought conferring a competitive advantage to Apo, which, also in this stage, showed the highest antioxidant power and was able to maintain a stable grain yield under stress.

KEY WORDS: *Oryza sativa*, drought, leaf primary metabolism, leaf oxidative stress status, vegetative stage, reproductive stage.

INTRODUCTION

Drought impacts on the morphology and physiology of plants and induces metabolic reprogramming to adapt to the stress. The extent of this reprogramming is key in the trade-off between growth and survival. Essential metabolism under stressful conditions needs to be maintained, while anti-stress agents such as compatible solutes, antioxidants, stress-responsive proteins and enzymes need to be produced (Obata and Fernie, 2012; Claeys and Inze, 2013). Growth reduction is an early response to water limitation and frequently occurs without any alteration in photosynthetic rate (Skirycz and Inzé, 2010; Fàbregas and Fernie, 2019). Prolonged drought induces stomatal closure to reduce water loss, but this also limits photosynthetic CO_2 assimilation, resulting in metabolic alterations and constraints (Chaves *et al.*, 2009; Pinheiro and Chaves, 2011). As a consequence, adjusting carbohydrate biosynthesis and translocation, for example for osmoregulation, plays a central role in the response to drought (Luquet *et al.*, 2008; Hummel *et al.*, 2010; Muller *et al.*, 2011). Previous research on drought stress, has shown that accumulation of particular metabolites in leaves (e.g. raffinose, trehalose, proline, and glycine betaine) can have a protective function whereas the accumulation of other metabolites may simply be a consequence of drought stress (for example the increase of free amino acids from protein breakdown) (Verslues and Juenger, 2011; Obata and Fernie, 2012; Krasensky and Jonak, 2012; Fàbregas and Fernie, 2019). Another effect of drought-induced stomatal closure and lower carbon availability, is an enhanced generation of reactive oxygen species (ROS) (Suzuki *et al.*, 2012; Noctor *et al.*, 2014) responsible for oxidative damage that drives the cell into senescence (Halliwell, 2006) and, in extreme cases, to death (Van Breusegem and Dat 2006). A complex enzymatic and non-enzymatic antioxidative system protects plants against this oxidative damage, and is essential for conferring drought tolerance (Mittler *et al.*, 2011; Baxter *et al.*, 2014; You and Chan, 2015).

An additional layer of complexity in plant stress responses to drought relates to the fact that the stress has a different impact on plant performance at different developmental stages. For this reason, the reproductive stage is often targeted in drought experiments involving cereals, as the occurrence of the stress at this stage results in the most severe grain yield reduction (Passioura, 2012; Biswal and Kohli, 2013; Reynolds *et al.*, 2016). This is particularly true for rice, in which even moderate stress during flowering can result in a strongly reduced grain yield (Liu *et al.*, 2006; Venuprasad *et al.*, 2007; Sandhu *et al.*, 2014). Nevertheless, in the

coming years, cultivation of rice under non-flooded conditions, in non-puddled and non-water-saturated soil ('aerobic rice') during the entire crop cycle is expected to expand (Venuprasad *et al.*, 2012). Such rice cultivation is particularly important, considering the expected water limitations resulting from an increased variability and unpredictability of precipitation due to global warming (Ray *et al.*, 2015). A better understanding of drought tolerance mechanisms in rice at the vegetative and reproductive stage will contribute to the selection of genotypes with better adaptation to water-limited production systems and can help to achieve a better use of water resources.

In this study, we investigated, under controlled conditions, how changes in leaf metabolism and oxidative stress status during vegetative development are associated with morphological and physiological changes (water consumption over time, biomass accumulation and leaf water status) during progressive drought and after re-watering. The study was done with three rice varieties with contrasting levels of drought tolerance, 1) IR64, a high-yielding genotype that is commonly grown under flooded conditions and is highly susceptible to drought as indicated by its considerable reduction of grain yield when grown under aerobic conditions (Mackill and Khush, 2018); 2) Apo, a drought tolerant aerobic-adapted variety with good yield potential (Venuprasad *et al.*, 2012); and 3) UPL Ri-7, an upland-adapted variety with improved yield potential (Atlin *et al.*, 2006). We compared the leaf metabolic and oxidative stress profiles during vegetative development with the profiles of the same varieties when exposed to drought in the reproductive stage, under field conditions, to get insight into the similarity of the drought stress coping strategies in the two different phenological stages and under controlled and field conditions.

MATERIALS AND METHODS

Pant material, stress treatment and sampling

Seeds of the three indica rice (O. sativa) genotypes were obtained from the International Rice Research Institute (IRRI) gene bank: (1) IR64 (IRGC 117268); (2) Apo (IRGC 115128); (3) UPL Ri-7 (IRTP 9897) (hereafter called Ri-7). Seeds were directly sown into 1.75 litre pots (14 cm ø, 18 cm height) filled with the same amount of dry field soil. Plants (one per pot) were grown in a controlled climate chamber (12h photoperiod, 28/23°C (d/n); 75/70% relative humidity (d/n); 600 µmol m^{-2} s^{-1} PAR) and watered every second day with ½ strength Hoagland's solution until

day 24 after sowing. On day 25 after sowing, all pots were watered to 140% field capacity (flooded). Stressed plants were not watered for 10 consecutive days before being re-watered to flooded conditions. Pots of stressed plants were covered with a reflecting disc to reduce water evaporation from the soil (Supplementary Fig. S1). Control plants were kept flooded during the entire experiment duration. The experimental design was a completely randomized single-block (trolleys) design, each comprising 24 pots with plants of the same genotype and the same treatment. The position of each trolley was rotated daily. Trolleys' height was adjusted to expose all plants to the same light intensity. Every day, twelve stressed plants per genotype were weighed (middle of the day, 6 hours of light) to determine water loss from the pots. Based on the progression of the drought treatment, four time points were selected for biomass measurements and leaf sampling: TP1, 4 days of stress; TP2, at around 40% field capacity; TP3, 10 days of stress; TP4, 2 days after re-watering (recovery). At each time point, for every genotype, 3 control and 3 stressed plants were randomly selected to determine the dry weight of leaves, tillers and roots, the number of tillers and the RWC of the last fully developed leaf of the main tiller. RWC was expressed as (Fresh-Dry weight)/(Saturated-Dry weight)*100%.Towards the end of the day (10 hours of light) of each time point, 12 plants of the same genotype and treatment were sampled for metabolite analyses. From 3 different plants, 2 fully developed leaves on the primary tiller were collected, pooled and immediately frozen in liquid nitrogen to generate a biological replicate. At TP4, necrotic leaf parts of drought-stressed plants were not considered for metabolite and redox state analysis nor for biomass determination. In total at each time point, 4 biological replicates were collected per genotype and per treatment.

The same three rice genotypes were grown in a large field trial (~300 accessions) conducted at the International Rice Research Institute (IRRI), Philippines, during the 2013 dry season (Kadam *et al.*, 2018), under control and drought stress (2 weeks of water withholding at the reproductive stage).

Analysis of primary metabolites

Sugars (sucrose, fructose, glucose, trehalose and xylose) levels were measured using a Dionex HPLC system according to Bentsink *et al.* (2000), with minor modifications. The levels of nutrient anions (phosphate, nitrate and sulfate) and organic acids (citrate, isocitrate and α-

ketoglutarate) were determined by a Dionex HPLC system as described by He *et al.* (2014), with minor modifications. Sugars, nutrient anions and organic acids were expressed in $mg \cdot L^{-1} \cdot g^{-1}$ of dry weight (DW). Amino acid (Ala, Ser, Pro, Val, Thr, Ile, Leu, Asp, Glu, Met, His, Phe, Arg, Tyr, Lys, Gly, GABA, Asn, Gln, Trp and Orn) detection and quantification using UPLC-MS/MS was performed according to Carreno-Quintero *et al.* (2014) with modifications. Amino acids were expressed in $pmol \cdot \mu L^{-1} \cdot g^{-1}$ of DW. Detailed protocols can be found in the Supplementary Methods.

Flag leaf samples of the three genotypes were collected in the field trial during the last day of stress and metabolites were quantified by GC-MS and main sugars by a colorimetric assay as described in Melandri *et al.* (2019).

Quantification of oxidative stress markers and antioxidant enzyme activities

The lipid peroxidation product malondialdehyde (MDA) was assayed according to *Hodges et al.* (1999) and expressed in $nmol(MDA) \cdot g^{-1}$ of DW. Total antioxidant capacity (TAC) was assayed by FRAP reagent (Benzie and Strain, 1999) and expressed in $\mu mol(trolox) \cdot g^{-1}$ of DW. The amount of soluble protein in each sample was quantified by the Lowry method *(Lowry et al.*, 1951). Protein carbonylation (ProtOx) was assayed according to Levine *et al.* (1994) and expressed as $mg(carbonyl) \cdot mg^{-1}$ of protein. Ascorbate peroxidase (APX), dehydroascorbate reductase (DHAR), superoxide dismutase (SOD), Catalase (CAT) activities were measured using a micro-plate reader (Dhindsa *et al.*, 1982; Aebi, 1984; Murshed *et al.*, 2008). Activities were expressed as $\mu mol(AsA) \cdot mg(protein)^{-1} \cdot min^{-1}$ for APX and DHAR, $\mu mol(H_2O_2) \cdot mg(protein)^{-1} \cdot min^{-1}$ for CAT and $\mu mol(SOD) \cdot mg(protein)^{-1} \cdot min^{-1}$ for SOD. The described procedures were applied to both the samples from the controlled experiment and for the field trial. Detailed protocols can be found in the Supplementary Methods.

Statistical analysis

Statistical analyses were performed using R statistical software (version 3.1.2; The R Foundation for Statistical Computing). Analysis of variance (One-way ANOVA) and Tukey Honest Significant Differences (TukeyHSD) test were used to compare soil moisture content of each genotype at each time point, setting the threshold for statistical significance at $P < 0.05$. Imputation of missing metabolic and oxidative stress values, prior to any other statistical test,

was performed by the knnImputation function in the DMwR package. Student's t-test was performed to evaluate the significant differences between metabolites, oxidative stress markers and enzyme activities, and growth-related traits under the two different treatments for each genotype at each time point. Fold-change analysis, prior to PCA, was performed by dividing each individual metabolic and oxidative stress marker/enzyme value under drought by the mean value of the same metabolite and redox state marker/enzyme under control conditions for each genotype at each time point. PCA was performed using the prcomp function in the stats package. Each metabolite value was $\log_{10}$ transformed to improve normality, centered (mean subtraction) and scaled (standard deviation division) before PCA.

RESULTS

Plant transpiration and leaf water status

During the experiment under controlled conditions, the daily variation in pot water content - expressed as percentage of Field Capacity (FC) - was used as proxy for plant transpiration. At time point (TP) 1 (4 days after water withholding), pot water content was significantly different ($P < 0.001$) between the genotypes, with IR64 showing the highest water consumption, followed by Ri-7 and Apo (Fig. 1 and Supplementary Fig. S2). TP2 was selected at a specific soil moisture content (~40% FC) between the genotypes to compare their physiological and biochemical responses at similar drought intensity (Supplementary Fig. S2). For IR64, 40% FC was reached one day earlier (6 days after water withholding) than for Apo and Ri-7 (7 days), suggesting a higher water consumption under drought for the first genotype (Fig. 1). At TP3 (10 days after water withholding), pot water contents were extremely low (~20% FC) for all genotypes (Fig. 1) with, however, Apo displaying a significantly ($P < 0.001$) higher value than the other two genotypes (Supplementary Fig. S2).

To get insight into the progression of plant dehydration, we determined the relative water content (RWC) of the youngest fully developed leaf on the primary tiller (Fig. 2A-C and Supplementary Table S1). Control plants of all genotypes displayed a stable leaf RWC, higher than 90%, during all four TPs. Drought stressed plants only showed a significant difference in RWC from their controls at the end of the stress (TP3), with IR64 displaying a much lower value (24%) than the other two (~70%).

Collectively, these results suggest that drought sensitivity of IR64 is associated with high water consumption that resulted in strong leaf dehydration at the end of the stress period. In contrast, the drought tolerant genotypes Apo and Ri-7 had a more conservative water consumption (particularly Apo) and suffered less dehydration.

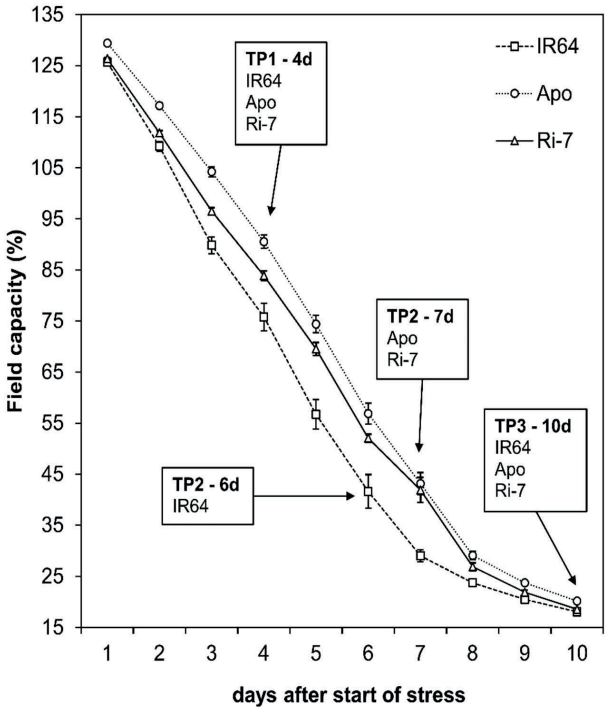


Fig. 1. Evolution of the soil water content in the drought treatment for three rice genotypes.
Each data point represents the mean field capacity value (%) of 12 different plants (±SE) per genotype. The first time point (TP1) for sampling was selected 4 days after water withholding for all the genotypes. The second time point (TP2) for sampling was selected 6 days after water withholding for IR64 and 7 days for Apo and Ri-7. The third time point (TP3) for sampling was selected 10 days after water withholding for all the genotypes.

Growth responses to increasing drought stress

At the end of the stress, drought resulted in visible growth reduction in the stressed plants, with clear differences between the genotypes (Fig. 3). Overall, under control conditions, IR64 accumulated a higher leaf biomass than the other two genotypes, at all TPs (Fig. 2D-F). Under drought stress, IR64 displayed a significant reduction in leaf biomass compared with the control, already at TP1 (Fig. 2G). This reduction remained significant and stable at TP2, and sharply increased at TP3 and TP4, when IR64 reached the largest reduction in leaf biomass (-

77.6 %) among the three genotypes. Leaf biomass of Apo was significantly reduced from TP2 onwards (Fig. 2E), and the difference with the control increased more gradually than in IR64 at TP3 and TP4 (Fig. 2H). Ri-7 did not show a significant response of leaf biomass in the first two TPs with even an increase (not significant) displayed at TP1 under drought. In Ri-7, the decrease in leaf biomass started at TP3 (Fig. 2F) and this continued at TP4 (Fig. 2I). Under severe drought (TP3), strong leaf rolling was observed in all plants of all genotypes. Rolling fully recovered after re-watering (TP4), but with leaf tips showing necrotic areas (Fig. 3) that were discarded and not considered for leaf biomass determination at TP4. The lowest reduction in leaf biomass between TP3 and TP4 was displayed by Apo (-12.7 %), followed by Ri-7 (-23.9 %) and IR64 (-32.7 %).

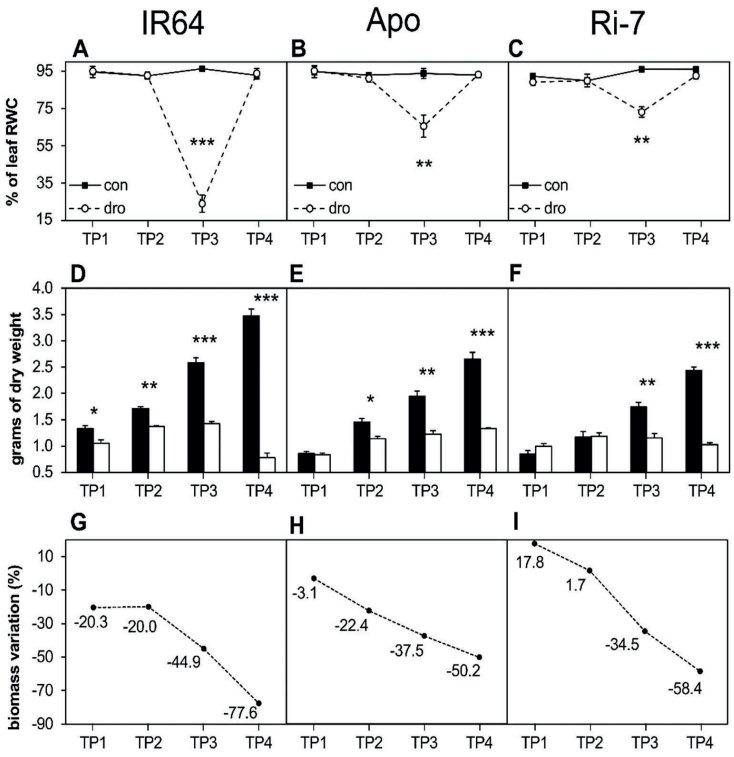


Fig. 2. Evolution of leaf water content and biomass in response to drought for three rice genotypes. Percentage of leaf relative water content (RWC) (A, B, C) of the top fully developed leaf in control and drought plants at each time point. Leaf biomass (dry weight) (D, E, F) at the four sampling time points (TP1, TP2, TP3, TP4). Leaf biomass difference (%) (G, H, I) between control and drought plants at each time point. Black and white columns represent the mean ± standard error (SE) of three biological replicates of control and drought stressed plants, respectively. Symbols represent the mean ± standard error (SE) of three biological replicates. ***, **, * represent t-test's *P*-values of significance between control and drought replicates with $P < 0.001$, $0.001 < P < 0.01$, $0.01 < P < 0.05$, respectively.

All genotypes showed a similar trend of biomass reduction for stem dry weight (DW), but with a delay of one time point and, in general, with less significant differences than for leaf weight (Supplementary Table S1). Stressed plants of all genotypes also had a significantly lower root DW compared with control plants, but only at TP4. Interestingly, in IR64 drought caused a significant reduction in tiller number under drought, between TP2 and TP4 (Supplementary Table S1).

Taken together, these results highlight contrasting constitutive growth rates between the three genotypes under control conditions and different growth responses to increasing drought as well as upon recovery.

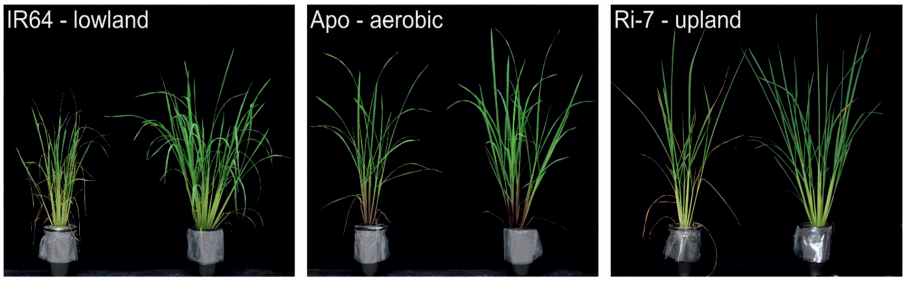


Fig. 3. Effect of drought treatment on the shoot phenotype of three rice genotypes after re-watering. Pictures of IR64 (left), Apo (middle) and Ri-7 (right) at the recovery time point (TP4 - 2 days after re-watering) of sampling. In each picture, the stressed plant is on the left with its respective control on the right.

Overall effect of drought and re-watering on metabolism and oxidative stress status

To get insight into the drought-induced changes in leaf metabolism and oxidative stress status, we analysed the levels of primary metabolites (29), nutrient anions (3) and oxidative stress markers/enzymes (7), at all time points. The levels (mean±st.dev) of the 39 variables under control and drought stress treatments are shown in Supplementary Table S2 and S3, respectively. First the effect of genotype and plant development on the metabolic, nutrient and oxidative stress variables under optimal conditions was analysed using principal component analysis (PCA) (Supplementary Fig. 3). The first two principal components (PCs) explained more than 50% of the total variation, but the samples are not clearly separated by developmental stage (TPs) and genotype (Supplementary Fig. S3A). Only TP1 and TP4 samples, with maximum plant age difference (8 days), separate along PC1. Nevertheless, the genotypes show a consistent ordering (Apo → IR64 → Ri-7, from left to right) at each TP. The loading plot

of the first two PCs (Supplementary Fig. S3B) shows which variables discriminate younger (e.g. high values of Thr, Glu, Ser, Ala, Asn, Asp) and older plants (e.g. high values of phosphate, sulphate, Ile, Leu, sucrose, trehalose, glucose). To normalize the effect, although limited, of plant age and genotype on the variables' levels of the samples and, therefore, truly assess drought-induced changes, we decided to use fold-change values (drought over control) of the variables in all further statistical analyses.

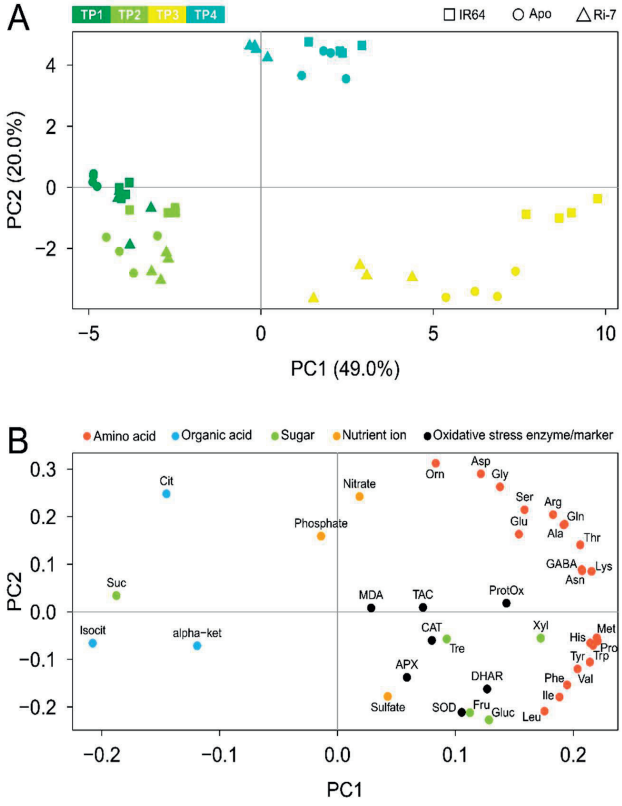


Fig. 4. Principle component analysis of the effect of transient drought on metabolites, nutrient ions and oxidative stress markers/enzymes of three rice genotypes.

Principal component analysis score plot (A) based on the fold change values (drought over control) of metabolites, nutrients and oxidative stress markers/enzymes of the biological replicates of IR64 (square), Apo (circle) and Ri-7 (triangle) at the four time points (TP1, TP2, TP3, TP4). Biological replicates are colored according to the time point of sampling (TP1: dark green; TP2: light green; TP3: yellow; TP4: light blue). Principal component 1 (PC1) explains 49.0% of the samples' variance, while principal component 2 (PC2) explains 20.0% of the samples' variance. Loading plot (B) of the 39 variables colored based on their class (amino acid: red; organic acid: light blue; sugar: green; nutrient ion: orange; oxidative stress marker/enzyme: black).

PCA based on these fold-change values showed that the first two PCs explained 69% of the total variation, with PC1 alone explaining 49% (Fig. 4A). Along PC1 the samples are distributed from left to right in accordance with the increasing severity of drought from TP1 to TP3. On this PC, TP1 (mild stress) samples of the three genotypes did not completely separate from the ones of TP2 (mild-severe stress). Samples of TP3 (severe stress) clearly separated from the others and displayed a wide distribution with Ri-7 replicates being closest to TP1 and TP2 samples (suggesting they are least stressed), followed by Apo and IR64 replicates (most stressed), respectively. Samples of TP4 (stress recovery) are positioned in between TP2 and TP3 along PC1. All this suggest that PC1 represents the metabolic and oxidative stress signature of drought stress.

Loadings of the variables on PC1 (Fig. 4B) showed that less severe drought stress (TP1 and TP2) is associated with elevated tricarboxylic acid (TCA) cycle intermediates (citrate, isocitrate and α-ketoglutarate) and sucrose. In contrast, severe drought stress (TP3) is associated with elevated levels of almost all amino acids (particularly Pro, Met, Lys, Trp, His, Tyr, Val, Asn, Thr and GABA) and oxidative stress markers/enzymes, as well as glucose, fructose and trehalose. Interestingly, along PC2 (Fig. 4A), explaining 20.0% of the total variation, there is a clear separation between drought stressed (TP1, TP2 and TP3) and re-watered samples (TP4), suggesting that PC2 represents metabolic and oxidative stress differences between drought stress and recovery. Loadings of the variables on PC2 (Fig. 4B) showed that increased levels of specific amino acids (mainly Orn, Asp, Gly and Ser), citrate, nitrate and phosphate are associated with stress recovery (TP4).

Stress-induced genotypic responses in metabolites, nutrient ions and oxidative stress status

To identify the drought-induced genotypic changes in primary metabolism and oxidative stress status during stress imposition and after re-watering, we analysed if the response to drought relative to the control was significant ($P < 0.05$) for every variable, genotype and TP (Table 1). Among the significant changes, we focused on the variables showing high deviations from control (fold-change decrease < 0.75 or increase > 1.5). Overall, the number of variables showing a significant deviation from control gradually increased between TP1 (18) and TP2 (32) before dramatically increasing at TP3 (80) and decreasing (59) again after re-watering (TP4).

At mild (TP1) and mild-severe (TP2) stress intensities, just few amino acids showed a significant response to drought and almost only a decrease relative to the control. A marked decrease (~0.5-fold or lower) in amino acids associated with photorespiration (Gly and Ser) and nitrogen remobilisation (Gln and Asn) was observed at TP2, especially in Apo and, to a lesser extent, in Ri-7. Different from the two earlier TPs, almost all the amino acids strongly increased at severe drought (TP3) with the highest fold-change values detected in IR64, followed by Apo and Ri-7. In particular, the stress-responsive amino acid Pro increased massively (almost 100-fold) in IR64, but less in Apo (55-fold) and Ri-7 (15-fold). In IR64, GABA showed the highest increase among all metabolites (115-fold) whereas it increased 10-fold less in Apo and Ri-7. After re-watering (TP4), the majority of the amino acids still showed increased values from control, but to a lower extent and with less distinct genotypic differences than at TP3. Nevertheless, it is interesting to note that, at TP4, the stress-responsive amino acid Pro was still significantly higher in the stressed IR64 plants, but not in the other two genotypes. Surprisingly, Orn, that did not change from control values under drought (TP1-TP3), markedly increased only after re-watering (TP4), showing the highest fold-change in Ri-7, followed by Apo and IR64.

The drought response of the organic acids followed the opposite trend of the amino acids. Except for a moderate (~2- to ~3-fold) increase of α-ketoglutarate at TP1 (in Apo) and TP2 (in IR64), no significant changes in organic acids were observed until severe stress (TP3) when the levels of TCA cycle intermediates, citrate and isocitrate, strongly decreased (fold < 0.5) in all genotypes. A similar strong decrease for isocitrate was observed also after re-watering (TP4) in all three genotypes.

Differently for the previous two classes of metabolites, sugars displayed a constant increase from mild (TP1) to severe (TP3) stress. At TP1, this increase was already quite marked (~4-fold) in Ri-7 for fructose and glucose with the latter showing an increase in IR64 too, albeit less strong (~2-fold). At TP2, the three genotypes displayed a high and quite similar increase in glucose (7- to 11-fold) and fructose (6- to 12-fold) compared with the control. These values increased even more at TP3 before decreasing after re-watering (TP4) when the two sugars still showed a higher value in the stressed (and re-watered) IR64 and Apo plants, but not in Ri-7. Sucrose did not show any important change from control at the four TPs whereas the hemicellulose-derived sugar xylose displayed a similar increase (7- to 10-fold) as the other sugars in all genotypes at severe drought (TP3) which decreased again at TP4 (2- to 4-fold).

Table 1. The effect of drought on the levels of primary metabolites, nutrient ions and oxidative stress markers/enzymes in the leaves of the three genotypes. Fold change (drought over control) of the mean values (Fold) of the 39 variables at each time point (TP1, TP2, TP3, TP4). Student's t-test *P*-value (*P*) of each variable is reported next to the fold change value. Fold changes that were significant (t-test $P < 0.05$) are highlighted in different colors (see color scale below).

Metabolic or redox state class		TP1						TP2						TP3						TP4					
		IR64		Apo		Ri-7		IR64		Apo		Ri-7		IR64		Apo		Ri-7		IR64		Apo		Ri-7	
		Fold	*P*	Fold	*P*	Fold	*P*	Fold	*P*	Fold	*P*	Fold	*P*	Fold	*P*	Fold	*P*	Fold	*P*	Fold	*P*	Fold	*P*	Fold	*P*
BCAAs	Leu	1.21	0.018	0.89	0.016	1.33	0.037	1.18	0.039	1.83	0.005	1.22	0.076	10.95	0.000	8.86	0.003	4.96	0.008	0.98	0.700	0.84	0.011	0.92	0.238
	Ile	1.20	0.017	0.89	0.024	1.41	0.010	1.31	0.010	1.71	0.010	1.17	0.128	19.24	0.000	12.45	0.003	4.83	0.010	1.13	0.154	1.03	0.632	0.98	0.715
	Val	1.02	0.681	0.63	0.005	1.42	0.006	1.26	0.012	1.36	0.080	1.22	0.063	30.43	0.001	17.15	0.003	5.88	0.010	2.16	0.003	1.77	0.021	1.55	0.004
AAAs	Phe	1.16	0.020	1.02	0.706	1.08	0.483	1.18	0.059	1.34	0.007	1.01	0.912	20.23	0.001	11.66	0.005	4.39	0.030	1.21	0.157	1.06	0.640	0.96	0.528
	Trp	1.07	0.170	0.96	0.521	1.46	0.047	1.62	0.018	1.58	0.039	1.49	0.006	33.18	0.001	29.75	0.004	11.80	0.032	4.51	0.022	7.54	0.001	1.11	0.208
	Tyr	1.04	0.554	0.98	0.808	1.06	0.477	1.03	0.623	1.34	0.000	0.95	0.508	7.47	0.002	5.18	0.005	2.57	0.033	1.35	0.004	1.19	0.063	1.14	0.068
Stress responsive AAs	Orn	0.97	0.562	0.85	0.097	1.07	0.111	1.01	0.864	0.90	0.147	0.98	0.702	2.34	0.000	1.14	0.018	0.95	0.143	5.43	0.000	8.71	0.001	9.32	0.000
	Pro	0.98	0.750	0.85	0.062	0.93	0.370	1.16	0.167	1.05	0.530	1.05	0.535	97.51	0.001	54.97	0.001	15.42	0.053	8.32	0.011	3.83	0.064	1.26	0.007
	GABA	1.07	0.783	0.69	0.120	1.34	0.297	0.74	0.046	0.85	0.380	1.13	0.424	115.10	0.001	6.11	0.007	3.72	0.009	10.57	0.001	7.17	0.006	4.43	0.000
Photorespiratory AAs	Gly	0.95	0.586	0.46	0.013	0.72	0.143	0.98	0.829	0.54	0.020	0.71	0.009	4.02	0.000	1.49	0.017	0.67	0.108	5.97	0.000	4.45	0.017	5.52	0.001
	Ser	0.86	0.027	0.92	0.131	0.73	0.029	0.79	0.049	0.48	0.007	0.51	0.000	1.65	0.004	1.66	0.009	1.67	0.041	2.56	0.000	2.23	0.001	2.32	0.000
N remobilisation AAs	Gln	0.88	0.372	0.80	0.051	0.96	0.910	1.17	0.481	0.34	0.012	0.55	0.069	48.41	0.002	18.18	0.004	3.33	0.023	32.43	0.001	36.19	0.002	16.64	0.000
	Glu	1.02	0.485	1.01	0.735	0.92	0.137	1.27	0.023	0.78	0.012	0.75	0.001	1.44	0.002	1.34	0.002	1.45	0.002	1.68	0.000	1.22	0.002	1.54	0.000
	Asn	0.62	0.007	0.47	0.008	0.93	0.592	0.63	0.003	0.34	0.015	0.84	0.003	39.78	0.001	45.77	0.002	31.41	0.014	49.24	0.001	33.26	0.002	4.67	0.005
	Asp	1.04	0.699	1.09	0.127	0.64	0.027	1.01	0.903	0.62	0.009	0.46	0.001	2.18	0.000	1.39	0.033	0.86	0.218	4.42	0.000	3.82	0.000	2.91	0.000
Other AAs	Ala	0.87	0.076	0.65	0.003	0.79	0.107	0.82	0.024	0.64	0.000	0.79	0.020	5.05	0.003	2.12	0.000	1.36	0.010	3.08	0.000	3.41	0.003	3.00	0.002
	Thr	0.71	0.014	0.82	0.008	0.85	0.061	0.95	0.441	0.58	0.009	0.83	0.017	7.63	0.001	3.98	0.005	3.01	0.021	4.12	0.001	4.52	0.000	4.28	0.000
	Met	0.88	0.055	0.82	0.002	0.99	0.691	0.92	0.025	0.95	0.246	0.96	0.307	4.76	0.001	4.07	0.003	1.89	0.032	1.77	0.002	1.29	0.017	1.20	0.005
	His	1.06	0.584	0.95	0.246	1.22	0.103	1.23	0.208	1.12	0.105	1.09	0.389	16.37	0.002	13.34	0.007	4.76	0.074	2.61	0.114	4.01	0.024	1.08	0.378
	Arg	0.90	0.187	0.72	0.008	0.96	0.235	1.08	0.381	0.76	0.009	0.95	0.127	10.76	0.005	4.36	0.030	2.16	0.017	8.36	0.009	11.71	0.000	7.45	0.000
	Lys	1.00	0.976	0.89	0.062	1.11	0.026	1.20	0.010	1.19	0.112	1.12	0.083	11.55	0.001	6.78	0.010	2.01	0.000	4.54	0.010	5.33	0.000	2.51	0.001
Sugars	Suc	1.10	0.018	1.10	0.083	1.17	0.056	1.03	0.425	1.07	0.077	0.96	0.386	0.81	0.004	0.76	0.003	0.80	0.004	0.79	0.000	0.97	0.392	1.03	0.412
	Gluc	1.95	0.001	1.06	0.692	4.34	0.005	11.15	0.008	10.34	0.026	7.21	0.001	18.47	0.001	22.69	0.001	17.80	0.000	3.56	0.014	5.51	0.018	0.88	0.265
	Fru	1.19	0.417	1.32	0.327	4.41	0.003	9.42	0.010	11.55	0.018	5.74	0.000	10.90	0.003	24.80	0.009	12.08	0.000	4.38	0.018	5.69	0.008	0.62	0.247
	Tre	1.26	0.160	0.80	0.369	1.63	0.078	1.62	0.065	1.24	0.062	2.18	0.000	1.55	0.115	3.41	0.006	2.52	0.025	0.92	0.759	3.48	0.004	2.04	0.001
	Xyl	0.60	0.143	0.80	0.160	1.39	0.268	3.36	0.112	1.32	0.544	4.52	0.100	7.20	0.000	7.31	0.001	10.63	0.019	2.89	0.004	4.15	0.000	2.70	0.108
Organic acids	Cit	1.09	0.285	1.45	0.082	1.02	0.894	1.11	0.457	1.03	0.660	0.67	0.015	0.44	0.005	0.35	0.001	0.46	0.008	1.43	0.022	1.39	0.008	1.24	0.010
	Isocit	1.17	0.059	1.26	0.106	1.00	0.977	1.88	0.075	1.00	0.992	1.05	0.730	0.09	0.004	0.29	0.002	0.55	0.036	0.37	0.008	0.32	0.000	0.51	0.001
	alpha-ket	0.75	0.354	3.18	0.001	2.68	0.125	1.75	0.015	1.56	0.095	2.39	0.018	0.48	0.126	1.02	0.936	1.23	0.281	0.63	0.089	0.73	0.127	0.75	0.002
Nutrient ions	Sulfate	1.41	0.027	0.95	0.552	1.25	0.087	1.16	0.534	1.16	0.282	1.64	0.002	1.70	0.013	1.25	0.092	1.62	0.002	0.93	0.434	0.89	0.068	1.31	0.006
	Nitrate	0.38	0.007	1.34	0.073	0.38	0.193	1.15	0.808	0.74	0.196	0.89	0.896	0.75	0.581	1.02	0.963	0.10	0.204	4.12	0.008	2.25	0.079	5.49	0.012
	Phosphate	0.61	0.018	0.75	0.000	0.61	0.000	0.39	0.001	0.56	0.000	0.44	0.001	0.56	0.022	0.40	0.000	0.58	0.002	0.90	0.327	0.77	0.016	0.42	0.004
Redox state marker	TAC	1.36	0.409	1.09	0.814	1.02	0.945	1.11	0.557	1.22	0.422	1.41	0.163	2.06	0.052	3.15	0.033	3.09	0.062	1.23	0.501	1.54	0.162	1.24	0.046
	ProtOx	1.04	0.584	1.00	0.949	0.97	0.545	1.12	0.238	1.03	0.832	1.19	0.047	1.73	0.013	1.12	0.312	1.14	0.094	1.27	0.160	1.07	0.318	1.15	0.043
	MDA	0.79	0.073	1.06	0.677	1.00	0.989	1.76	0.042	1.94	0.250	2.43	0.169	2.27	0.092	3.08	0.114	3.77	0.155	1.27	0.356	2.38	0.172	2.11	0.161
Redox state enzyme	DHAR	0.88	0.044	1.11	0.512	0.81	0.188	1.11	0.272	0.86	0.235	0.94	0.705	1.17	0.120	2.61	0.001	1.68	0.007	1.03	0.842	1.14	0.619	0.67	0.002
	APX	1.08	0.444	1.26	0.110	0.75	0.150	0.95	0.907	1.16	0.770	1.28	0.342	1.20	0.059	1.47	0.027	1.34	0.380	0.82	0.156	0.97	0.658	1.09	0.657
	CAT	0.65	0.036	0.88	0.403	0.69	0.495	0.81	0.123	1.05	0.714	2.47	0.027	1.69	0.003	1.36	0.001	0.86	0.256	0.91	0.701	0.95	0.794	1.36	0.289
	SOD	1.04	0.732	1.16	0.371	0.91	0.642	0.83	0.027	0.87	0.313	1.07	0.535	1.20	0.158	1.97	0.048	1.32	0.187	0.87	0.460	0.82	0.254	0.79	0.007

Mean fold change of drought over control replicates:

< 0.5 | 0.5 ~ 0.75 | 1.5 ~ 5.0 | 5.0 ~ 50 | > 50

Among the nutrient ions, phosphate decreased almost equally in the three genotypes from TP1 to TP3 whereas, at TP4, it was no longer affected in the re-watered plants except for Ri-7. Only after re-watering, there was a change in nitrate, which was increased in stressed and re-watered IR64 and Ri-7.

Considering the oxidative stress markers and enzymes, few changes were observed in the first two TPs with a decrease of catalase (CAT) in IR64 at mild stress (TP1) and an increase in Ri-7 at mild-severe stress (TP2). IR64 also showed an increase in malondialdehyde (MDA) at TP2. At severe drought (TP3), oxidative stress markers and enzymes showed the strongest upregulation by drought. Protein oxidation (ProtOx) increased in IR64 and total antioxidant capacity (TAC) increased in Apo. Among the antioxidant enzymes, dehydroascorbate reductase (DHAR) showed a higher activity in Apo and Ri-7 but not in IR64. Apo also displayed an increase in the activity of superoxide dismutase (SOD) while CAT activity increased in IR64. After re-watering (TP4), oxidative stress markers and enzymes were not different from the control, except for a slightly lower activity of DHAR in Ri-7.

Overall, these results indicate that the different classes of primary metabolites (amino acids, sugars, organic acids) and the oxidative stress markers/enzymes showed a similar kind of response to increasing drought and re-watering in the three genotypes. However, the intensity of this response was different between the genotypes and, considering each TP separately, we identified genotype-specific responses of single metabolites and oxidative stress markers/enzymes.

Metabolic signatures of leaf senescence

Gln and Asn are two amino acids involved in nitrogen recycling and export from senescent leaves (Chrobok *et al.*, 2016). To evaluate the level of leaf senescence among the genotypes, we calculated the ratio of glutamine to glutamate (Gln/Glu) and asparagine to aspartate (Asn/Asp) (Watanabe *et al.*, 2013) for each genotype at the four TPs (Fig. 5). Under control conditions, the Gln/Glu and Asn/Asp ratios were consistently low (ratio < 0.3) for all genotypes at all TPs. Both ratios did not respond to the drought treatment at mild (TP1) and mild-severe (TP2) stress intensity. At severe stress (TP3), however, the Gln/Glu ratio strongly increased in IR64 and to a lesser extent in Apo and Ri-7 whereas the Asn/Asp ratio showed a similar marked increase in all genotypes. After re-watering (TP4), both ratios decreased again with stressed

and re-watered Ri-7 showing the lowest values (almost the same as for control plants), followed by stressed and re-watered Apo and IR64 that still had a higher Asn/Asp ratio than the control plants.

The Gln/Glu and Asn/Asp values suggest that premature drought-induced leaf senescence only occurred in all genotypes under severe drought stress and decreased again upon re-watering. The values of these ratios at TP3 and TP4 suggest a higher degree of leaf senescence in IR64, and much lower in Apo and Ri-7.

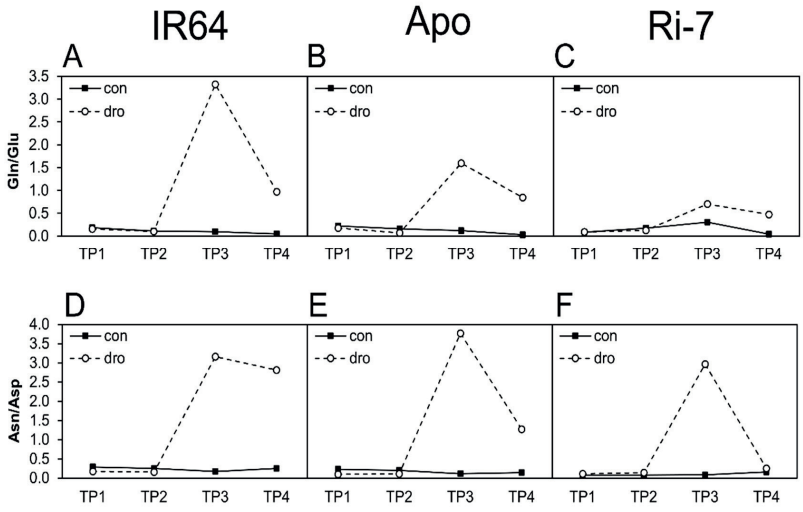


Fig. 5. Drought-induced variation of the leaf senescence markers Gln/Glu and Asn/Asp ratios in three genotypes.
Mean values of the glutamine to glutamate (Gln/Glu) and asparagine to aspartate (Asn/Asp) ratios at the four time points (TP1, TP2, TP3, TP4) under control (con) and drought (dro) conditions in IR64 (A, D), Apo (B, E) and Ri-7 (C, F).

Comparison of the vegetative and reproductive stage

The same three genotypes were also included in a large field trial conducted at the International Rice Research Institute (IRRI), Philippines, during the 2013 dry season (Kadam *et al.*, 2018). In the trial, 300 indica rice genotypes were grown in two separate fields, one served as control (flooded) and in the other drought stress (14 days of water withholding) was applied at the reproductive stage (50% flowering) before re-watering all the accessions to let them reach maturity for harvest. Melandri *et al.* (2019) describe how the flag leaf metabolic and

oxidative stress profiles of the drought stressed accessions were used to efficiently predict grain yield loss – due to drought - at harvest. In the field trial, drought exposure in the reproductive stage induced a decrease in grain yield of -27.1% in IR64 and -20.3% in Ri-7 whereas Apo did not show grain yield reduction but a slight increase of +5.8%. It is noteworthy that drought-induced grain yield loss in Ri-7 in the reproductive stage was similar to that of IR64, whereas the latter showed a more tolerant biomass response (more similar to Apo) when drought stress was imposed during vegetative development (Fig. 2G-I).

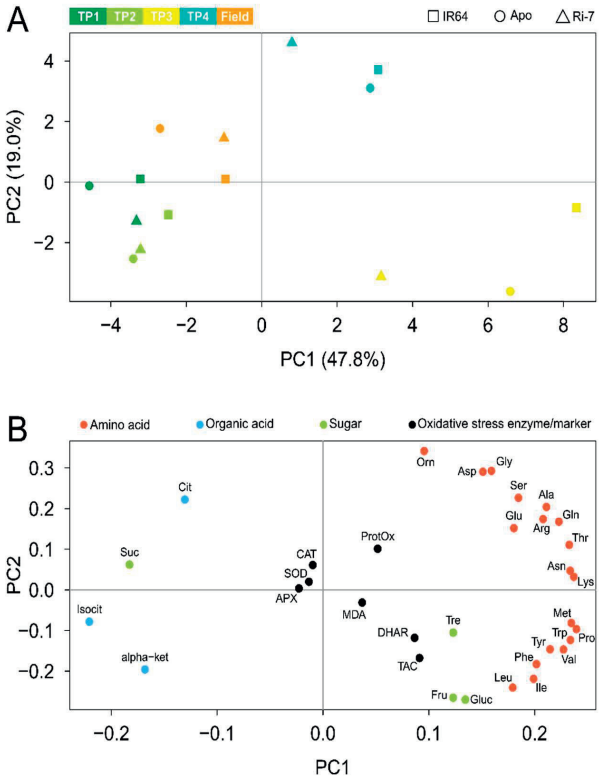


Fig. 6. Principle component analysis of the effect of drought at the vegetative and reproductive stage on metabolites and oxidative stress markers/enzymes of three rice genotypes.
Principal component analysis score plot (A) based on the fold change values (drought over control) of metabolites and oxidative stress markers/enzymes of the biological replicates of IR64 (square), Apo (circle) and Ri-7 (triangle). Samples at the vegetative stage were collected at the four time points (TP1, TP2, TP3, TP4) and samples at the reproductive stage in one time point (Field). Samples are colored as follows, TP1: dark green; TP2: light green; TP3: yellow; TP4: light blue; Field: orange. Principal component 1 (PC1) explains 47.8% of the samples' variance, while principal component 2 (PC2) explains 19.0% of the samples' variance. Loading plot (B) of the 36 variables colored based on their class (amino acid: red; organic acid: light blue; sugar: green; oxidative stress marker/enzyme: black).

Using PCA, we compared the drought-induced differences in leaf metabolic and oxidative stress profiles of the three genotypes in the field, during the reproductive stage, and in the climate room, during the vegetative stage. Since we only had average values for the field experiment (Melandri *et al.*, 2019), in the PCA we used the mean fold-change values also for the data from the climate room. Similar as in Fig. 4, the first two PCs explained a high percentage of the total variation, with PC1 alone explaining 47.8% (Fig. 6A). The first two PCs determined a distribution of the vegetative stage samples almost identical to the one in Fig. 4A, which was further supported by the very similar variables' loadings (Fig. 4B and Fig. 6B). As discussed above for Fig. 4, also in Fig. 6 PC1 represents the metabolic and oxidative stress signature of drought stress (from left to right). The position of the field samples along PC1 is between the ones of TP2 and TP3, but closer to TP2, with Apo more on the left (lower stress) and, IR64 and Ri-7 more shifted to the right (more stress) (Fig. 6A). These results suggest that the drought stress intensity in the field in the reproductive stage was in between the mild-severe (TP2) and severe (TP3) drought stress observed for the same genotypes at the vegetative stage under climate room conditions. Furthermore, our results agree with the observation that drought stress intensity in the field trial was moderately severe (-46 kPa, from tensiometers reads) (Kadam *et al.*, 2018).

Table 2 shows the drought-induced fold-changes of the 36 primary metabolites and oxidative stress markers/enzymes detected in the three genotypes at the reproductive stage. Confirming the PCA (Fig. 6A), the amino acids showed an overall increase stronger than at TP2 during the vegetative development, but lower than at TP3 (Table 1 and 2). Interestingly, IR64 and Ri-7 displayed a similar number of amino acids with an increased value (9 and 8, respectively) whereas in Apo only few amino acids increased (2). Considering the organic acids, in the reproductive stage the three genotypes showed a reduction similar but less intense than observed under severe drought (TP3) in the vegetative stage with the exception of Ri-7 that showed similar values. Surprisingly, sugars did not show any important deviation from control in the reproductive stage (Table 2). This is in contrast with what we observed in all genotypes at mild-severe (TP2) and severe (TP3) drought stress during vegetative development (Table 1). Interestingly, oxidative stress markers and enzymes (with the exception of TAC) showed an overall strong upregulation in all genotypes in the reproductive stage, even stronger than observed under severe drought (TP3) in the vegetative stage (Table 1 and 2). MDA increased

in IR64 and Ri-7, and ProtOx in Ri-7 and Apo. Apo displayed a very strong increase in all the antioxidant enzymes, and it was the only one showing an increase in DHAR and ascorbate peroxidase (APX). CAT and SOD increased strongly in all three genotypes with Apo displaying the highest values for the first enzyme and Ri-7 for the second.

Table 2. The effect of drought on the levels of primary metabolites and oxidative stress markers/enzymes in the leaves of the three genotypes at the reproductive stage.
Fold change (drought over control) of the mean values (Fold) of the 36 variables for the three genotypes. Fold changes are highlighted in different colors (see color scale in Table 1).

		FIELD – Flowering		
Metabolic or oxidative stress class		**IR64**	**Apo**	**Ri-7**
		Fold	**Fold**	**Fold**
BCAAs	Leu	2.52	1.19	1.46
	Ile	1.62	0.81	1.26
	Val	0.94	0.72	0.97
AAAs	Phe	2.69	1.30	1.55
	Trp	2.53	1.02	1.16
	Tyr	1.78	1.14	1.73
Stress responsive AAs	Orn	1.96	1.79	4.84
	Pro	2.59	1.32	1.27
Photorespiratory AAs	Gly	0.97	1.49	1.95
	Ser	0.89	1.20	1.18
N remobilisation AAs	Gln	1.23	1.20	2.30
	Glu	1.21	1.09	0.97
	Asn	0.97	0.93	0.83
	Asp	1.40	1.04	1.02
Other AAs	Ala	1.41	1.64	1.72
	Thr	1.20	0.99	0.90
	Met	1.46	0.78	1.64
	Arg	0.55	1.12	1.52
	Lys	1.61	0.80	0.90
Sugars	Sucrose	0.92	1.40	0.80
	Glucose	1.19	1.23	1.54
	Fructose	1.12	1.29	1.27
	Trehalose	1.53	1.03	0.98
Organic acids	Citrate	1.25	0.79	0.48
	Isocitrate	0.58	0.60	0.36
	α-ketoglutarate	1.21	0.88	0.87
Oxidative stress marker	TAC	1.42	0.63	0.58
	ProtOx	1.42	2.29	1.65
	MDA	2.29	0.93	2.33
Oxidative stress enzyme	DHAR	1.48	2.56	1.10
	APX	1.09	6.72	1.15
	CAT	4.16	6.92	3.91
	SOD	3.26	5.89	9.66

In conclusion, the metabolic and oxidative stress profiles of the genotypes in the reproductive stage showed a similar response to drought in amino acids (increase) and organic acids (decrease) as when exposed to mild-severe to severe drought stress during vegetative development, but a different response in the sugars (no change) and in the activity of antioxidant enzymes (strong increase).

DISCUSSION

Drought stress in the vegetative stage has a stronger impact on the 'water-spender' IR64 than on the 'water-savers' Apo and Ri-7

The transpiration differences between the three genotypes suggest that the high and sustained water consumption of IR64 under water-limiting conditions exposed the genotype to a more prolonged and severe drought stress than the other two genotypes, which displayed a more conservative water use (Fig. 1). These results are in agreement with a recent study by Ouyang *et al.* (2017) who showed that IR64 displays a higher stomatal conductance and a lower transpiration efficiency than Apo and Ri-7, both under well-watered and mild drought stress conditions. Our findings also point towards the constitutive higher leaf biomass in IR64 than in Apo and Ri-7 under optimal growth conditions (Fig. 2D-F and Supplementary Table S1) as an additional cause for its higher water loss through transpiration, especially during the first days of water withholding (until TP1) (Fig. 1).

The higher water consumption of IR64 compared to Apo and Ri-7, which is also maintained after the initiation of the water withholding, is therefore likely the cause of the extremely strong dehydration (to about 25% RWC) suffered by the lowland genotype after 10 days (TP3) of stress (Fig. 2A). Maintenance of a high cellular hydration under drought is crucial to maintain essential metabolic functions under stressful conditions (Obidiegwu, 2015). The lower RWC in IR64 than in Apo and Ri-7 (~70% RWC) at TP3 (Fig. 2B-C) might have increased the incidence of leaf cell death (leaf necrotic lesions were visible in all genotypes at this TP) more in the first genotype than in the other two. This hypothesis is supported by the stronger leaf biomass loss observed in IR64 than in Apo and Ri-7 after re-watering (TP4) (Fig. 2D-I and Fig. 3). Our leaf biomass reduction results indirectly confirm the different levels of drought tolerance (lower for lowland and higher for aerobic and upland rice) reported in the literature for these three genotypes (Atlin *et al.*, 2006; Venuprasad *et al.*, 2012; Mackill and Khush, 2018).

Relationships between growth dynamics, primary metabolism and oxidative stress status

Leaf growth reduction in response to drought is one of the mechanisms used by plants to limit the expansion of their evaporation surface and it generally occurs before reduction of photosynthesis (Hummel *et al.*, 2010; Claeys and Inze, 2013). Different rates of leaf expansion under drought were observed before in rice genotypes with contrasting drought tolerance

(Cabuslay *et al.*, 2002; Parent *et al.*, 2010) and confirmed by our leaf growth results (Fig. 2G-I). In our study, the absence of increased values for photorespiratory amino acids (Gly and Ser) and oxidative stress markers/enzymes in all genotypes at both TP1 and TP2 (Table 1) support the presence of an active photosynthetic metabolism until severe stress (TP3). Indeed, a drought-induced reduction of photosynthesis should have increased the activity of the photorespiratory pathway, resulting in higher production of Gly and Ser (Maurino and Peterhansel, 2010; Bauwe *et al.*, 2010; Hodges *et al.*, 2016), and/or should have enhanced ROS generation and oxidative stress in leaf cells (Suzuki *et al.*, 2012; Noctor *et al.*, 2014). The consequence of maintaining photosynthetic active metabolism coupled with reduced growth often results in an increase in the concentration of carbohydrates in the leaves (Muller *et al.*, 2011; Blum, 2017). At mild (TP1) and mild-severe (TP2) drought stress, the analysis of sugars and leaf growth dynamics in the context of an active photosynthetic metabolism revealed a different response to drought between the three genotypes.

Reduced leaf growth occurred in IR64 at the earliest stage of stress (TP1) and only at more severe drought intensities in Apo and Ri-7 (Fig. 2G-I). Surprisingly, the early reduction of leaf biomass in IR64 was not associated with an increase in the concentration of the main sugars (Table 1). A possible explanation for this might be that, at early-mild drought (TP1), the sampled mature leaves of IR64 were still exporting carbon to the growing sink organs. Similarly, Apo did not display any increase in sugars at early-mild drought (Table 1), but this might simply be determined by the low level of stress experienced by this water saving genotype (Fig. 1). Indeed, at TP1 Apo displayed no leaf biomass reduction (Fig. 2E and 2H). Similar to Apo, Ri-7 did not show a reduction in leaf biomass at early-mild drought (Fig. 2F and 2I) but its fructose and glucose levels markedly increased (Table1). Sugar-mediated osmotic adjustment in response to drought helps to sustain cell turgor, stomatal opening and photosynthesis, but at the expense of growth (Blum, 2017). In Ri-7, the increased accumulation of fructose and glucose at early-mild drought (TP1) might represent a strategy of acclimation to drought that does not affect growth and is possibly responsible for the tolerance of this genotype.

Progression of drought severity from mild (TP1) to mild-severe (TP2) resulted in a sudden and strong increase of fructose and glucose (~10-fold) in Apo and IR64 and less marked (~6-fold) in Ri-7 (Table 1). Despite the similar strong increase of sugars among the genotypes, they

displayed a very different leaf growth response to drought at TP2. Fructose and glucose accumulation in Apo was associated with its first marked reduction in leaf growth (Fig. 2H) and, therefore, it might represent an osmotic adjustment originated from the carbon surplus deriving from the maintained activity of photosynthesis (Muller *et al.*, 2011; Blum, 2017). Surprisingly, at TP2 IR64 did not display a further increase of leaf biomass loss from TP1 (Fig. 2G), even though the higher severity of the stress induced a strong accumulation of carbohydrates in this genotype (Table1). A possible explanation for this may reside in the reduction of stem weight that started at TP2 in IR64 only, and at later stages of stress in Apo and Ri-7 (at TP3 and TP4, respectively) (Supplementary Table S1). In rice, the excess of photo-assimilates produced in leaves is stored in the stem as carbohydrate reserves that can be used to buffer leaf performance under stress at the reproductive stage (Yang *et al.*, 2001; Morita and Nakano, 2011; Wang *et al.*, 2017). Our results suggest that, in IR64, non-structural carbohydrates stored in the stems were remobilised (starting from TP2) to stabilise leaf growth and provide osmotic protection under drought already at the vegetative stage. Nevertheless, the earlier use of these stem reserves, compared with the other two genotypes, might have undermined stem vigour of IR64 and contributed to the early reduction in tiller number, observed only in this genotype (Supplementary Table S1). Different from IR64 and Apo, the increase of fructose and glucose in Ri-7 at TP2 was not coupled with any reduction in the biomass of leaves, stems or roots (Fig. 2I and Supplementary Table S1). In Ri-7, the absence of leaf growth reduction and the simultaneous presence of active carbon assimilation at this TP further supports the hypothesis of a sugar-mediated osmotic adjustment to drought as the origin for the high levels of the two sugars (already observed at TP1) in the leaves of this genotype.

Prolonged and severe drought (10 days, TP3) caused a marked reduction in growth in all genotypes (Fig. 2G-I). This reduction was reflected in a state of strong metabolic alteration and associated with reduced photosynthesis, as displayed by the increased levels of Ser, Gly and oxidative stress markers/enzymes (Table 1). At this TP, the extremely high levels of almost all amino acids displayed by the three genotypes have been described before to occur in leaves of many crop species exposed to drought (Obata and Fernie, 2012; Krasensky and Jonak, 2012; Fàbregas and Fernie, 2019). This strong amino acid increase is associated with protein catabolism occurring during premature stress-induced leaf senescence (Araújo *et al.*, 2011;

Watanabe *et al.*, 2013; Hildebrandt *et al.*, 2015). The hypothesis of a strong catabolic activity at severe drought (TP3) is supported by the increased asparagine to aspartate (Asn/Asp) and glutamine to glutamate (Gln/Glu) ratios observed in the three genotypes at this sampling point (Fig. 5). In senescent leaves, protein degradation-derived aspartate and glutamate are converted to asparagine and glutamine to act as transport molecules for long-distance nitrogen remobilisation through the phloem (Watanabe *et al.*, 2013; Avila-Ospina *et al.*, 2014). A further confirmation of a high catabolic activity at severe drought (TP3) in the three genotypes is represented by the shared decrease in the level of two TCA cycle intermediates, isocitric and citric acid (Table 1). These two organic acids are involved in mitochondrial respiration (Shi *et al.*, 2016). Under severe drought, the decrease in leaf growth might have caused a reduction in mitochondrial respiration, resulting in lower production of isocitric and citric acid, because of an overall reduced energy availability for biosynthesis (Atkin and Macherel, 2008). Similar to TP2, under severe drought (TP3), the three genotypes displayed a comparable and very high accumulation of fructose and glucose (Table 1), but, differently from the previous TP, this sugar increase is likely the result of the initiation of starch degradation induced by reduced photosynthetic carbon assimilation (Wingler *et al.*, 2006; Pinheiro and Chaves, 2011; Stitt and Zeeman, 2012). Despite the similar and shared state of metabolic alteration observed among the three genotypes at TP3, the extent of this alteration was more severe in IR64, followed by Apo and Ri-7 as overall indicated by PCA (samples of TP3 in Fig. 4A) and, more specifically, by the marker of leaf senescence Gln/Glu (TP3 in Fig. 5A-C). In addition, the different levels of certain metabolites between the genotypes under severe drought (TP3) are indicative of their stress status. Among these metabolites, the highest increase of Pro was in the leaves of IR64 (~100-fold), and less in Apo (~55-fold) and even less in Ri-7 (~15-fold) (Table 1). Accumulation of Pro under drought stress has often been reported in the literature (Fàbregas and Fernie, 2019) and also in rice (Todaka *et al.*, 2017). Pro is thought to play a role in counteracting the enhanced generation of ROS to protect cellular functions against the damage caused by dehydration (Verslues and Juenger, 2011; Krasensky and Jonak, 2012; Nakabayashi and Saito, 2015). A similar, even more marked trend was displayed by GABA, an amino acid that strongly accumulates, like Pro, during many abiotic stresses (Obata and Fernie, 2012; Fàbregas and Fernie, 2019). Even if its role under stress remains unclear, this amino acid has been proposed as a regulator of osmolarity and coordinator of the carbon-nitrogen

balance under carbon limitation (Michaeli and Fromm, 2015; Hildebrandt *et al.*, 2015; Fàbregas and Fernie, 2019).

A different level of oxidation and antioxidant activity was specifically observed between the three genotypes only under severe drought (Table 1). It has been shown that under persistent water-limitation, drought-induced stomatal closure increases ROS production as a result of excess of light and reduced CO_2 assimilation (Asada, 2006; Miller *et al.*, 2010). Among the three genotypes, Apo displayed simultaneously high levels of DHAR, SOD and TAC suggesting a stronger antioxidant capacity under drought than in the other two genotypes that showed only a higher activity of DHAR (Ri-7) and CAT (IR64). DHAR regenerates oxidized ascorbate with reduced glutathione as electron donor (Das and Roychoudhury, 2014) through the ascorbate-glutathione pathway (Foyer and Noctor, 2011). Increased activity of DHAR and SOD under drought stress was described before in rice seedlings of different cultivars (Selote and Khanna-Chopra, 2004; Sharma and Dubey, 2005). In plants, SOD converts highly oxidative superoxide (O_2^-) to less harmful H_2O_2 (Halliwell, 2006). Increased activity of SOD, present only in Apo, could represent a safeguarding mechanism to generate H_2O_2 that, in turn, triggers more effective antioxidant defences. This would explain the finding of increased TAC levels, once more, only in Apo. TAC represents a proxy for non-enzymatic metabolites of the antioxidant defence system, like ascorbate, glutathione, carotenoids, flavonoids and tocopherols (Sharma *et al.*, 2012).

After re-watering, the differences in the leaf metabolic and oxidative stress profiles of re-watered plants are more similar to the control than under severe drought (TP3) and with less marked differences between the genotypes (Fig. 4A; Table 1). This shows that a similar recovery from stress occurred in all genotypes and suggests that their leaf biomass reduction (-77.6% in IR64, -58.4% in Ri-7 and -50.2% in Apo) after re-watering is entirely due to the differences in the metabolic and antioxidative responses during drought.

Different drought-induced metabolic and antioxidative responses between vegetative and reproductive stage in the three genotypes

Comparative analysis of metabolic and oxidative stress profiles of the genotypes in the two different phenological stages (PCA in Fig. 6) revealed that samples of the field trial (reproductive stage) experienced a stress intensity higher than mild-severe (TP2) stress at the

vegetative stage but lower than severe (TP3). The flag leaf metabolic profiles of the three genotypes in the reproductive stage (Table 2) had a similar signature of drought-induced leaf senescence (increased amino acids, but less than at TP3) and of lower biosynthetic activity (decreased organic acids like at TP2) as during the vegetative stage. This signature was particularly marked in IR64 and Ri-7 but much less extreme in Apo. Different from TP2 and TP3 in the vegetative stage, the levels of sugars did not increase compared with the control in the reproductive stage (Table 1 and 2). These finding highlight the importance, for these three yield improved genotypes, of maintaining a stable sugar export under drought in the reproductive stage when the flag leaf is the most important source of assimilates for the developing panicles and thus for yield stability (Yoshida, 1972; Biswal and Kohli, 2013). Therefore, by behaving as a carbon exporter in a context of drought-induced stomatal closure and lower carbon availability, the flag leaves of the three genotypes were likely exposed to oxidative stress by the enhanced generation of ROS (Suzuki *et al.*, 2012; Noctor *et al.*, 2014). This hypothesis is confirmed by the high levels of the lipid peroxidation product MDA in IR64 and Ri-7, and by high ProtOx in Apo and Ri-7 (Table 2). All the three genotypes counteracted the oxidative stress by markedly increasing the activity of antioxidant enzymes, which displayed much higher values than at severe drought (TP3) during the vegetative stage (Table 1 and 2). Among the three genotypes, Apo, similar to the vegetative stage results, displayed the strongest antioxidant response under drought also in the reproductive stage. This response helped Apo to minimize grain yield loss, unlike IR64 and Ri-7 in which less antioxidant enzymes increased in activity. Surprisingly, grain yield loss in this stage was about the same for IR64 and Ri-7 while, in the vegetative stage, Ri-7 suffered much less growth reduction under drought than IR64 (Fig. 2G-I). This is likely due to the inability of Ri-7 to activate a protective leaf osmotic acclimation response to drought (in contrast to during vegetative development) while actively exporting sugars to the developing panicles. This finding confirms the prevalent role of leaf antioxidant enzyme activities for grain yield stability of rice under drought stress at the reproductive stage (Melandri *et al.*, 2019).

CONCLUSIONS

The aim of the present research was to determine how drought-induced leaf metabolic and oxidative stress responses in drought tolerant rice genotypes differed from a susceptible one to better understand the mechanisms of adaptation to drought during the vegetative and

reproductive stage. In the vegetative stage, a conservative water use coupled with sugar-mediated osmotic acclimation in Ri-7 (upland rice) or strong antioxidative response in Apo (aerobic rice) helped these two tolerant genotypes to limit drought-induced biomass loss compared with the susceptible genotype IR64 (lowland rice). In the reproductive stage, limitations in sugar-mediated osmotic adjustments (sugar were exported to panicles) under drought stress conferred a competitive advantage to the genotype showing the highest antioxidant power, Apo, that was the only one able to maintain a stable grain yield.

By enhancing our understanding of physiological, metabolic and antioxidant strategies associated with drought tolerance in rice, this study provides a framework for the exploration of the genetic control of these mechanisms and their exploitation in breeding for new varieties adapted to water-limited environments.

ACKNOWLEDGMENTS

This work is part of the Growing Rice like Wheat research programme financially supported by an anonymous private donor, via the Wageningen University Fund, for the first author's PhD fellowship. HA is a postdoctoral researcher of the Research Foundation – Flanders (FWO; 12U8918N).

SUPPLEMENTARY DATA

Supplementary data are available at:

https://drive.google.com/open?id=1rFJkrHHCb38Yd9nqsAemOZ82SH6sD28R

Fig. S1. The experimental setup.

Fig. S2. Effect of the treatment on soil water content.

Fig. S3. Global effects of plant age and genotype on values of leaf metabolites, nutrient ions and oxidative stress markers/enzymes in the control samples.

Table S1. Dry weight (DW) of leaves, stems, roots, and total biomass, leaf relative water content (RWC) and the number of tillers of the three genotypes under control and drought conditions at the four time points of sampling.

Table S2. Levels of the 39 metabolites, nutrient ions and oxidative stress markers/enzymes measured in leaf of control samples of the three genotypes at each time point.

Table S3. Levels of the 39 metabolites, nutrient ions and oxidative stress markers/enzymes measured in leaf of drought stressed samples of the three genotypes at each time point.

Supplementary Methods

CHAPTER

3

Biomarkers for grain yield stability in rice under drought stress

Giovanni Melandri, Hamada AbdElgawad, David Riewe, Jos A. Hageman, Han Asard, Gerrit T. S. Beemster, Niteen N. Kadam, S. V. Krishna Jagadish, Thomas Altmann, Carolien Ruyter-Spira, Harro J. Bouwmeester

Published in Journal of Experimental Botany (2019), https://doi.org/10.1093/jxb/rz221

ABSTRACT

Crop yield stability requires an attenuation of the reduction of yield losses caused by environmental stresses such as drought. Using a combination of metabolomics and high-throughput colorimetric assays, we analysed central metabolism and oxidative stress status in the flag leaf of 292 *indica* rice (*Oryza sativa*) accessions. Plants were grown in the field and were, at the reproductive stage, exposed to either well-watered or drought conditions to identify the metabolic processes associated with drought-induced grain yield loss. Photorespiration, protein degradation and nitrogen recycling were the main processes involved in the drought-induced leaf metabolic reprogramming. Molecular markers of drought tolerance and sensitivity in terms of grain yield were identified using a multivariate model based on the values of the metabolites and enzyme activities across the population. The model highlights the central role of the ascorbate-glutathione cycle, particularly dehydroascorbate reductase, in minimizing drought-induced grain yield loss. In contrast, malondialdehyde was an accurate biomarker for grain yield loss, suggesting that drought-induced lipid peroxidation is the major constraint under these conditions. These findings highlight new breeding targets for improved rice grain yield stability under drought.

KEY WORDS: *Oryza sativa*, drought, leaf primary metabolism, leaf oxidative stress status, reproductive stage, PLSR.

INTRODUCTION

Drought-induced closure of stomata not only reduces water loss, but also limits CO_2 diffusion into the leaf intercellular spaces and thus decreases photosynthetic carbon assimilation. This alteration results in the disruption of cellular homeostasis and leads to an enhanced generation of reactive oxygen species (ROS), mainly in the peroxisomes and chloroplasts (Suzuki *et al.*, 2012; Noctor *et al.*, 2014). In leaf tissues of C_3 plants exposed to light and under drought stress, peroxisomes are considered the major production site of hydrogen peroxide (H_2O_2), primarily because of the enhanced activity of the photorespiratory pathway (Noctor 2002). In chloroplasts, ROS production arises when excitation energy exceeds the level required for CO_2 assimilation (Asada 2006), a condition that is favoured by drought-induced stomatal closure (Miller *et al.*, 2010). Under persistent water-limited conditions, the cellular ROS scavenging capacity is exceeded by ROS production that, in turn, leads to oxidative damage that drives the cell into senescence (Halliwell, 2006) and, in extreme cases, cell death (Van Breusegem and Dat, 2006). To reduce oxidative damage, plants employ a complex enzymatic and non-enzymatic antioxidative system, which is triggered by ROS (Mittler *et al.*, 2011; Baxter *et al.*, 2014; You and Chan, 2015).

The decreased photosynthetic carbon assimilation, caused by stomatal closure, results in a reprogramming of plant central metabolism and growth to try to maintain the activity of essential metabolic pathways while simultaneously adapting to the stressful conditions (Obata and Fernie, 2012; Claeys and Inze, 2013). Under drought stress, accumulation of leaf metabolites thought to have protective functions (e.g. fructose, glucose, raffinose, proline, and glycine betaine) or to accumulate as a consequence of the stress (e.g. protein breakdown causing the overall increase in free amino acids) have been reported for a number of plant species (Verslues and Juenger, 2011; Obata and Fernie, 2012; Krasensky and Jonak, 2012; Fàbregas and Fernie, 2019). This metabolic response to drought varies during different developmental stages (Hummel *et al.*, 2010; Skirycz *et al.*, 2010) . Among them, the reproductive stage is considered the most drought sensitive in plants (Passioura, 2012) and particularly in rice, which, among the cereal crops, shows greatest sensitivity to water limitation (Venuprasad *et al.*, 2007). The top leaves of rice, and primarily the flag leaf, are the most important source of assimilates for the developing panicles (Yoshida, 1972). Drought-induced alteration of metabolism and redox state in these leaves during the reproductive

stage was shown to be linked with a reduction in grain yield (Biswal and Kohli, 2013; Sandhu *et al.*, 2014).

In the last decade, an increasing number of studies used metabolomics as a large-scale screening tool to identify markers for plant trait improvement (Fernandez *et al.*, 2016; Kumar *et al.*, 2017). This interest mainly relies on the fact that metabolic markers showed equal or even higher predictive power for plant traits than traditional genetic markers (Fernandez *et al.*, 2016). Metabolomics-based prediction of biomass was used in *Arabidopsis* recombinant inbred lines and accessions grown under optimal (Meyer *et al.*, 2007; Steinfath *et al.*, 2010; Sulpice *et al.*, 2009) and sub-optimal (Sulpice *et al.*, 2013) conditions. These studies were performed under controlled conditions, thus reducing environmental effects and increasing the likelihood of finding strong relationships between metabolite levels and the trait of interest (Fernandez *et al.*, 2016). In crop species, trait prediction based on metabolic profiling of large field grown populations of genetically diverse accessions remains rare and even more rare in experiments simultaneously conducted under optimal and non-optimal conditions (Riedelsheimer *et al.* 2012, 2013; Xu *et al*. 2016).

Here we present a large field study aimed at improving our understanding of the drought-induced reprogramming of metabolism and oxidative stress status in rice and its effect on grain yield. The flag leaf central metabolome together with a range of oxidative stress markers and redox state-related enzymes were analysed in a collection of 292 phenotypically and genetically diverse rice lines that were exposed to well-watered and drought conditions during the reproductive stage. The dataset was used to analyse the relationship between oxidative stress and central metabolism under the two different treatments, to generate multivariate models that predict grain yield loss under drought stress and to identify metabolic and oxidative stress markers of tolerance and sensitivity to drought. The combination of oxidative stress markers and enzyme activities with the metabolomics dataset for the prediction, and the large number of field-grown accessions, make this study an extensive and valuable source of information to identify biomarkers for improved grain yield stability of rice under drought.

MATERIALS AND METHODS

Genetic resources and plant growth

Two-hundred ninety-two accessions of *Oryza sativa* subsp. *indica* were used in a field experiment at the International Rice Research Institute (IRRI), Los Baños, Philippines during the 2013 dry season. The accessions are largely as those in the PRAY-indica panel (http://ricephenonetwork.irri.org) including traditional and improved *indica* lines, from tropical and sub-tropical regions. The experiment comprised a control field and drought stress field, each with three replicates of the population arranged in a serpentine design (Supplementary Fig. S1). To synchronize flowering, the accessions were divided into six groups according to days required to flower (previously collected data), and progressively sown and transplanted with intervals of 10 days between each group. Drought stress consisted of 14 consecutive days of water withholding applied only to the stress field at the reproductive stage (targeting 50% flowering). At the end of stress, the field was re-watered until all the accessions reached maturity for harvest (further details in Kadam *et al.*, 2018).

Phenotyping

Percentage of grain yield loss (GY loss) of each accession was calculated ($GY_{control}$-$GY_{drought}$)/($GY_{control}$*100), as the mean values of the GY loss of all replicates (3 for drought and 2 for control). A variable, Sam-Flow, was calculated as the date of leaf sampling (Sam), minus the date of 50% flowering (Flow) for every genotype, under control and drought treatment, separately. The genotypes together with their GY loss, Flow and Sam-Flow values are shown in Supplementary Table S1.

Leaf sampling

Eight flag/top leaves from the main tiller of 8 plants per plot (that were not used for yield determination) were sampled and immediately frozen in liquid nitrogen. Three drought field replicates of all accessions were collected (9:30- 11:00 am), on day 14 of the stress treatment. Two control field replicates of the entire population were collected, two days later, during the same time window. Samples were ground in liquid nitrogen, shipped to the Netherlands on dry ice and stored at -80°C until further analysis in Germany and Belgium, where samples were also shipped to on dry ice.

Metabolite profiling

Metabolite profiling was performed as described by Riewe *et al.* (2012) and Riewe *et al.* (2016). For each accession and treatment, equal amounts of replicates (3 for drought, 2 for control) were pooled, resulting in 584 samples (292 each for drought and control). Samples were extracted in MeOH/H_2O (15.0±1.0 mg fresh weight), dried and in-line derivatized (MPS2 autosampler, GERSTEL) prior to GC-MS analysis (AGILENT/LECO). Metabolites were identified using ChromaTOF software (LECO) and the library provided by the Golm Metabolome Database (GMD, http://gmd.mpimp-golm.mpg.de/download/). Peak intensities were determined using the R package *TargetSearch* (Cuadros-Inostroza *et al.*, 2009) and normalised against an internal standard (D8-Valine), fresh weight and detector response variation.

Glucose, fructose and sucrose detection

Glucose, fructose and sucrose were quantified spectrophotometrically (Riewe *et al.*, 2008). In brief, NADPH production at 340 nm was converted to saccharide content after the sequential addition of hexokinase (glucose), phosphoglucoisomerase (fructose) and invertase (sucrose) into a reaction mix, containing extract, glucose-6-phosphate dehydrogenase, ATP and $NADP^+$.

Oxidative stress markers

Malondialdehyde (MDA) is a by-product of lipid peroxidation. Its content was assayed according to Hodges *et al.* (1999). Leaf material (50 mg FW) was homogenized (80% v/v ethanol), using a MagNA Lyser (Roche, Vilvoorde, Belgium). After centrifugation, the supernatant was allowed to react with thiobarbituric acid to produce the chromogen, thiobarbituric acid-malondialdehyde (TBA-MDA). Absorbance of TBA-MDA was measured at 440, 532, and 600 nm, using a micro-plate reader (Synergy Mx, Biotek Instruments Inc., Vermont, VT, USA). Protein oxidation (ProtOx) was estimated through measuring the protein carbonyl content, utilizing dinitrophenylhydrazine (DNPH) derivatization (Levine *et al.*, 1994).

Molecular antioxidants

Total non-enzymatic antioxidant capacity (TAC) was assayed after homogenizing and extracting leaf tissue (50 mg FW), in 80% ethanol (v/v). The extract was centrifuged, and the tripyridyltriazine (TPTZ) assay-reagent was mixed with the extract (Benzie and Strain, 1999).

Absorbance change at 600 nm was measured using a microplate reader. Trolox (0 to 650 µM) was used as standard. Total polyphenol content (Poly) was assayed after the same homogenizing, extracting ad centrifugation steps, using a Folin–Ciocalteu reagent for phenol detection (Zhang *et al.*, 2006), with gallic acid as standard.

Antioxidant enzymes

Soluble proteins were extracted according to Murshed *et al.* (2008) and quantified by the Lowry method (Lowry *et al.*, 1951). Enzyme activities were determined in a semi-high-throughput set-up (Zinta *et al.*, 2014; AbdElgawad *et al.*, 2016). Ascorbate peroxidase (APX), dehydroascorbate reductase (DHAR), monodehydroascorbate reductase (MDHAR) and glutathione reductase (GR) activities were measured in extracts obtained from 100 mg of frozen tissue, in 1 mL of extraction buffer: 50 mM MES/KOH (pH 6.0) containing 0.04 M KCl, 2 mM $CaCl_2$, and 1 mM ASC, homogenized by MagNALyser (Roche, Vilvoorde, Belgium). APX, MDHAR, DHAR and GR activities were determined in microplates according to the method of Murshed *et al.* (2008). Their activities were assayed in 50 mM K-phosphate, 50 mM HEPES pH 7.6, 50 mM HEPES pH 7 and 50 mM HEPES pH 8, respectively. APX, MDHAR, and GR activities were measured by monitoring the decrease in ASC (ε290 = 2.8 $mM^{-1}cm^{-1}$), NADH (ε340 = 6.22 $mM^{-1}cm^{-1}$) and NADPH (ε340 = 6.22 $mM^{-1}cm^{-1}$). Peroxidase (POX) activity was determined by the oxidation of pyrogallol in 100 mM phosphate buffer (ε430 = 2.46 $mM^{-1}cm^{-1}$) (Kumar and Khan, 1983). Superoxide dismutase (SOD) activity was analysed by measuring the inhibition of nitro-blue tetrazolium (NBT) reduction (ε550= 12.8 $mM^{-1}cm^{-1}$) (Dhindsa *et al.*, 1982). Glutathione peroxidase (GPX) activity was measured as described by Drotar *et al.* (1985), in a coupled enzyme assay with glutathione reductase, measuring the decrease in NADPH absorption. Catalase (CAT) activity was assayed by monitoring the H_2O_2 decomposition at 240 nm (ε240 = 39.4 $M^{-1}cm^{-1}$) (Aebi, 1984). Ascorbate oxidase (AO) activity was measured as the rate of decrease of ascorbate (absorbance at 290 nm, Yoshimura *et al.*, 1998). Glutathione S-transferase (GST) activity was determined by measuring conjugation of GSH to 1-chloro-2,4-dinitrobenzene (CDNB) at 340 nm (Habig *et al.*, 1974). Peroxiredoxin (PRX) activity was determined according to Horling *et al.* (2003), by measuring the decrease in H_2O_2 concentration in the reaction mixture. Glutaredoxin (Grx) activity was determined by measuring the reduction of 2-hydroxy-ethyl-disulfide by GSH in the presence of NADPH and yeast glutathione reductase (Lundberg *et al.*, 2001). Thioredoxin (Trx) activity

was determined by measuring NADPH oxidation (Wolosiuk *et al.*, 1979) at 340 nm. Ferredoxin-NADP(H) reductases (Frxs) activity was determined as the reduction of potassium ferricyanide at 420 nm (Rodriguez *et al.*, 2006).

Photorespiration enzymes

Glycolate oxidase (GOX) activity was determined by the formation of a glyoxylate complex with phenylhydrazine (ε324 = 17 mM-1 cm-1, Feierabend and Beevers, 1972). Hydroxypyruvate reductase (HPR) activity was determined according to Schwitzguebel and Siegenthaler (1984), as the oxidation of NADH that was followed at 340 nm upon hydroxypyruvate addition.

Data analysis

Results of all metabolites and oxidative stress markers/enzymes were $\log_{10}$ transformed to improve normality. Metabolites showing more than 5% missing values among the samples were excluded from the analysis. A list of the results of metabolites and oxidative stress markers/enzymes used for statistical analyses is included in Supplementary Table S2.

Statistical analyses and graphical representations were performed using R (version 3.4.3; The R Foundation for Statistical Computing). Fold-change was calculated (on non $\log_{10}$-transformed data) dividing drought values by control values. Hierarchical clustering analysis was conducted using the *hclust* function (*stats* package) and based on complete linkage analysis of pairwise dissimilarity, calculated as $1\text{-}r_s$ (r_s, Spearman rank correlation coefficient). Dendrograms were created by using the *dendextend* package, and heat maps of correlations were created using the *heatmap.2* function (*gplot* package). Imputation of missing values, prior to PCA and PLSR analyses, was performed by the *knnImputation* function in the *DMwR* package. PCA was performed using the *prcomp* function (*stats* package) and the value of each variable was centred (mean subtraction) and scaled (standard deviation division) before analysis.

To identify the variables predictive for grain yield loss, a cross-validated partial least squares regression (PLSR, *pls* package) was used (Mevik and Wehrens, 2007; Mumm *et al.*, 2016). Observations were auto-scaled in the PLSR procedure. The number of latent variables to include in the model was selected by testing the predictability value (Q^2) using an increasing

number of latent variables from 1 to 10. The relative importance of the metabolites in the models was summarized using rank-products.

RESULTS

Relationship between differences in flowering time and grain yield loss

The rice population was grown in a field trial as part of a study aimed to collect information on phenotypic trait performance under well-watered and drought stress conditions (Kadam *et al.*, 2018). To evaluate the impact of drought stress on grain yield stability, the percentage of grain yield loss (GY loss), under drought versus control conditions, was calculated across the 292 genotypes as indicator of stress tolerance/susceptibility (Supplementary Table S1).

Even though the accessions were sown and transplanted on different dates to minimise flowering time differences, flowering synchronisation was not perfect, and could represent a confounding effect on yield results under drought (Kadam *et al.*, 2018). Correlation analysis was performed to evaluate the influence of flowering time differences (Supplementary Table S1) on the drought-induced GY loss performance of the 292 accessions (Supplementary Fig. S2). Flowering (Flow) under drought significantly and negatively correlated (p-value < 0.001) with GY loss (r_s = -0.35) but only 12% (R^2 = 0.12) of the variation was explained by the corresponding linear model. In general, the correlation trend shows that accessions that already flowered before stress imposition (less than the 10% of the total) displayed a relatively higher severity of GY loss than the ones that nearly or already flowered (booting stage for 60% and heading stage for 30% of the total) during stress imposition. Interestingly, a significant (p-value < 0.001) and almost identical negative correlation was observed between Flow under control (r_s = -0.37; R^2 = 0.13) and GY loss. This similarity is determined by the almost perfect correlation (r_s = 0.96; R^2 = 0.94; p-value < 0.001) observed between Flow under control and drought (Supplementary Fig. S2). Nevertheless, drought stress significantly affected (Paired t-test's p-value < 0.001) the date of 50% flowering with a delay of around three days (mean±sd: 83.9±10.6) compared to control (mean±sd: 81.0±10.3). The almost perfect correlation between Flow under the two treatments indicates that the flowering delay under drought is virtually identical in the 292 accessions.

Drought induces accumulation of amino acids and affects the level of antioxidant enzymes and organic acids

Leaf samples of the 292 accessions were analysed by untargeted GC-MS-based metabolite profiling to assess the variation in polar metabolites under well-watered and drought conditions. A total of 88 metabolites were identified, predominantly primary metabolites (amino acids, sugars and organic acids). The amount of the three most abundant sugars (sucrose, fructose and glucose) was determined spectrophotometrically (glucose also by mass spectrometry). The same leaf materials were analysed for the oxidative stress status of the different accessions under the two treatments. For this, the level of molecular antioxidants (2), oxidative stress markers (2) and the activity of enzymes (16) involved in ROS scavenging mechanisms and photorespiration were quantified. The 111 variables, metabolites and oxidative stress markers/enzymes considered in this study (Supplementary Table S3) are hereafter referred to as MetabOxi.

Principal component analysis (PCA) was used to gain insight in the overall effect of drought on the MetabOxi profiles of the accessions. Among the first three principal components (PCs), PC1 explained 29.5% of the total variation, and almost completely separated the control and drought stressed samples (Supplementary Fig. S3), suggesting a strong influence of drought on the MetabOxi profiles of the accessions. To evaluate the treatment effect on the level of the individual MetabOxi variables, one-way ANOVA was conducted. This showed that drought significantly influenced (Bonferroni-corrected p-value < 0.05) most (91 out of 111) of the MetabOxi variables across the 292 genotypes (Supplementary Table S4). To quantify the magnitude of these alterations, we conducted a fold-change (f.c.) analysis (stress over control values) for the 91 MetabOxi variables that changed significantly between treatments (Table 1). The majority (75 out of 91) displayed a fold-change increase (f.c. > 1) with allantoin and 2-aminoadipic acid showing the largest increase (f.c. > 10). Interestingly, the level of all amino acids significantly increased under drought with the highest f.c. values for Pro, Asn-H_2O and Orn. Organic acids, sugars, oxidative stress markers and enzymes displayed a more diverse response. Particularly, most of the organic acids showed a decrease (f.c. < 1) with two tricarboxylic acid cycle (TCA) intermediates, isocitric and citric acid, and two glycolysis intermediates, glyceric acid-3-phosphate and phosphoenolpyruvic acid, displaying the largest decrease (f.c. < 0.6). Of the sugars, galactinol and raffinose displayed a strong increase (f.c. >

4), just as the ROS scavenging enzymes superoxide dismutase (SOD), ascorbate peroxidase (APX) and catalase (CAT) (3 > f.c. > 4). Collectively, these results highlight the strong influence of drought stress on the metabolome and oxidative stress status of the population, mainly characterized by an increase in the level of amino acids and activity of specific antioxidant enzymes and by a decrease in the level of organic acids.

Table 1. The effect of drought on the MetabOxi variables showing a significant response to stress. Fold change mean values (drought/control) of the 91 variables showing a significant effect of treatment (Bonferroni-corrected *p*-value < 0.05) by one-way ANOVA. The fold change values are colored according to their fold change increase or decrease class (≤ 0.5 dark blue, 0.5~1.0 light blue, 1.0~2.0 light red, 2.0~4.0 red, ≥ 4.0 purple). st.dev.: standard deviation.

Variable class	Name	fold change	st. dev	Variable class	Name	fold change	st. dev
Amino acid	Alanine	1.78	0.87	Organic acid	2-Piperidinecarboxylic acid	1.99	1.35
	Alanine, beta-	3.10	2.44		Adipic acid, 2-amino-	10.46	13.79
	Arginine	2.64	3.35		Butanoic acid, 4-amino-	2.62	2.22
	Asparagine -H2O	6.16	12.32		Butanoic acid, 4-hydroxy-	1.98	0.88
	Aspartic acid	1.78	0.92		Citric acid	0.58	0.34
	Cysteine	1.67	0.87		Erythronic acid	1.38	0.38
	Glutamic acid	1.34	0.33		Fumaric acid	1.16	0.39
	Glutamine	3.57	3.82		Galactonic acid-1,4-lactone	1.52	0.75
	Glycine	2.59	2.25		Gluconic acid-1,5-lactone	1.18	0.36
	Isoleucine	3.03	3.43		Glutaric acid, 2-hydroxy-	1.48	0.61
	Leucine	3.09	2.98		Glyceric acid	0.77	0.21
	Lysine	2.56	3.46		Glyceric acid-3-phosphate	0.51	0.32
	Methionine	2.42	3.11		Isocitric acid	0.39	0.22
	Ornithine	5.81	9.51		Maleic acid	1.30	0.53
	Phenylalanine	3.44	3.86		Malic acid, 2-methyl-	0.67	0.27
	Proline	7.29	7.94		Malonic acid	3.33	2.34
	Pyroglutamic acid	1.65	0.79		Phosphoenolpyruvic acid	0.56	0.33
	Serine	1.83	0.90		Phosphoric acid	0.77	0.21
	Serine, O-acetyl-	1.00	0.59		Prephenic acid	0.93	0.35
	Threonine	2.54	2.19		Quinic acid	1.75	0.94
	Tryptophan	2.86	2.22		Quinoline-2-carboxylic acid,…	1.25	0.83
	Tyrosine	1.77	1.40		Salicylic acid	1.55	1.21
	Valine	2.03	1.98		Shikimic acid	1.52	0.73
Sugars	Erythritol	2.00	1.14		Shikimic acid, 3-dehydro-	1.72	0.97
	Fructose (abs)	1.48	0.50		Succinic acid	1.23	0.44
	Fructose-6-phosphate	0.95	0.37		Threonic acid	1.46	0.55
	Fucose	1.55	0.62	Oxidative stress enzyme	GOX	1.94	0.95
	Galactinol	5.16	2.89		HPR	2.73	2.24
	Galactosamine, N-acetyl-	1.61	0.84		POX	1.35	0.78
	Glucosamine, N-acetyl-	1.48	0.58		APX	3.10	7.78
	Glucose (abs)	1.69	0.73		CAT	3.08	6.37
	Glucose	2.34	1.72		GST	0.94	0.31
	Glucose-6-phosphate	0.94	0.41		SOD	4.08	2.64
	Glycerol-3-phosphate	1.69	1.30		DHAR	1.05	0.88
	Inositol, myo-	1.52	0.34		MDHAR	0.76	0.59
	Isomaltose	3.90	1.88		AO	1.82	4.26
	Maltose	1.60	1.05		GR	0.91	0.78
	Mannosamine, N-acetyl-	1.59	0.82		Trxs	1.45	0.75
	Raffinose	4.40	2.00		Prxs	1.42	0.74
	Ribitol	1.46	0.63		Frxs	1.43	0.56
	Ribose-5-phosphate…	0.97	0.41	Oxidative stress marker	MDA	2.18	0.96
	Trehalose	1.36	1.28		TAC	0.78	0.47
Other known	Allantoin	11.18	15.19		Poly	0.91	0.76
	Guanosine	2.70	3.42		ProtOx	1.46	0.39
	Secologanin	2.00	0.91				
	Urea	2.04	1.11				
	Uridine	0.94	0.26				

Drought stress increases the correlations within and between metabolite classes

Correlation analysis was performed to assess the nature and strength of the associations between flag leaf metabolites under control and drought conditions (Supplementary Table S5 and S6). Fig. 1 presents the degree of correlation among metabolites, together with their hierarchical clustering under the two treatments (see Supplementary Fig. S4 for a detailed representation of the dendrograms). By comparing the two heat maps, similarities as well as stress-induced population-wide differences in the correlations between leaf metabolites can be identified. Metabolites of the same class (mainly amino acids, sugars and organic acids) clustered together in both treatments (clusters of the same colour in Fig. 1). Only a single drought-specific cluster (yellow), including five amino acids (Asp, Glu, Gly, Ser and pyroglutamic acid) and allantoin, was identified. Drought stress resulted in stronger correlations between metabolites within each cluster than for control conditions as evidenced, particularly, by the amino acid cluster (red). Drought stress also increased the correlations (positive or negative) between clusters representing different metabolite classes. For example, under drought the amino acid (red) and sugar cluster (green) displayed an increased correlation with the latter also showing an increased negative correlation with the organic acid clusters (light blue and dark blue). Particularly, between the two organic acid clusters, the one containing three TCA intermediates, citric, isocitric and malic acid (light blue), showed the strongest negative correlation with the amino acid (red) and sugar (green) clusters. Weaker correlations with the same clusters were displayed by the other organic acid cluster (dark blue), enriched in metabolites of the shikimate pathway (shikimic acid, quinic acid and 3-dehydroshikimic acid). Interestingly, the drought-specific amino acid-enriched cluster (yellow) displayed a similar correlation pattern with the other metabolite clusters as the major amino acid cluster (red) to which it was strongly positively correlated. Under control conditions, the abovementioned clusters showed very low correlations with each other. In summary, drought stress resulted in a generally stronger correlation within and between all the leaf metabolite classes than under control conditions and induced one new (stress-specific) cluster, containing Asp, Glu, Gly, Ser, pyroglutamic acid and allantoin.

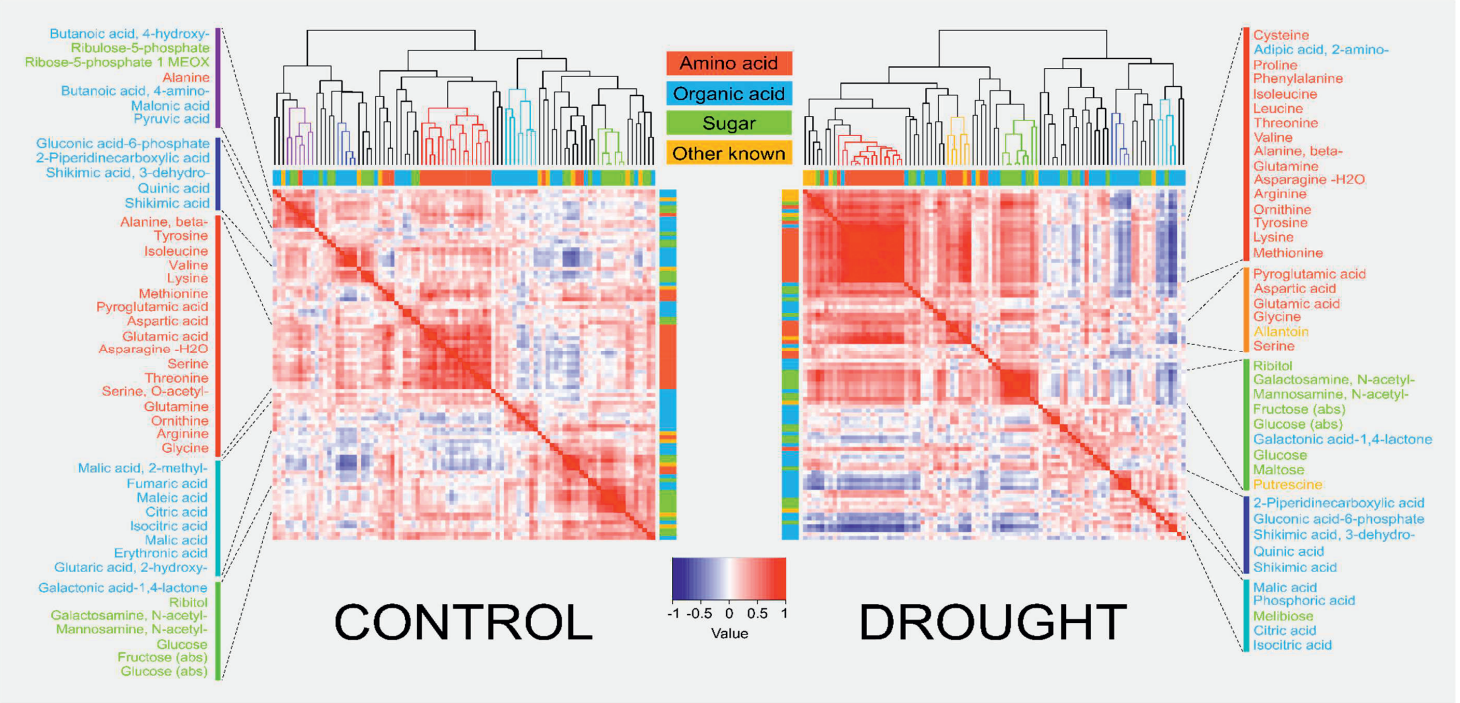


Fig. 1. Hierarchical clustering and heatmap of pairwise correlations between metabolites under the two treatments for the 292 rice accessions. Spearman correlations under control (left) and drought stress (right) conditions. Color bars on the top and side of the heatmaps represent the four main classes of metabolites (amino acid, organic acid, sugar, other known). In the dendrograms, main clusters are colored according to the class of the majority of the metabolites they included: red (amino acids), light blue (TCA cycle organic acids), blue (shikimic acid pathway), green (sugars), purple (mixed cluster), yellow (mainly amino acid). All the other minor clusters are colored in black. The metabolites included in each cluster are displayed at the side of the two heatmaps. Metabolite names are colored according to the four main classes of known metabolites mentioned above.

Leaf amino acid metabolism is strongly linked with stress-induced photorespiratory and antioxidative enzymes

To investigate the relationships between leaf oxidative stress status and metabolism, correlation analysis and hierarchical clustering were conducted on the combined datasets (Fig. 2, Supplementary Table S7 and S8). Strikingly, a cluster formed by two photorespiratory enzymes, hydroxypyruvate reductase (HPR) and glycolate oxidase (GOX), and the lipid peroxidation product malondialdehyde (MDA) was present under both control and drought conditions (brown). HPR and GOX showed the strongest positive correlation with a number of metabolites under both treatments, and, especially, with the main cluster of amino acids (red). The two enzymes also displayed a strong positive correlation with the drought-specific cluster of amino acids (yellow), which contains Gly and Ser, known to be produced by the photorespiratory pathway. MDA showed low correlations with all other metabolites under control, whereas, under stress, these correlations became stronger with an overall pattern very similar to that of HPR and GOX (Fig. 2). Under drought stress, the HPR-GOX-MDA cluster also showed a stronger negative correlation with the TCA cycle cluster (light blue) than under control conditions. Considering the other oxidative stress markers/enzymes, almost all (15 out of 17) displayed negative correlations with the single amino acid cluster (red) under control conditions. Under drought, a more limited number (7 out of 17) showed a negative correlation with the two amino acid clusters (red and yellow). However, among them, specific ROS scavenger enzymes such as dehydroascorbate reductase (DHAR), CAT and APX showed very strong negative correlations with the amino acid clusters (red and yellow), stronger than in control conditions. In summary, these results show the presence of a strong correlation between the flag leaf oxidative stress status and metabolome in both treatments but with strongest and more specific associations under drought, particularly between oxidative stress enzyme activities/markers and amino acids. The best correlations between the variables of the two datasets were the same under drought and control conditions: HPR with Gly and GOX with Ser (Fig. 3). For GOX and Ser, the correlation value (r_s) was 0.76 under control and 0.91 under drought stress, with the two linear models able to explain 58% and 83% of the variation, respectively. For HPR and Gly the correlation was even stronger, with r_s = 0.84 under control and r_s = 0.96 under drought, and with 75% and 93% of the variation, respectively, explained by the linear models.

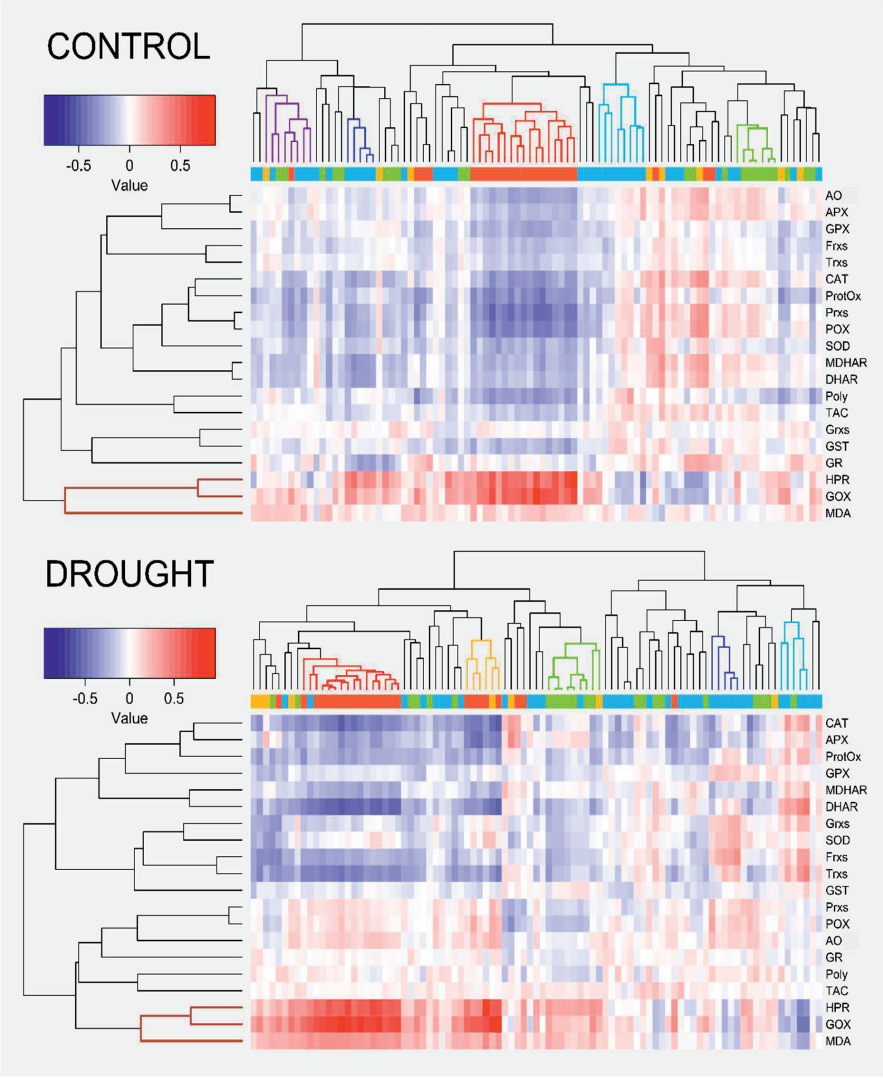


Fig. 2. Hierarchical clustering and heatmap of the oxidative stress markers/enzymes versus metabolites pairwise correlations under the two treatments for the 292 rice accessions.
Spearman correlations under control (top) and drought stress (bottom) conditions. The metabolite color code of the bar below the dendrograms as well as the color of the main clusters are the same as in Fig. 1. Malondialdehyde (MDA), polyphenols (Poly), protein oxidation (ProtOx), total antioxidant capacity (TAC), ascorbate oxidase (AO), ascorbate peroxidase (APX), catalase (CAT), dehydroascorbate reductase (DHAR), ferredoxins (Frxs), glycolate oxidase (GOX), glutathione peroxidase (GPX), glutathione reductase (GR), glutaredoxins (Grxs), glutathione S-transferase (GST), hydroxypyruvate reductase (HPR), monodehydroascorbate reductase (MDHAR), peroxidase (POX), peroxiredoxins (Prxs), superoxide dismutase (SOD), thioredoxins (Trxs).

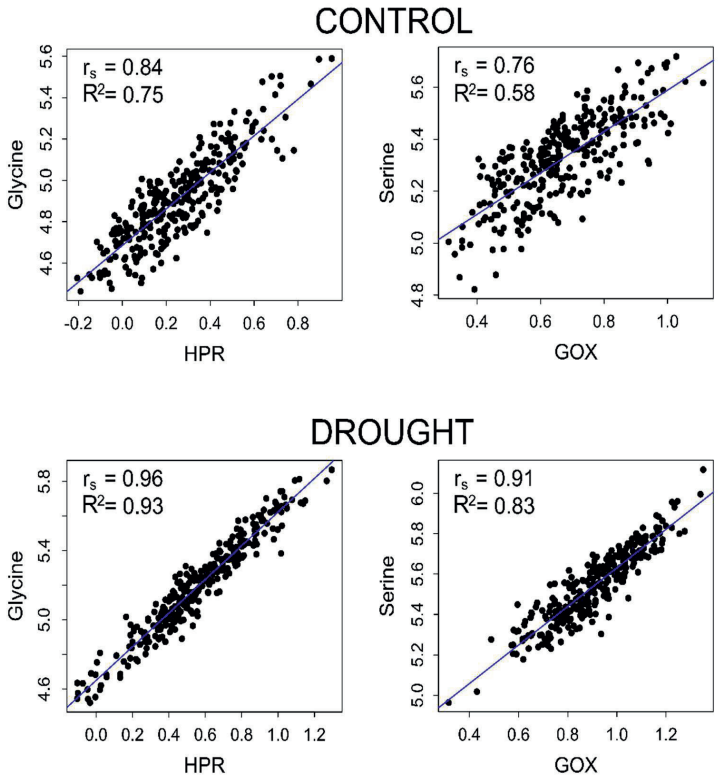


Fig. 3. Linear models between the photorespiratory enzymes and hydroxypiruvate reductase and glycolate oxidase and the amino acids glycine and serine under the two treatments.
Linear models of the two best correlations between MetabOxi variables under the two treatments. Axes are expressed as $\log_{10}$ transformed detector response values for the metabolites glycine and serine and $\log_{10}$ enzyme activity for hydroxypyruvate reductase (HPR) and glycolate oxidase (GOX). In each plot are reported the Spearman correlation value (r_s) between the two variables and the variation explained by the specific linear model (R^2).

Single MetabOxi variables are highly correlated with the genotypic variation in grain yield stability

Next, a correlation analysis was carried out on the control and drought values of the 111 MetabOxi variables and GY loss (Supplementary Table S9) to assess if single variables could be associated with yield stability of the 292 rice accessions.

A higher number of significant correlations (Bonferroni corrected *P*-value < 0.05) with GY loss were found using drought values of the MetabOxi variables (53) than the control (25) values (Supplementary Table S9). Under both treatments, positive correlations outnumbered the negative ones (23 out of 25 positive correlations under control and 48 out of 53 under

drought). The variables that displayed the strongest correlations differed between the two treatments. Under control conditions (Table 2), erythritol showed the best correlation (r_s = 0.41) with GY loss, followed by a number of amino acids (Phe, Leu, beta-Ala as top ones) and 2-amino-adipic acid, all displaying similar and positive values (r_s ~ 0.30). The percentage of GY loss variance explained by the linear models (R^2) under control was low (below 20%), with erythritol showing the highest contribution (17%).

Table 2. MetabOxi variables with the best correlations between control values and grain yield loss. List of the 15 best (highest significance) Spearman correlations between control values of the MetabOxi variables and grain yield loss (GY loss). Spearman rank correlation values and significance between the same variables and flowering at sampling (Sam-Flow) are displayed. R^2: variance explained by the linear model created between the trait and each variable. Star (*) indicates a MetabOxi variable significantly correlated with both grain yield loss and flowering at sampling.

	CONTROL						
	GY loss			Sam-Flow			
Variable	Spearman's rho	Bonf.corr. *p*-value	R^2	Spearman's rho	Bonf.corr. *p*-value	R^2	
Erythritol	0.41	4.28E-11	0.17	0.75	5.17E-34	0.42	*
Phenylalanine	0.36	4.69E-08	0.13	0.57	7.20E-16	0.23	*
Alanine, beta-	0.33	1.75E-07	0.12	0.04	1	0.00	
Adipic acid, 2-amino-	0.31	5.48E-07	0.11	0.11	1	0.01	
Leucine	0.33	1.62E-06	0.11	0.27	0.0313	0.05	*
Maltose	0.31	1.65E-05	0.09	0.17	1	0.01	
Proline	0.31	2.83E-05	0.09	0.21	1	0.00	
Lysine	0.26	0.0005	0.07	-0.06	1	0.02	
Glucosamine, N-acetyl-	0.30	0.0010	0.07	0.76	7.95E-43	0.50	*
Isoleucine	0.25	0.0019	0.06	0.00	1	0.00	
Ribitol	0.25	0.0021	0.06	0.43	5.03E-10	0.15	*
Mannosamine, N-acetyl-	0.26	0.0029	0.06	0.41	1.55E-08	0.13	*
Ornithine	0.23	0.0061	0.06	-0.10	1	0.00	
Tyramine	0.26	0.0063	0.05	0.44	2.96E-09	0.14	*
Poly	-0.24	0.0092	0.05	-0.01	1	0.00	

Under drought, correlation values strongly increased (Table 3 and Supplementary Table S10) with seventeen MetabOxi variables displaying higher values than the top one under control conditions. The two best correlations with GY loss under drought were displayed by the lipid peroxidation product MDA (r_s = 0.63) and the antioxidant enzyme DHAR (r_s = -0.56), both showing a similar percentage of GY loss variance explained by their respective linear models (R^2 = 0.38 and 0.37). Interestingly, DHAR was the only top ranked variable to show a negative correlation with GY loss. Under drought, many more amino acids, than under control conditions, ranked among the top correlated variables with GY loss (Thr, Arg, Val as top ones).

Table 3. MetabOxi variables with the best correlations between drought values and grain yield loss. List of the 15 best (highest significance) Spearman correlations between drought values of the MetabOxi variables and grain yield loss (GY loss). Spearman rank correlation values and significance between the same variables and flowering at sampling (Sam-Flow) are displayed. R^2: variance explained by the linear model created between the trait and each variable. Star (*) indicates a MetabOxi variable significantly correlated with both grain yield loss and flowering at sampling.

	DROUGHT						
	GY loss			Sam-Flow			
Variable	Spearman's rho	Bonf.corr. *p*-value	R^2	Spearman's rho	Bonf.corr. *p*-value	R^2	
MDA	0.63	3.65E-30	0.38	0.23	0.1106	0.04	
DHAR	-0.56	5.81E-29	0.37	-0.06	1	0.00	
Threonine	0.48	6.17E-18	0.25	0.33	1.21E-06	0.11	*
Arginine	0.48	6.75E-18	0.25	0.36	1.73E-06	0.11	*
Valine	0.47	9.99E-17	0.24	0.39	2.40E-09	0.14	*
Ornithine	0.46	2.44E-16	0.23	0.35	1.79E-07	0.12	*
Isoleucine	0.46	1.13E-15	0.22	0.41	3.79E-12	0.18	*
Serine	0.46	2.03E-15	0.22	0.25	0.0248	0.05	*
Methionine	0.45	2.54E-15	0.22	0.41	1.34E-11	0.17	*
Leucine	0.45	7.23E-15	0.21	0.40	1.20E-12	0.19	*
Glutamine	0.44	8.17E-15	0.21	0.38	6.60E-11	0.16	*
Phenylalanine	0.44	1.96E-14	0.21	0.45	1.38E-16	0.24	*
Tyrosine	0.44	5.84E-14	0.20	0.39	1.65E-12	0.19	*
Alanine, beta-	0.43	3.82E-13	0.19	0.32	3.48E-08	0.13	*
Lysine	0.41	9.76E-13	0.19	0.35	1.39E-08	0.13	*

Again, all these amino acids displayed similar correlation values with GY loss (r_s ~ 0.45) but the correlations under drought, as well as the R^2 values of their linear models (~ 20%), were substantially higher than under control conditions.

Considering the previously described significant and similar correlation values between GY loss and flowering under both treatments (Supplementary Fig S2), we also decided to investigate if values of the MetabOxi variables were associated with differences in flowering among the accessions at sampling time. For this purpose, a new flowering variable (Sam-Flow) was created by subtracting the date of 50% flowering from the date of leaf sampling for every genotype (Supplementary Table S1) and correlation analysis was then performed between Sam-Flow and the control and drought values of the 111 MetabOxi variables (Supplementary Table S10). Interestingly, the top two correlated metabolites with GY loss under control conditions, erythritol and Phe, displayed significant and particularly high correlations (r_s = 0.75 and 0.57, respectively) with Sam-Flow. Different from the control conditions, the two top correlated variables with GY loss, MDA and DHAR, showed no significant correlation with Sam-Flow (Table 3).

Collectively, these results suggest that only drought values of the MetabOxi variables, and primarily of MDA and DHAR, are highly associated with GY loss performance across the 292 genotypes. In addition, the non-significant correlation of MDA and DHAR with Sam-Flow further validates their true association with GY loss only. Nevertheless, the trait prediction accuracy of these two variables remains below 40% of the GY loss phenotypic variance.

Prediction of grain yield loss by multivariate partial least squares regression models reveals the importance of leaf antioxidant system for yield stability

Two models based on multivariate partial least squares regression (PLSR) were generated to predict GY loss performance of the population considering the 111 MetabOxi variables. In the first model, control values of variables were used for prediction of GY loss under drought. The best model (1LV, see Supplementary Fig. S5) showed a low predictability (Q^2 = 0.15) (Fig. 4). This suggests that non-stressed values of flag leaf MetabOxi variables are not very accurate markers of tolerance/sensitivity for drought-induced GY loss across the population. In contrast, the best PLSR model based on values of the MetabOxi variables under drought conditions (6LVs, see Supplementary Fig. S5) displayed a high predictability (Q^2 = 0.64) (Fig. 4), indicating a strong association between stress-induced changes in leaf metabolism/oxidative stress and GY loss.

Both the PLSR models were based on ten different single sub-models generated by the cross-validating procedure. By multiplying the ten ranks of each MetabOxi variable in the single sub-models, their overall ranking was calculated (Supplementary Table S11 and S12). A low rank-product implies that the variable has a high importance for the model. Among the top five predicting variables of the control model (Table 4), galactaric acid and erythritol ranked 1st and 2nd, both showing a similarly low rank-product value (32 and 64, respectively). Interestingly, galactaric acid poorly correlates with GY loss whereas erythritol, as mentioned before, showed the highest correlation with the trait under control conditions. Additionally, erythritol showed a strong correlation with Sam-Flow whereas galactaric acid did not (Table 4). The next highest ranking predictors of the model, 2-aminoadipic acid, Trp and allantoin, showed a much lower importance than the first two, as represented by their high rank-product values. Interestingly, Trp, like erythritol, showed a strong positive correlation with Sam-Flow.

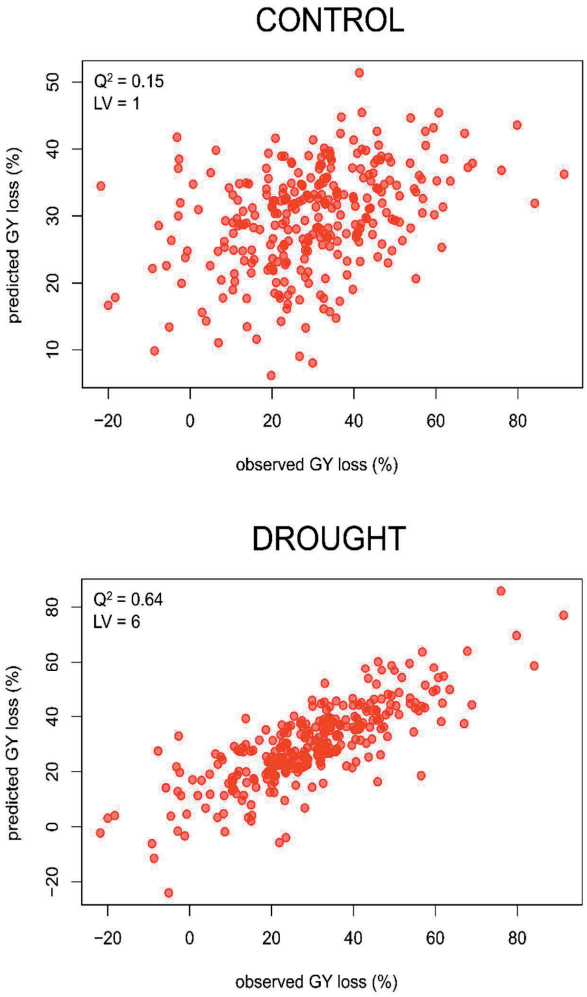


Fig. 4. Predicted versus observed values of the control and drought PLSR models for the prediction of grain yield loss in the 292 rice accessions.
PLSR plot of the cross-validated models for GY loss based on the control (top) and drought (bottom) values of the 111 MetabOxi variables. Predictability (Q^2) and linear latent variables (LVs) of the two models are reported.

In the drought-based PLSR model of GY loss, all the five top predicting variables are oxidative stress enzymes or markers (Table 4). The most important predicting variable of the model, by far, is the antioxidant enzyme DHAR that ranked first in all the ten single sub-models (rank-product = 1). The second highest ranking predicting variable of the model is the lipid peroxidation product MDA that displayed a still relatively low rank-product value (1024) albeit

already considerably higher than DHAR. DHAR and MDA are also the two variables that showed the highest correlation coefficients with GY loss, the first negative and the second positive. The model selected the antioxidant enzyme MDHAR and the level of total non-enzymatic antioxidant capacity (TAC) as third and fourth best predicting variables even if both poorly correlate with GY loss. The fifth most important predicting variable of the PLSR drought model, ascorbate oxidase (AO) is, like DHAR and MDHAR, an enzyme involved in the ascorbate-glutathione redox cycle. The high rank-product value displayed by MDHAR, TAC and AO indicates they have a substantially lower importance in the model than DHAR and MDA. Differently from the control model, the top predictors of the drought model did not show significant correlations with Sam-Flow (Table 4), except from AO that displayed a weak correlation.

PLSR prediction models of grain yield stability in the population highlighted that only the drought stressed values of the MetabOxi variables could provide an accurate prediction of the trait. In addition, only antioxidant enzymes and oxidative stress markers ranked among the top ranked predicting variables of the drought model and they showed low correlations with Sam-Flow. This indicates that stress-induced alteration in the leaf oxidative stress status - much more than the metabolome - is tightly linked to the prevention of GY loss under drought.

Table 4. **Best predicting variables of the control and drought PLSR models for the prediction of grain yield loss in the 292 rice accessions.**

Top five ranked predicting variables of the double cross-validated PLSR models for grain yield loss prediction based on control (left) and drought (right) values. Variables are ranked based on their rank-product value. Variables with the lower rank-product value are the ones with the larger discriminative power. r_s: Spearman's rho of correlation with grain yield loss (GY loss) and flowering at sampling (Sam-Flow). Star (*) indicates a significantly correlated variable.

CONTROL PLSR MODEL			
Variable	Rank-product	GY loss r_s	Sam-Flow r_s
Galactaric acid	32	0.09	0.00
Erythritol	64	0.41 *	0.75 *
Adipic acid, 2-amino-	9.841E+04	0.31 *	0.11
Tryptophan	7.680E+05	0.27 *	0.73 *
Allantoin	2.540E+07	0.18 *	0.18

DROUGHT PLSR MODEL			
Variable	Rank-product	GY loss r_s	Sam-Flow r_s
DHAR	1	-0.56 *	-0.06
MDA	1024	0.63 *	0.23
MDHAR	5.905E+04	-0.03	0.17
TAC	1.806E+07	0.08	0.07
AO	3.931E+07	0.22 *	-0.17 *

DISCUSSION

Drought affects flag leaf central metabolism and induces leaf senescence, protein degradation and nitrogen recycling

Drought-induced stomatal closure reduces leaf photosynthetic activity, which induces metabolic reprogramming aimed to simultaneously adapt to the stressful condition and maintain active essential metabolic pathways (Obata and Fernie, 2012; Claeys and Inze, 2013). In this study, the imposition of drought stress during flowering resulted in a population-wide alteration of flag leaf central metabolism and oxidative stress status. The stress induced an increase in the level of most metabolites with a marked increase in the level of all the free amino acids and a, slightly less marked, increase of almost all the sugars, and decrease in the organic acids (Table 1).

The increase in amino acid levels is a response to water limitation that has been observed before in leaves of many plant species (Obata and Fernie, 2012; Krasensky and Jonak, 2012; Obata *et al.*, 2015; Fàbregas and Fernie, 2019) when exposed to severe drought stress in the vegetative stage. This accumulation has been associated with protein catabolism induced by premature stress-induced leaf senescence (Araújo *et al.*, 2011; Watanabe *et al.*, 2013; Hildebrandt *et al.*, 2015). The presence of a strong drought-induced catabolic activity is further supported by the large increase under drought in the level of two metabolites, allantoin and 2-aminoadipic acid (Table 1). Allantoin is an intermediate in purine catabolism known to be important for nitrogen remobilization and more recently postulated to have a role in stress tolerance by activating the production of abscisic acid (Watanabe et al. 2014). 2-Aminoadipic acid is the central metabolite involved in the plant catabolic pathway of lysine (Zhu *et al.*, 2000).

Accumulation of specialized metabolites able to induce water retention and positive turgor pressure and to counteract the enhanced generation of ROS helps to protect cellular functions against the damage caused by drought-induced dehydration (Verslues and Juenger, 2011; Krasensky and Jonak, 2012; Nakabayashi and Saito, 2015). We found that the amino acid showing the highest population-wide increase was Pro, widely reported in the literature to accumulate under stressful conditions and considered as an osmolyte, ROS scavenger and stabilizer of protein structure (Hare and Cress, 1997; Verslues and Juenger, 2011). Similarly, the highest increase in the sugars was displayed by two members of the RFOs family, galactinol

and raffinose. These two sugar alcohols have been reported to accumulate, like Pro, in leaves exposed to environmental stress and to exert an osmoprotective action and scavenging activity against hydroxyl radicals (Nishizawa *et al.*, 2008; Ende, 2013; Fàbregas and Fernie, 2019). Interestingly and contrary to the literature, in our study the strong increase in Pro, galactinol and raffinose seemed not to be associated with a signature of stress tolerance. Indeed, drought values of these three metabolites neither showed negative correlations with GY loss (Supplementary Table S9) nor ranked as good model predictors for GY loss (Supplementary Table S12).

Only few metabolite levels decreased under drought and most of them belonged to the class of organic acids with TCA cycle (isocitric and citric acid), and glycolysis (glyceric acid-3-phosphate and phosphoenolpyruvic acid) intermediates displaying the highest decrease (Table 1). The TCA cycle and glycolysis are two interconnected pathways involved in the production of metabolic intermediates used in biosynthesis elsewhere in the cell (Araujo *et al.*, 2012). Under drought, the strong reduction in the levels of metabolites of these two pathways could be due to an overall reduced biosynthetic activity of stressed leaves. This indirectly supports the fact that the increased levels of all the amino acids must be due to stress-induced protein degradation.

Drought stress alters the relationships between central metabolic pathways

Consistent with previous studies, hierarchical clustering analysis (Fig. 1) displayed the tendency of metabolites of the same class (amino acids, organic acids and sugars) to cluster together as evidence of commonly shared biosynthetic or catabolic pathways (Carreno-Quintero *et al.*, 2012; Obata and Fernie, 2012; Riedelsheimer *et al.*, 2012; Obata *et al.*, 2015). Drought induced a unique stress-specific cluster including allantoin and four amino acids, Asp, Glu, Gly and Ser (yellow in Fig. 1) that separated from the main cluster of the amino acids (red). Above, we already discuss the role of allantoin in nitrogen remobilisation as intermediate product of purine catabolism. A similar role is exerted by two amino acids of the cluster, Asp and Glu, both involved in nitrogen remobilisation as major long-distance phloem transport nitrogen forms (after their conversion to Asn and Gln) in senescent leaves (Watanabe *et al.*, 2013; Avila-Ospina *et al.*, 2014). The other two amino acids of the stress-specific cluster, Gly and Ser, are well known markers of photorespiration as an increased activity of that process

results in higher production of these compounds (Maurino and Peterhansel, 2010; Bauwe *et al.*, 2010). The presence of this cluster, under drought conditions only, further supports the idea that physiological processes such as stress-induced senescence, nutrient recycling and photorespiration are co-ordinately enhanced by drought and have an impact on leaf central metabolism (Avila-Ospina *et al.*, 2014; Hildebrandt *et al.*, 2015; Hodges *et al.*, 2016). Confirming this hypothesis are the increased negative correlations that we found under drought between the amino acid clusters (red and yellow in Fig. 1) and the TCA cycle cluster (light blue). These negative correlations link high levels of photorespiration and presence of leaf senescence to low levels of mitochondrial respiration under drought stress (Atkin and Macherel, 2008).

Photorespiration and ROS scavenging activity are linked with metabolic and oxidative damage induced by drought

In addition to metabolic alterations, the drought-induced decrease in carbon assimilation results in enhanced generation of ROS in leaf cells (Suzuki *et al.*, 2012; Noctor *et al.*, 2014). Overall, we found that the activity of two photorespiratory enzymes, HPR and GOX, strongly positively correlated with the main amino acid clusters (Fig. 2). More specifically, the amounts of Gly and Ser in the flag leaf of the 292 genotypes almost perfectly correlated with the activity of the photorespiratory enzymes HPR and GOX under stress and showed very high correlations also under control conditions (Fig. 3). These results support the hypothesis of a direct interaction between photorespiration and amino acid metabolism (Fernie *et al.*, 2013; Hodges *et al.*, 2016) and they help to shed light on the contribution of the photorespiratory pathway to the overall supply of Gly and Ser in rice leaves under stressed and non-stressed conditions (Ros *et al.*, 2013). The strength of these correlations (Fig. 3) highlights the importance of photorespiration in rice leaf central metabolism and suggests that its reduction to improve crop yield (Betti *et al.*, 2016; Walker *et al.*, 2016; South *et al.*, 2019; Weber and Bar-Even, 2019) should be approached with great care as this could lead to leaf metabolic impairments resulting in even higher yield losses under drought.

In the presence of limited CO_2 supply, the photorespiratory pathway has often been considered as a protective mechanism able to reduce ROS generation, and the consequent oxidative damage, by consuming the excess of energy produced in the chloroplast (Wingler *et*

al., 2000; Takahashi and Badger, 2011; Hodges *et al.*, 2016). However, the MDA values detected in this study do not support the ROS protective action of photorespiration. MDA is a lipid peroxidation product widely accepted as a marker of membrane oxidative damage (Møller *et al.*, 2007). In the present study, under stress conditions, MDA values display a very similar correlation pattern with the main metabolic clusters as the one of GOX and HPR (Fig. 2). This suggests that during enhanced photorespiration, a high rate of H_2O_2 production in the peroxisomes results in oxidation of cellular membrane lipids as previously described in maize leaves exposed to drought (Avramova *et al.*, 2017).

Plants have evolved a complex enzymatic and non-enzymatic antioxidative system to protect cells from the enhanced generation of ROS and their oxidative action (You and Chan, 2015). Under drought stress, we found that only the activity of some antioxidant enzymes, DHAR, CAT and APX, showed a strong negative correlation with the metabolic clusters associated with stress (Fig. 2) suggesting a specialized role of these enzymes in counteracting the adverse effects induced by drought. CAT and APX are considered the main antioxidant enzymes involved in H_2O_2 removal in leaves (Noctor *et al.*, 2014). In particular, CAT directly converts H_2O_2 to water and oxygen (Mhamdi *et al.*, 2010) and exerts most of its activity in the peroxisomes, where this enzyme counteracts H_2O_2 generation by GOX during photorespiration (Bauwe *et al.*, 2010). APX protects chloroplast membranes by reducing H_2O_2 to water through the oxidation of ascorbate (Das and Roychoudhury, 2014). An efficient regeneration of reduced ascorbate is therefore essential for H_2O_2 scavenging. DHAR regenerates the oxidized ascorbate by using reduced glutathione as electron donor (Das and Roychoudhury, 2014) and the enzyme participates, like APX, in the ascorbate-glutathione pathway, the main redox hub in plants (Foyer and Noctor, 2011). CAT and APX activity displayed a strong population-wide increase under drought stress whereas, surprisingly, DHAR activity did not (Table 1). This could indicate that DHAR is not as ubiquitously important as CAT and APX, which are upregulated across the whole population or, alternatively, other regulatory mechanisms such as posttranslational modification or allosteric interactions could control DHAR *in vivo,* but are not discriminated in the *in vitro* assay (Sulpice *et al.*, 2010). Nevertheless, differences in - *in vitro* analysed - DHAR activity in the vegetative stage between rice genotypes differing in drought tolerance have been described before (Selote and Khanna-Chopra, 2004). Conversely, and even more surprisingly, SOD, the enzyme showing the highest drought-induced increase

(Table 1), did not show any strong positive or negative association with the metabolic clusters representative of stress (Fig. 2). In plants, SODs, localized in chloroplasts, mitochondria and cytosol, catalyse the conversion of the highly oxidative anion superoxide (O_2^-) to the less harmful H_2O_2 (Halliwell, 2006). Apparently, under drought stress, SOD activity increases independent of the genotype, thus representing a commonly-shared mechanism to generate high amounts of H_2O_2 that, in turn, can be detoxified by more genotype-dependent, localized and effective antioxidative responses.

Multivariate modelling of GY loss reveals that the enzymes of the ascorbate-glutathione cycle are essential for GY stability under drought

Differently from univariate statistics, multivariate analysis considers the simultaneous relationships between all the variables of a given dataset, thus increasing its predictive power, and it was used before in metabolomics-based plant trait prediction (Meyer *et al.*, 2007; Mumm et *al.*, 2016; Steinfath *et al.*, 2010; Sulpice *et al.*, 2009, 2013). The PLSR model based on control values of the MetabOxi variables showed low predictability for GY loss (Fig. 4). This observation suggests that basal levels of flag leaf primary metabolites, oxidative stress markers/enzymes and their interactions have little influence in determining the genotypic GY loss sensitivity upon the introduction of drought stress. Despite its low predictability, the PLSR control model suggested the highest importance for erythritol and galactaric acid. Consistent with the literature (Obata and Fernie, 2012; Fàbregas and Fernie, 2019), we found increased levels of erythritol under drought (Table 1) but only its control values displayed a positive correlation with drought-induced GY loss and an even stronger correlation with flowering time differences (Table 2 and 4). Considering the almost perfect correlation between flowering time differences between the two treatments and its influence on GY loss (Supplementary Fig. S2), it seems possible that the high ranking of erythritol in the control model reflects the relative importance of flowering differences at the time of stress imposition on GY loss performance of the 292 accessions. On the other hand, control values of the other equally important model's predictor, galactaric acid, were not significantly correlated with GY loss, nor with flowering differences (Table 2 and 4). In a previous study on 21 rice genotypes, galactaric acid was described to correlate positively with plant growth, under both control and drought stress conditions (Degenkolbe *et al.*, 2013). Its presence as top predictor in the control PLSR model

might reveal a hidden link between genotypic-induced differences in plant growth rate under optimal conditions and GY loss tolerance under drought stress.

Different from the control model, the PLSR model based on drought values of the MetabOxi variables showed high predictability of GY loss (Fig. 4). This clearly indicates that the stress-induced interaction between metabolites and oxidative stress markers/enzymes is tightly linked to the GY loss performance of the population. The top predictors of the drought model highlighted the importance of the enzymatic reduction of oxidized ascorbate as key mechanism to reduce the negative effect of oxidative stress on GY loss (Table 4). Particularly, DHAR outclassed all the other predictors in terms of model contribution. Above, we have already discussed the role of this enzyme in counteracting the adverse metabolic changes induced by drought (Fig. 2 and Table 3). Interestingly, another antioxidant enzyme, MDHAR, ranked 3rd in the model. MDHAR, just as DHAR, regenerates reduced ascorbate so that it can be used again by APX in the scavenging of H_2O_2 (Das and Roychoudhury, 2014). MDHAR, by the direct conversion of monodehydroascorbate to ascorbate before it spontaneously converts to dehydroascorbate (Smirnoff, 2000), might reduce the DHAR workload thus increasing the efficiency of ascorbate reduction. For this reason, MDHAR activity under drought might have been selected as a good predictor by the PLSR model despite its null correlation with GY loss (Table 4) and this demonstrates the value of this modelling approach above simple correlation analysis. The presence of MDA among the top three predictors of the drought model reinforces the link between leaf oxidative damage and GY loss. MDA values under drought showed the highest positive correlation with GY loss (Table 3 and 4) and displayed a strong correlation with photorespiratory activity and leaf senescence (Fig. 2). All these findings suggest that MDA represents the best biomarker of GY loss sensitivity under drought and that oxidative damage of leaf membrane lipids is among the most damaging processes caused by drought.

CONCLUSIONS

The metabolic and oxidative stress profiles of the rice flag leaf changed dramatically during drought stress in the reproductive stage. These changes proved to be highly informative for the grain yield loss sensitivity of the different rice accessions at harvest time. Multivariate modelling of grain yield loss revealed that the coordinated activity of enzymes involved in the

ascorbate-glutathione cycle, and among them primarily DHAR, is an essential mechanism of drought tolerance in rice. Our study suggests that the co-expression of specific antioxidant enzymes of the ascorbate-glutathione cycle (DHAR and MDHAR) could represent a robust mechanism of tolerance that can minimize grain yield losses under drought. Finally, the genetic diversity of the 292 rice accessions used in this study offers the possibility to find genomic associations for the identified key enzymatic and metabolic determinants of grain yield loss under drought. These associations could be developed into genetic markers to be used in breeding for grain yield stability under drought in rice.

ACKNOWLEDGMENTS

This work is part of the Growing Rice like Wheat research programme financially supported by an anonymous private donor, via Wageningen University Fund, for the first author's PhD fellowship. GC-MS analysis was enabled by the Transnational Access capacities of the European Plant Phenotyping Network (EPPN, grant agreement no. 284443) funded by the FP7 Research Infrastructures Programme of the European Union. The authors wish to thank Andrea Apelt (IPK-Gatersleben) for excellent technical assistance during the GC-MS analysis.

SUPPLEMENTARY DATA

Supplementary data are available at:

https://drive.google.com/open?id=1BNJ-1qlagmNVwOM7wlqfKht691hbVIUT

Fig. S1. Experimental design of the field trial.

Fig. S2. Correlations between flowering differences and grain yield loss.

Fig. S3. Principal component analysis based on the 111 MetabOxi variables.

Fig. S4. Hierarchical clustering results for the 91 metabolites under the two treatments.

Fig. S5. Predictability of the PLSR models with increasing LVs

Table S1. Percentage of grain yield loss (GY loss %), days to 50% flowering (Flow) and leaf sampling date minus date of 50% flowering (Sam-Flow) under control (CON) and drought stress (DRO) conditions, for the 292 genotypes.

Table S2. Leaf values (log_{10} transformed) of the 111 MetabOxi variables in the 292 genotypes of the population under control and drought stress conditions.

Table S3. List of the 111 MetabOxi variables considered in this study.

Table S4. Effect of treatment on the 111 MetabOxi variables.

Table S5. Spearman rank correlation matrix of the 91 metabolites under control conditions.

Table S6. Spearman rank correlation matrix of the 91 metabolites under drought conditions.

Table S7. Spearman rank correlation matrix of the 21 oxidative stress markers/enzymes and the 91 metabolites under control conditions.

Table S8. Spearman rank correlation matrix of the 21 oxidative stress markers/enzymes and the 91 metabolites under drought conditions.

Table S9. List of the 111 MetabOxi variables ranked by significance of their Spearman rank correlation with grain yield loss (GY loss) under control and drought conditions.

Table S10. List of the 111 MetabOxi variables ranked by significance of their Spearman rank correlation with flowering at sampling (Sam-Flow) under control and drought conditions.

Table S11. Ranking of the cross-validated PLSR model based on control values of the variables.

Table S12. Ranking of the cross-validated PLSR model based on drought values of the variables.

CHAPTER

Association mapping and genetic dissection of drought-induced canopy temperature differences in rice

Giovanni Melandri, Ankush Prashar, Susan R. McCouch,
C. Gerard van der Linden Hamlyn G. Jones, Niteen N. Kadam,
S. V. Krishna Jagadish, Harro J. Bouwmeester, Carolien Ruyter-Spira

Under revision in Journal of Experimental Botany

ABSTRACT

Drought-stressed plants display reduced stomatal conductance, which results in increased leaf temperature by limiting transpiration. In this study, thermal imaging was used to quantify the differences in canopy temperature under drought in a rice diversity panel consisting of 293 *indica* accessions. The population was grown under paddy field conditions and drought stress was imposed for 2 weeks at flowering. The canopy temperature of the accessions during stress negatively correlated with grain yield (r = -0.48) and positively with plant height (r = 0.56). Temperature values were used to perform a genome-wide association (GWA) analysis using a 45K-SNP map. A QTL for canopy temperature under drought was detected on chromosome 3 and fine-mapped using a high-density imputed SNP map. The candidate genes underlying the QTL point towards differences in the regulation of guard cell solute intake for stomatal opening as the possible source of temperature variation. Genetic variation for the significant markers of the QTL was present only within the tall, low-yielding landraces adapted to drought-prone environments. The absence of variation in the shorter genotypes, which showed lower leaf temperature and higher grain yield, suggests that breeding for high grain yield in rice under paddy conditions has reduced genetic variation for stomatal response under drought.

KEYWORDS

Oryza sativa, drought, canopy temperature, thermal imaging, GWAS, haplotype analysis.

INTRODUCTION

The increasing variation in temperature, precipitation, and their interaction, resulting from global climate change is predicted to increase the variability in global crop yield by more than 30% (Ray *et al.*, 2015). Among the cereals, rice is especially sensitive to water limitation and heat stress, particularly during the reproductive stage (Jagadish *et al.*, 2007; Sandhu *et al.*, 2014). Climate change and the increasing probability of both prolonged and intermittent periods of drought are therefore likely to seriously affect rice production, particularly in rain-fed lowland farmlands which account for more than 30% of the world's total rice cultivation area (Bailey-Serres *et al.*, 2010). Thus, plant breeders aim to develop varieties with improved yield performance under both favourable and water-limited conditions (Kumar *et al.*, 2014).

Crop germplasm collections stored in gene banks worldwide represent a large and potentially valuable reservoir of favourable alleles that can be used to develop new crop varieties that provide yield stability under both favourable and stressful environments (Tester and Langridge, 2010; Huang and Han, 2014). Over the last 10-20 years, rapid improvements in the throughput and cost-effectiveness of sequencing and genotyping have made it possible to generate extensive information about plant genetic variation at the genome level. This genomic information can be combined with phenotypic data for genetic analyses. The development of phenotyping tools has not progressed as rapidly, resulting in a 'phenotyping bottleneck' (Furbank and Tester, 2011; Cobb *et al.*, 2013) which limits the genetic dissection of complex traits such as tolerance to drought stress. However, new, non-destructive, non-invasive, image-based approaches to phenotyping in both the field and in controlled environments are increasingly available, greatly enhancing the potential to phenotype large populations (Furbank and Tester, 2011; White *et al.*, 2012; Cobb *et al.*, 2013; Reynolds *et al.*, 2016). The use of indirect 'proxy' indicators for stress can be particularly useful and a powerful resource for field-based phenotyping (Jones, 2014). Among them, canopy temperature, measured by thermal imaging, has already proven to be a good indicator of drought stress in the field, as it indirectly measures stomatal conductance (Leinonen et al. 2006; Munns et al. 2010), one of the main physiological traits involved in the regulation of water loss (Schroeder *et al.*, 2001).

Several recent field studies successfully utilized infrared thermography to measure genotypic variation in stomatal conductance in a large number of genotypes (Jones *et al.*, 2009;

Rebetzke *et al.*, 2012; Zia *et al.*, 2013; Prashar *et al.*, 2013; Rutkoski *et al.*, 2016). Critical to the success of such studies was the use of appropriate normalisation techniques to overcome the environmental fluctuations (air temperature, humidity, wind speed and incident radiation) that induce variation in canopy temperature during the process of imaging. The same studies also suggest that thermal image analysis of crop canopies is maximally effective in water-limited environments, as the genotypic differences in stomatal conductance are maximised under these conditions. As a consequence, thermal imaging provides a potentially useful phenotyping strategy for the selection of drought tolerant genotypes (Jones *et al.*, 2009; Prashar *et al.*, 2013).

In the present study, we assessed the effectiveness of thermal imaging to quantify genetic variation in canopy temperature/stomatal conductance in tropical rice, using a population of 293 *indica* accessions grown in the field under control and drought conditions at the International Rice Research Institute (IRRI) in the Philippines. Statistical analyses revealed a relationship between canopy temperature during flowering, plant height and grain yield. We also report what is, to our knowledge, the first genome-wide association analysis of leaf temperature in rice, demonstrating that there is genetic variation for this trait and pinpoint genomic loci and *a priori* candidate genes that underlie this variation.

MATERIALS AND METHODS

Description of the field experiment

A population consisting of 293 accessions of *Oryza sativa* subsp. *indica* was used in a field trial experiment at the International Rice Research Institute (IRRI), Los Baños, Philippines (14°11'N, 121°15'E; elevation 21 m above sea level) during the 2014 dry season. The accessions are largely the same as those in the PRAY-indica panel (http://ricephenonetwork.irri.org) which includes traditional and improved *indica* rice lines originating from rice-growing countries in tropical and sub-tropical regions around the world. The same panel was recently used in studies where a number of diverse traits were phenotyped as the basis for GWAS (Qiu *et al., 2015;* Al-Tamimi *et al.*, 2016; Rebolledo *et al.*, 2016; Kadam *et al.*, 2017, 2018; Kikuchi *et al.*, 2017). The field trial was carried out in two separate fields, one that served as control and the other for the drought stress treatment. Each field comprised three replicates of the population (A, B, C for control and D, E, F for drought) arranged into a serpentine design

(Supplementary Fig. S1). Each replicated accession consisted of 48 plants covering 2.5*0.8m area and arranged as four rows of 12 plants each. To manage the differences in flowering phenology, the accessions were sown at 7-day intervals and transplanted to create subgroups that allowed us to synchronize flowering. Eight subgroups were created according to the number of days required to reach 50% flowering. Each group was progressively sown and transplanted into the field with an interval of 7 days between each group. Inside each subgroup, accessions were transplanted from the shortest to the tallest one to minimise the positioning of short and tall genotypes next to each other (plant height data collected during the dry season 2013). The imposed drought stress treatment consisted of 14 consecutive days of water withholding applied only to the stress field at the reproductive stage (targeting 50% flowering). Weather data was collected during the entire experiment by a weather station located in the middle of the two fields. 26 tensiometers were randomly distributed over the stress field to record soil water potential. At the end of the stress period, the stress field was re-watered until all accessions reached the maturity stage for harvest. The control field was constantly kept in a flooded condition (paddy field). At harvest (on average 30 days after re-watering) the following traits were scored for all replicates: plant height (cm), grain yield (grams/m^2) and shoot biomass (grams/m^2), harvest index (ratio between grain yield and total biomass). The dates of initial flowering, 50% flowering and 100% flowering were also recorded for replicated trials of each accession under both treatments. A more detailed description of the experiment, including the description of the same field trial conducted during the 2013 dry season, can be found in Kadam *et al.*, 2018.

Thermal imaging

A FLIR B660 (FLIR systems, USA) infrared camera was used for taking both infrared and visual images. The thermal camera is assembled with a focal plane array (FPA) uncooled microbolometer that operates in the spectral range of 7.5-13 µm with a resolution of 640 x 480 pixels. The thermal camera is also equipped with a digital camera with a resolution of 3.2 Megapixels. All pictures were taken from 3.5 m height (Jones *et al.*, 2009) with each image covering approximately 50 m^2 (Fig. 1 and Supplementary Fig. S1). The distance between the camera and the centre point of field in the image was kept constant, resulting in a camera angle of approximately 20° from the ground. Thermal pictures were taken, during the morning, 8, 9 and 10 days after the stress was applied (from 2nd to 4th April 2014). We

collected images to fully cover 'Rep B' (control), 'Rep E' and 'Rep F' (drought). For each replicate we collected images on two consecutive days at two different times during the morning period. To image an entire replicate, it took on average, 45 minutes. 'Rep B&E' were covered by 18 pictures whereas 'Rep F' by 27-28 (Supplementary Fig. S1 and Supplementary Table S1A and 1B). Camera settings were kept constant during the entire process of imaging with atmospheric temperature set to 30 °C and emissivity set at 0.95 according to Jones *et al.* 2003, Prashar and Jones 2014 and Prashar and Jones 2016.

Plot identification and picture analysis

Plot identification was achieved following the experimental design and by the use of three T-boards placed at known positions in each image (Fig. 1). In addition, every plot in the field trial was marked by a stick placed between two consecutive plots. The stick was characterized by a relatively higher temperature than the plant canopies and it was visible in the thermal images (Fig. 1B and 1C). Temperature quantification was performed by loading the images into the ThermaCAM Researcher Professional 2.10 software (FLIR systems), selecting a rectangular area for each plot canopy and using the mean temperature of the pixels in the enclosed rectangular area as representative for the specific plot (Fig. 1C and 1D). The temperature of each T-board reference surface (black and white) was determined in the same way.

Plot image normalisation

We considered three methods to normalise plot temperatures. In the first method ('image mean'), plot temperatures in each image were multiplied by the ratio between the mean temperature of all the plots in the replicate and the mean temperature of all the plots in the image. The normalised temperature of plots occurring in two consecutive images was calculated as the mean of the two resulting values (Prashar *et al.*, 2013). The second and third methods are based on the same procedure but using the mean temperature of the reference surfaces ('white reference' and 'black reference') to calculate the ratios. Correlation analysis between normalised data was performed to evaluate which normalisation method produced the highest degree of reproducibility for the same replicate imaged over two consecutive days.

Statistical analysis

Statistical analysis of the data was conducted by using R statistical software (version 3.4.3; The R Foundation for Statistical Computing). Correlation analysis and graphical matrices were produced using a modified function of the 'corrplot' R package. Box-Cox transformation of not normally distributed traits was calculated using the 'forecast' R package. SNP-based principal component analysis (PCA) was performed using the *prcomp* function in the 'stats' R package.

Genome-wide association (GWA) analysis with 45K SNP map

Residuals' distribution of the single replicates and mean values for control and drought stress replicates were analysed first. In case a replicate showed a non-normal distribution of residuals (Shapiro-Wilk's p-value < 0.05), temperature data were Box-Cox transformed before being used for association mapping (Supplementary Fig. S2). Genome-wide association studies (GWAS) were performed using a linear-mixed model in EMMAX (Kang *et al.*, 2010), which corrects for population structure by including a kinship matrix (IBS matrix) as covariate. EMMAX also provides an estimate of the phenotypic variance (pseudo-heritability, h^2) explained by the IBS matrix. Of the 293 accessions in the field experiment, 271 matched the original panel (*indica* Pray panel of 339 accessions) that was used to generate a 47K SNP map using Genotyping-By-Sequencing (GBS) (8.75% missing data imputed by Fast Phase Hidden Markov Model, Scheet and Stephens 2006) as reported by Kadam *et al.* (2017, 2018). The reduced number of accessions (271) altered the minor allele frequency (MAF) threshold of the 339 accessions panel, originally set at 0.05. To exclude rare alleles from the present study (n=271 accessions), the 47K SNP map was re-filtered for MAF > 0.05, resulting in 45,505 SNPs available for GWAS. Principle Component Analysis (PCA) based on the 45,505 SNPs was conducted to quantify subpopulation structure (Fig. S3). The main component (PC1) explained only 8.72% of the genetic variation and a combination of the first three components failed to clearly separate groups of accessions. Therefore, no PC covariates were added to the linear-mixed model to correct for subpopulation structure, following the approach used in McCouch *et al.*, (2016). GWAS results are presented as Manhattan and Quantile-Quantile plots using the 'qqman' R package. To avoid Type 1 error, a stringent significance threshold of $p < 0.00001$ (i.e. $-\log_{10} p > 5.0$) was used to identify marker-trait associations. This significance threshold

was higher than in other GWAS on rice using SNP maps of similar density (Zhao *et al.*, 2011; Dimkpa *et al.*, 2016; Kadam *et al.*, 2017, 2018).

GWA analysis using a high-density imputed SNP map

Imputation of the 47K SNP map was conducted using the Rice Imputation Server (RIS) following Wang *et al.* (2018). The 47K SNPs map (.hmp format) was first converted to Plink format (.ped/.map) and then to Oxford format (.gen/.sample) before being uploaded as a compressed folder (.tar.gz) in the RIS (http://rice-impute.biotech.cornell.edu). The RIS-imputed map was downloaded as single Plink file (.bed/.bim./.fam format) and divided into 12 individual chromosomes. Focusing only on chromosome 3, missing SNPs were imputed with Beagle version 4.1 (Browning and Browning, 2007). Finally, the Beagle-imputed map of chromosome 3 was filtered at MAF > 0.05, resulting in a set of 186,012 SNPs available for mapping on this chromosome. GWA analysis using the chromosome 3 imputed map was conducted as described above for the 45K SNP map. The IBS matrix of kinship used as a model covariate was calculated based on the 45K GBS SNP map (Wang *et al.*, 2018).

Linkage disequilibrium (LD) analysis and *a priori* candidate gene selection

Local pairwise linkage disequilibrium (LD) pattern near the significant SNPs was calculated and graphically represented by the 'snp.plotter' R package (Luna and Nicodemus, 2007). The annotations of genes located within LD blocks were obtained from the MSUv7 rice genome database (http://rice.plantbiology.msu.edu/). Exact localization and functional annotation of significant SNPs was conducted using SNPEff version 2.05 (Cingolani *et al.*, 2012) with MSUv7 as the reference genome.

RESULTS

Thermal imaging and data normalisation

Fluctuations in environmental conditions are the main obstacle to the use of thermal imaging to reliably analyse plant canopy temperature. Supplementary Table S1B shows the changes in air temperature, humidity, wind speed, and solar irradiance measured on the days and in the time windows during which the thermal imaging was performed. To reduce the overall effect of these factors on the analysis of canopy temperature, we imaged the field replicates only during the mornings on three consecutive days. Mornings were selected for imaging due

to the sharp increase in wind speed experienced every afternoon in the field location, and previous reports describing wind as a major factor strongly impacting stomatal conductance values (Jones, 1999; Maes and Steppe, 2012). Nevertheless, our data documented variation caused by environmental fluctuations between images (Supplementary Fig. S4) and therefore normalisation was needed (Prashar and Jones, 2014). We applied three different procedures to reduce the variation caused by environmental fluctuations ('image mean', 'white reference' and 'black reference') (see Material and Methods). By comparing the same field replicate imaged over two consecutive days we found that 'image mean' normalisation produced higher correlation values (B03-B04: from 0.075 to 0.37; E02-E03: from 0.23 to 0.69; F03-F04: from 0.25 to 0.65) than normalisation based on white and black references (Supplementary Fig. S5). Hence, 'image mean' normalisation was used in all subsequent analyses.

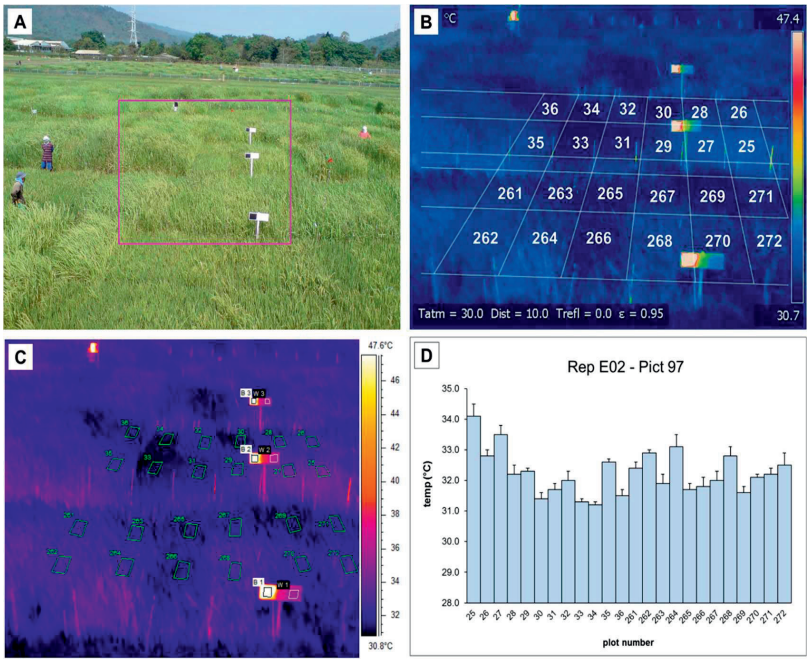


Fig. 1. Plot identification and picture analysis.
Example of a digital picture with the corresponding thermal picture area as indicated by a rectangle (A). Thermal picture used for plot identification (B). Selection of specific polygonal areas for the quantification of the genotypes' temperature (C). Temperature of the polygonal areas selected in C (D).

Canopy temperature

Drought stress induced a strong increase in canopy temperature. The mean value of the normalised stressed replicates E and F was 2.27 °C higher than that of the control B replicate (Fig. 2). Together with the treatment, the time of day strongly impacted leaf temperature. Control replicate B03, which was imaged in the early morning, showed a lower canopy temperature range (mean diff.: -0.90 °C) than B04 which was imaged in the late morning (approx. 2 hours later) the following day. As with the control replicates, drought replicate E02 showed a lower temperature (mean diff.: -1.85 °C) than E03. The temperature difference between the two E replicates was due to the combined effect of time of day (E02 measured earlier in the morning than E03) and an additional day of drought stress. Tensiometer readings showed that the soil water potential of the drought field decreased sharply during the days of imaging, moving from an average of -34 kPa on 2 April to -53 kPa on 4 April (Supplementary Fig. S6). Of the two stress replicates, F03 was imaged half an hour earlier than F04 but the latter was exposed to one more day of stress, resulting in a higher temperature for F04 (mean diff.: 0.33 °C) than for F03.

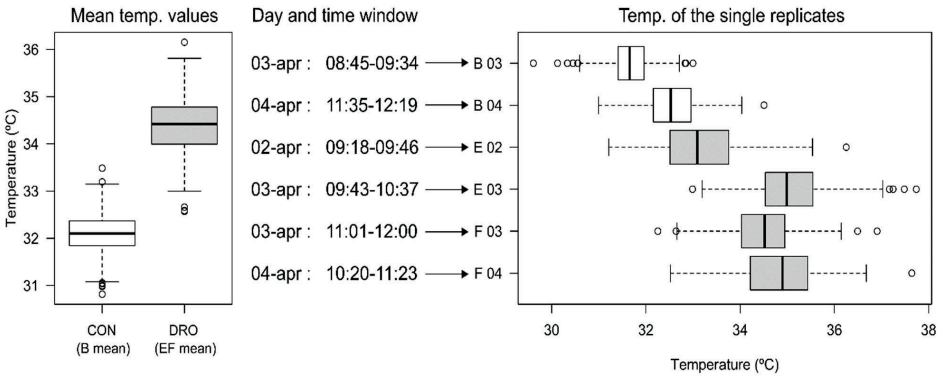


Fig. 2. Canopy temperature of the different field replicates.
Boxplots representing the mean temperatures of control (B mean) and drought (EF mean) replicates (on the left), and of the six single replicates (on the right), together with the information on the days and time windows of picture taking (in the middle). White and grey boxplots are representing control and stress values, respectively.

Relationships between canopy temperature and agronomic traits

Canopy temperature was measured in the second of two years (2013 and 2014 dry seasons) in which a field experiment was conducted to collect information on phenotypic trait performance of all 293 rice accessions evaluated under well-watered and drought-stress

conditions (Kadam *et al.*, 2018). Plant height (PHT), grain yield (GY), shoot biomass (Shoot bio.) and harvest index (HI) were among the agronomic traits recorded at the time of harvest in both years. A 'flowering' variable (FLW) was calculated during the stress period of the 2014 field trial only, by subtracting the date of 50% flowering for every genotype in each replicate from the date of thermal imaging. Drought stress similarly affected all traits in both years (Supplementary Fig. S7) and particularly reduced GY and HI, but minimally affected Shoot bio. and FLW (2014 only). To investigate the relationship between canopy temperature at the time of stress exposure and plant traits at harvest time, we conducted a Spearman correlation analysis between normalised temperature and agronomic trait values (Fig. 3).

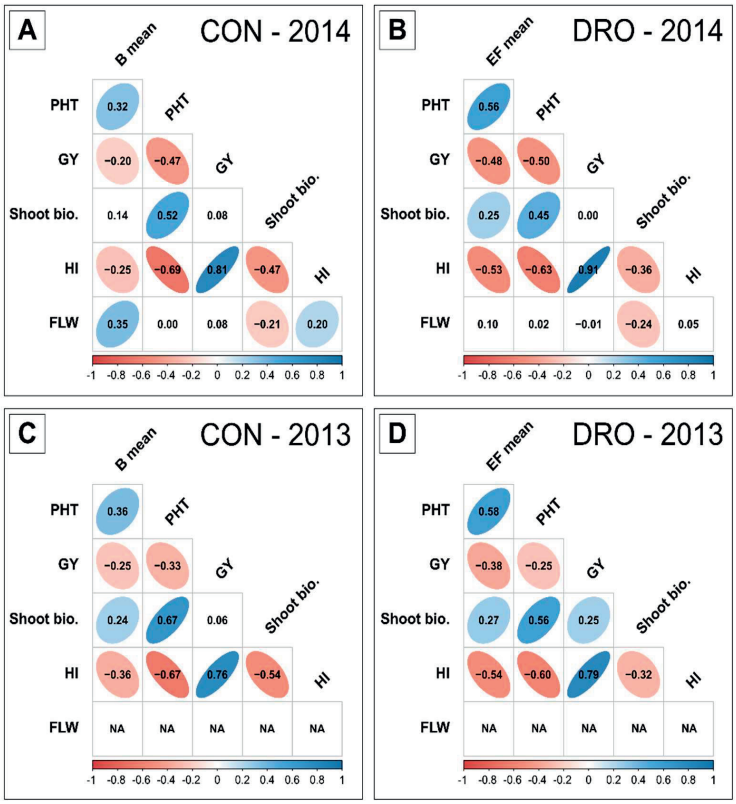


Fig. 3. Correlation analysis between mean values of canopy temperature and agronomic traits. Mean canopy temperature values selected under control (CON) and drought stress (DRO) conditions were determined in the 2013 dry season while agronomic traits were determined in both field trials conducted during the 2013 and 2014 dry seasons. Correlation values were determined by Spearman's correlation. White cells indicate a non-significant correlation value ($p > 0.001$). PHT: plant height; GY: grain yield; Shoot bio.: shoot biomass; HI: harvest index; FLW: flowering; B mean: mean temperature of control field replicates; EF mean: mean temperature of drought field replicates.

Canopy temperature negatively correlated with GY and HI and positively with PHT, Shoot bio. and FLW, under both conditions but with higher correlation coefficients under stress than control; FLW was not significantly correlated with canopy temperature under drought (Fig. 3A and 3B). Remarkably, almost identical correlations were found between canopy temperature measured in the 2014 dry season, and the agronomic traits scored during the field trial conducted in the 2013 dry season (Fig. 3C and 3D). In both years, the highest correlation coefficients were found between the mean values of canopy temperature under drought stress and PHT (r = 0.56 in 2014; r = 0.58 in 2013), GY (r = -0.48 in 2014; r = -0.38 in 2013) and HI (r = -0.53 in 2014; r = -0.54 in 2013). For these three traits, the percentage of variance explained by the linear models (R^2) associating temperature and trait was equal to 34% for PHT in both years, 25% and 16% for GY, 30% and 28% for HI in 2014 and 2013, respectively (Supplementary Fig. S8). Overall, these results show that, under drought stress, thermal imaging of rice canopies at flowering time can detect canopy temperature differences that correlate with plant performance at the time of harvest. Furthermore, the almost identical correlations between the agronomic traits scored during the 2013 field trial and canopy temperature measured in 2014 indirectly validate the robustness of the temperature results across two seasons of field trials.

GWAS and LD analysis using the 45K SNP map

The results described above demonstrate the effectiveness of thermal imaging in detecting quantitative differences in canopy temperature. Hence, we decided to try to use canopy temperature as a trait for association mapping analysis. Considering the strong influence of time of day and the day itself on canopy temperature (Fig. 2), the values of the separate field replicates were analysed, in addition to the mean temperature values of control and drought replicates. GWA mapping was conducted using a 45K SNP map and a stringent threshold of genome-wide significance ($-\log_{10} p > 5.0$) to detect only highly significant marker-trait associations. Quantile-Quantile plots relative to the GWA analyses are reported in Supplementary Fig. S9.

Mean temperature variation of control replicates (B mean) as well as the temperature of separate control replicates (B03 and B04) was characterized by a low fraction of pseudo-heritability (h^2 = 0.17~0.24) (Supplementary Table S2), and, indeed, GWA analysis did not find

marker-trait associations above the threshold of significance (Supplementary Fig. S10). Low heritability levels and absence of significant marker-trait associations suggest that no major genetic determinants are responsible for temperature variation under control conditions in this population.

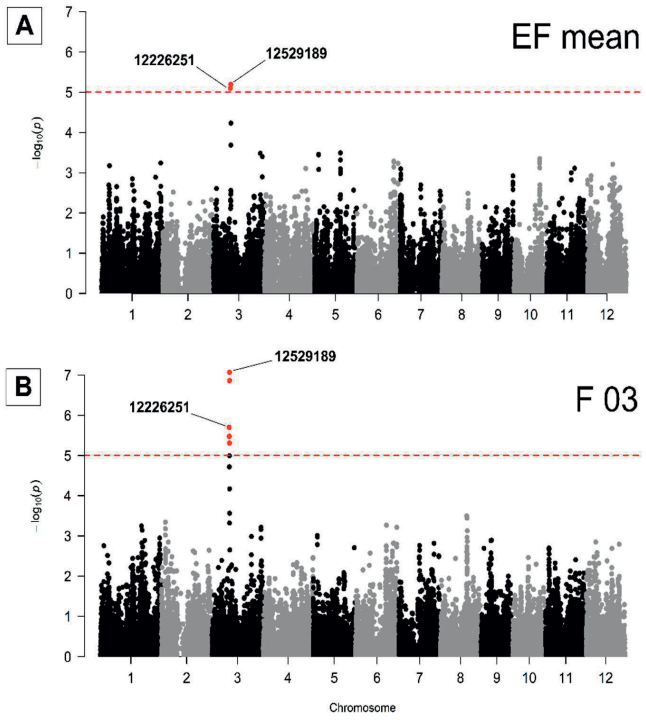


Fig. 4. Manhattan plots of the significant GWA mapping results using the 45K SNP map.
Manhattan plots of the GWA mapping results using the 45K SNP map for mean canopy temperature under drought (EF mean) and for the single field replicate F03. The red dashed line indicates the genome-wide threshold for significant associated markers ($-\log_{10} p > 5.0$). SNPs above the red line are highlighted in red.

Pseudo-heritability of mean temperature values under drought, 'EF mean', was equal to 0.5 (Supplementary Table S2) and, by GWA mapping, two significant markers ($-\log_{10} p > 5.0$) were identified on chromosome 3 (Fig. 4A). The distance between the two markers (SNP_12262251 and SNP_12529189) is 267 kbp, they show a similar level of significance and their minor alleles are associated with higher canopy temperature values than the major alleles (Supplementary Table S3). Similar to their mean, canopy temperature values of the separate drought stress

replicates displayed a substantially higher pseudo-heritability than control replicates (Supplementary Table S2). F03, the field replicate characterized by the highest heritability (h^2 = 0.45), was the only one showing significant marker-trait associations by GWA mapping (Fig. 4B and Supplementary Fig. S10). The five significant markers identified for F03 are located in the same region of chromosome 3 defined by the two significant markers identified for 'EF mean' (Fig. 4A). These two markers are commonly shared as significant between F03 and 'EF mean'. Among them, SNP_12529189 is the top marker in both replicates and for F03 it is highly significant with a LOD score above the Bonferroni corrected threshold ($-\log_{10} p > 5.96$). The minor alleles of all five significant markers of F03 are associated with higher canopy temperature values than the major alleles (Supplementary Table S3). The higher heritability of canopy temperature values under drought suggests a stronger genetic control for the trait under this stress condition. Mapping results indicate that the significant marker-trait associations on chromosome 3 are largely determined by the temperature results of replicate F03. Nevertheless, the detection of two significant markers in the same region using the drought mean values suggests that the allelic effect on canopy temperature for these markers is similar to F03 also in the other replicates, despite the fact that the LOD score in those replicates is below our stringent significance threshold. In agreement with this hypothesis, the minor alleles of both markers SNP_12529189 and SNP_12262251 were associated with higher canopy temperature values than the major alleles in all the individual field drought replicates (Supplementary Table S4).

Local linkage disequilibrium (LD) analysis was conducted considering a region of 500 kbp upstream and downstream of the two significant markers (SNP_12262251 and SNP_12529189). The pairwise LD estimates (r^2) of the 102 SNPs in this region (~1.2 Mbp) revealed that the two significant markers map to different LD blocks (r^2 > 0.6-0.7) (Supplementary Fig. S11A). Only 16 SNPs are present in the more localized region (~600 Mbp), including the two LD blocks. It is noteworthy that SNP_12262251 delimits its own LD block, with its closest upstream marker (SNP_11994173) located ~ 270 kbp away, and that these two markers show a very low value of pairwise LD (Supplementary Fig. S11B). A low density of GBS markers around the two significant SNPs does not allow a precise determination of the LD configuration, and leads to a likely underestimation of the size of the LD block containing SNP_12262251.

Fine mapping of chromosome 3 using a high-density imputed SNP map

GWA mapping of 'EF mean' canopy temperature values was performed again only for chromosome 3 using a high density SNP map generated by haplotype-based genetic imputation using the RIS (Wang *et al.*, 2018). Following imputation, the number of SNP markers available for mapping of chromosome 3 increased 46X, from 4,039 (on the 45K SNP map) to 186,012 on the high-density imputed map. Overall, the imputed map drastically improved the mapping resolution. The number of significant SNPs ($-\log_{10} p > 5.0$) associated with the 'EF mean' canopy temperature increased from 2 (45K map) to 65 (Fig. 5A and Supplementary Table S5). All newly imputed markers localized in the same region as the two significant markers (SNP_12262251 and SNP_12529189) previously identified using the GBS map (Supplementary Table S5).

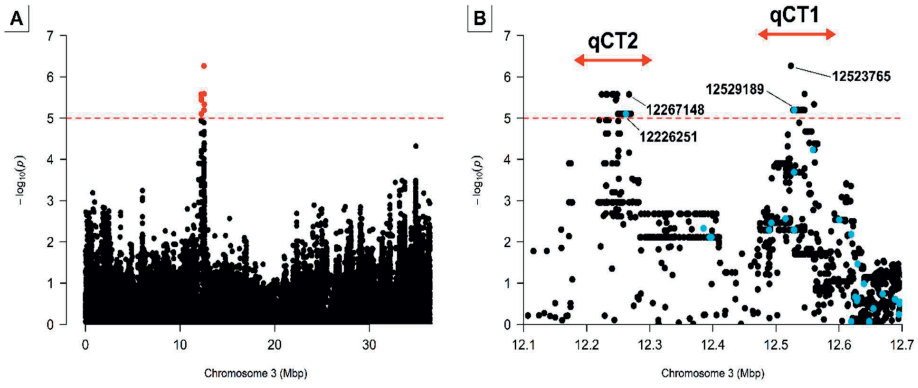


Fig. 5. Manhattan plots of the GWA mapping results for chromosome 3 using the imputed map. Manhattan plot of GWA mapping results for chromosome 3 using 186K SNPs for 'EF mean' canopy temperature. SNPs above the red line are highlighted in red (A). Zoom in in the region of the significant SNPs. The red arrows indicate the two QTLs (*qCT1* and *qCT2*). SNPs of the GBS map are highlighted in light blue (B). In both figures, the red dashed line indicates the genome-wide threshold for significant associated markers ($-\log_{10} p > 5.0$).

Zooming into the region of significance (12.1-12.7 Mbp) on the high-density map showed two distinct QTLs, *qCT1* and *qCT2* (Fig. 5B). Previously, *qCT2* could not be clearly detected due to the lack of markers around SNP_12262251 (Supplementary Fig. S11B). The significant SNPs define both QTLs are characterized by minor alleles associated with higher canopy temperature values than the major alleles (Supplementary Table S5). In the imputed high-

density map, the top SNP associated with *qCT1*, SNP_12523765 ($-\log_{10} p = 6.26$), is 5.4 kbp away from the top SNP identified using the GBS map (SNP_12529189, $-\log_{10} p = 5.19$). In *qCT2*, a series of 15 SNPs show the same significance ($-\log_{10} p = 5.57$), and among them, SNP_12267148 is closest (4.9 kbp away) to the top SNP on the GBS map, SNP_12262251 ($-\log_{10} p = 5.10$) (Fig. 5B and Supplementary Table S5). Thus, the imputed map supports the location of the QTL identified on chromosome 3 using the 45K SNP map, but provides improved resolution, making it possible to differentiate two, closely linked genomic associations.

To determine if the two QTLs are independently associated with canopy temperature, we again performed GWA mapping, but this time we fitted the most-significant SNP of *qCT1,* and subsequently of *qCT2,* as a covariate in the linear-mixed model. Fixing either SNP_12523765 (qCT1) or SNP_12267148 (*qCT2*) as model covariate yielded a similar loss of signal from all markers in the region ($-\log_{10} p < 3$), causing both QTLs to disappear (Supplementary Fig. S12). This suggests the presence of linkage between the two QTLs.

Haplotype analysis of *qCT1-qCT2* and their relationships with agronomic traits

Using 20 significant SNPs within the two QTLs, we analyzed the regional haplotypes to determine whether recombinant genotypes display a different association with canopy temperature, as well as with two agronomic traits, PHT and GY that we showed to be significantly correlated with the temperature results (Fig. 3). For the analysis, we selected the imputed SNPs showing the highest association with 'EF mean' in the two QTLs (3 SNPs for *qCT1* and 15 SNPs for *qCT2*) plus the two most significant GBS SNPs (SNP_12529189 and SNP_12262251, in *qCT1* and *qCT2*, respectively). In total, we identified 3 major haplotypes (present in at least 5% of the 246 accessions analyzed) (Table 1). The most common haplotype, *Haplotype I*, is harbored by 206 accessions that are fixed for the major alleles at the 20 SNPs defining the two QTLs. These accessions always display a lower canopy temperature (Supplementary Table S6). The other two haplotypes, *Haplotype II* and *Haplotype* III, are harbored by 22 and 18 accessions, respectively. *Haplotype III* is the mirror image of *Haplotype I,* carrying the minor alleles at all 20 SNPs for the two QTLs.

Table 1. Haplotype analysis of *qCT1-qCT2* and relationship with agronomic traits.
Haplotypes for *qCT1* and *qCT2* identified among the accessions considering 'EF mean' canopy temperature values. The 20 most significant SNPs for the two QTLs are listed together with the information relative to the minor and major allele (min./maj.), physical position (bp) and significance ($-\log_{10} p$). For each haplotype (I, II, III) are reported the number of lines carrying the haplotype and their mean value of canopy temperature (Temp.), plant height (PHT) and grain yield (GY) under control and drought stress conditions.

	qCT2																*qCT1*									
min./maj.	C/G	C/T	C/T	C/T	G/A	G/A	C/T	G/A	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/A	C/T	A/T	G/A	A/G	G/A					
Position	12223622	12230423	12230722	12232636	12240232	12240847	12240881	12241732	12242733	12243764	12248148	12248149	12248238	12248347	12262251	12267148	12523765	12529189	12545362	12560502						
Sign.	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.57	5.10	5.57	6.26	5.19	5.59	5.33	**Haplo type**	**No. of lines**	**DROUGHT**			**CONTROL**		
																							Temp. (°C)	**PHT (cm)**	**GY (gr./m²)**	**Temp. (°C)**	**PHT (cm)**	**GY (gr./m²)**
	G	T	T	T	A	A	T	A	T	T	T	T	T	T	A	T	T	A	G	A	**I**	**206**	34.24	119	296	32.02	132	475
	G	T	T	T	A	A	T	A	T	T	T	T	T	T	A	T	A	G	A	G	**II**	**22**	34.85	137	169	32.39	156	382
	C	C	C	C	G	G	C	G	C	C	C	C	C	C	C	C	A	G	A	G	**III**	**18**	35.05	135	187	32.29	152	417

Accessions carrying this haplotype always display a higher canopy temperature (Supplementary Table S6). Interestingly, *Haplotype II* is a recombinant between *Haplotype I* and *Haplotype III* and carries major alleles for the SNPs in *qCT2* (like *Haplotype I*) and minor alleles for the SNPs in *qCT1* (like *Haplotype III*). Therefore, the recombination breakpoint of *Haplotype II* falls between *qCT1* and *qCT2*. Comparing the phenotypic performance of the accessions carrying *Haplotype II* with the performance of those carrying *Haplotypes I* and *III* can offer insight into the separate effect of the two QTLs on canopy temperature, PHT and GY. For this purpose, and considering the unbalanced sample sizes of the haplotype groups, we conducted a series of Welch's t-tests (Fig. 6).

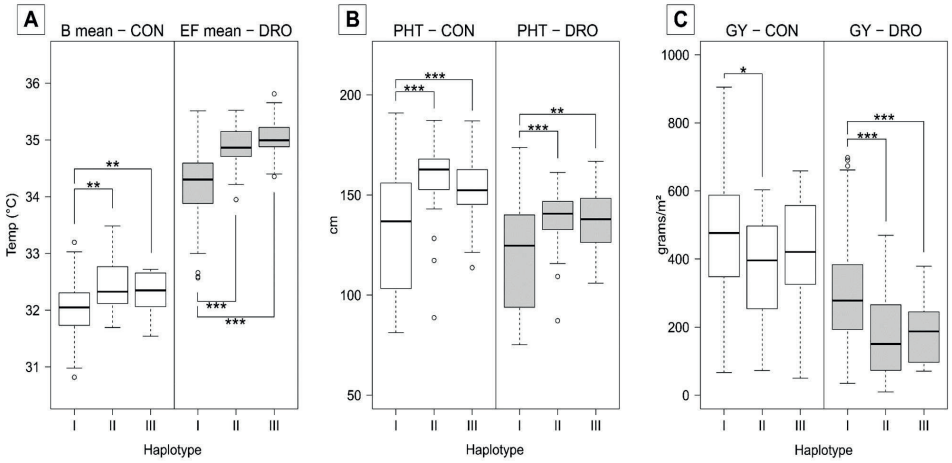


Fig. 6. Canopy temperature, plant height and grain yield performance of the accessions carrying the different *qCT1*-*qCT2* haplotypes.
Boxplots representing the range of variation of mean canopy temperature (A), plant height (B) and grain yield (C) for the accessions of the three haplotypes (I, II, III) under control and drought stress conditions. White and grey boxplots are representing control and stress values, respectively. *, **, *** indicate the level of significant difference ($p < 0.05$, $p < 0.01$, $p < 0.001$) between the groups. 'B mean': mean canopy temperature under control conditions; 'EF mean': mean canopy temperature under drought conditions; PHT: plant height; GY: grain yield.

A highly significant ($p < 0.001$) canopy temperature difference under drought (EF mean) was detected between accessions carrying *Haplotypes I* and *III* (mean 34.24 and 35.05 °C, respectively). Interestingly, the same significant difference for 'EF mean' temperature was shown between those carrying *Haplotypes I* and *II* (mean 34.85 °C) but no difference was

detected between accessions carrying *Haplotypes II* and *III* (Fig. 6A). The very similar mean temperatures of *Haplotypes II* and *III* (recombinant for the two QTLs) suggests that the main effect on canopy temperature under drought is exerted by *qCT1*. Similar results, but less significant ($p < 0.01$), were found for canopy temperature under control conditions (B mean). Under control conditions, the three haplotype groups displayed lower temperature differences (mean of *Haplotype I* = 32.02 °C, *Haplotype II* = 32.39 °C and *Haplotype III* = 32.29 °C) than under drought stress. We detected an almost identical pattern of significant differences between the three haplotype groups for PHT under both treatments (Fig. 6B). These differences are dependent on the fact that mean PHT of accessions carrying *Haplotype II* and *III* is ~20 cm higher than in those carrying *Haplotype I* (Table1).

Among the 48 accessions shorter than 100 cm (based on PHT in control conditions), 47 (98%) carry *Haplotype I* and only a single accession carries *Haplotype II* (cv Binulawan from Philippines) while none carries *Haplotype III* (Supplementary Table S6). The presence of only tall accessions among those carrying *Haplotypes II* and *III* suggests that the short semi-dwarf accessions of this population are fixed for *Haplotype I* and, thus, carry only major alleles at *qCT1* (associated with lower canopy temperature). Finally, we considered the differences in GY between the accessions of the different haplotype groups (Fig. 6C). Those carrying *Haplotypes II* and *III* showed a lower mean GY under control conditions (mean: 382 and 417 grams/m^2, respectively) than accessions carrying *Haplotype I* (mean: 475 grams/m^2) (Table 1), but these differences were either not significant, or only slightly significant ($p < 0.05$) (Fig. 6C). Under drought stress, GY differences between *Haplotype I* (mean: 296 grams/m^2) and the other two groups (mean of *Haplotype II* = 169 and *III* = 187 grams/m^2) became highly significant ($p < 0.001$) (Fig. 6C) suggesting a negative effect of the minor alleles at *qCT1* for GY under stress.

LD analysis of *qCT1* and *a priori* candidate gene identification

LD analysis of *qCT1* was performed to identify possible candidate genes underlying the LD blocks surrounding the most significant SNPs associated with 'EF mean' temperature. We considered SNPs of the imputed map showing values of association of $-\log_{10} p > 2$ in regions 125 kbp upstream and 125 kbp downstream of the most significant marker, SNP_12523765.

Pairwise LD estimates (r^2) of the 418 SNPs in this region showed that all the SNPs in *qCT1* form a single, large LD block characterized by r^2 values greater than ~0.5 (Fig. 7A). However, the most significant marker of the QTL is preceded by many markers with lower significance ($-\log_{10} p < 5$) and all map within a sub-block characterized by high pairwise LD ($r^2 > 0.8$). Furthermore, the 20 most significantly associated SNPs with 'EF mean' ($-\log_{10} p > 5$, red in Fig. 7A) are localized between two highly significant markers, SNP_12523765 and SNP_12460502 within *qCT1* (dashed black line in Fig. 7A). Thus, we considered the region delimited by these two markers (~42 kbp) as the most interesting for the identification of *a priori* candidate genes. Seven genes are included in this region (Fig. 7B and Supplementary Table S7).

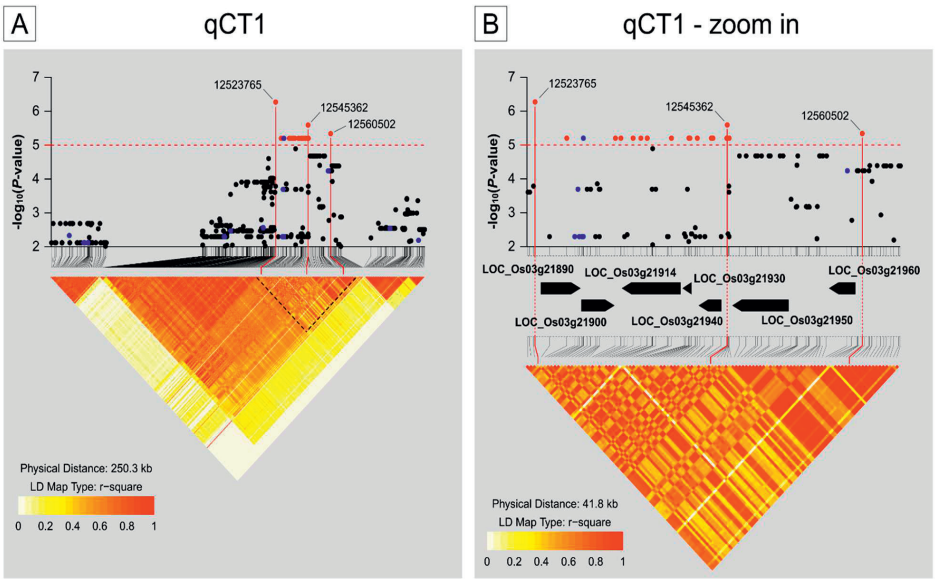


Fig. 7. Localized linkage disequilibrium analysis of *qCT1*.
Manhattan plots displaying the level of significance (y-axis) over genomic positions (x-axis) in a window of 125 kbp upstream and downstream of the marker (12523765) most significantly associated with canopy temperature of 'EF mean' and located on chromosome 3 in *qCT1* (A). Localized region (zoom in, black dashed triangle in A) showing the genes (black arrows) underlying the most significantly associated ($-\log_{10}$ *P*-value > 5.0) markers' loci (B). Different colors are used to represent the pairwise LD estimates (r^2) for each genomic location. Genomic locations of the most significant markers are projected on the LD matrix and on gene positions (only in B) by red lines. SNPs of the GBS map are highlighted in dark blue.

Among these 7, two were considered interesting candidate genes for their possible role in physiological processes regulating stomatal function. One gene is known to be responsive to abiotic stress (GOSlim terms) and encodes a putative mitochondrial fumarate hydratase (LOC_Os03g21950). This gene is located between SNP_12545362 and SNP_12560502, two of the three most significant markers in *qCT1* (Fig. 7B). Fumarate hydratase (fumarase) is responsible for the conversion of fumarate to malate, a solute involved in the mechanism of stomatal opening/closure. The other interesting gene in this region (LOC_Os03g21890) encodes for a plasma membrane high-affinity potassium (HAK) transporter (Bañuelos *et al.*, 2002) and is located in close proximity to the most significant marker (SNP_12523765) within *qCT1* (Fig. 7B). HAK transporters are involved in guard cell K^+ flux that controls stomatal opening/closure (Jezek and Blatt, 2017). Finally, we considered the predicted functional effect of each of the 20 significant SNPs ($-\log_{10} p > 5$) in this region to determine whether any of them could result in an amino acid change or a putative regulatory change affecting a specific gene candidate (Supplementary Table S8). None of the SNPs were associated with predicted non-synonymous mutations that could point towards a particular candidate, but many are located upstream of the gene coding region and thus potentially associated with changes in regulation of gene expression, including the most significant one, SNP_12523765, which is located 711 bp upstream of the HAK transporter.

DISCUSSION

Normalisation and physiological implication of canopy temperature results

The main determinants of canopy temperature in plants include genetic components affecting stomatal aperture and canopy structure (which may also affect aerodynamic resistance and radiation interception) and a range of environmental factors. The ability to reduce environmental fluctuations in humidity, irradiance and wind speed is key to screening for true genetic variation in stomatal conductance (Prashar *et al.*, 2013). Jones *et al.* (2009) demonstrated that the variation in canopy temperature between different rice genotypes can be detected in field experiments by thermal imaging with the use of appropriate normalisation techniques. In this paper we took this approach to the next level by screening a diversity panel consisting of 293 *indica* accessions. To date, only a few studies have followed a similar

extensive approach using thermal imaging of plant canopies under field conditions (Zia *et al.*, 2013; Prashar *et al.*, 2013; Rutkoski *et al.*, 2016). Normalisation of the raw canopy temperature data by 'image mean' reduced the influence of environmental factors. This is clearly shown by the increased correlation between the normalised canopy temperature data of the same field measured over two consecutive days (Supplementary Fig. S5A and 5B). Our results show that drought stress strongly increased canopy temperature. Other factors that may significantly affect canopy temperature are the time window and day of imaging and, for the stressed replicates only, the drought exposure time (Fig. 2). These results confirm that canopy temperature determined by thermal imaging is a reliable field proxy for stomatal conductance (Prashar and Jones, 2014; Jones, 2014), and that this trait is characterised by a very dynamic response to changing environmental conditions (Vico *et al.*, 2011; Drake *et al.*, 2013). The results of this study also suggest that water limitation reduced this dynamic response, as evidenced by the higher correlation coefficients under drought between the temperatures for the same field replicate imaged at different moments (Supplementary Fig. S5B).

Canopy temperature and agronomic traits at harvest time

Genotypic variation in stomatal conductance in rice may be responsible for differences in photosynthesis, even under optimal growing conditions (Ohsumi *et al.*, 2006; Ouyang *et al.*, 2017). Reduction of stomatal conductance is a well described physiological response to drought stress in rice (Centritto *et al.*, 2009; Ji *et al.*, 2012). The resulting limitation in leaf CO_2 diffusion has been shown to cause grain yield reduction in rice genotypes grown under water-limited conditions with the stress imposed at the flowering stage in the field (Centritto *et al.*, 2009; Lauteri *et al.*, 2014). There is evidence - obtained with a limited number of genotypes - that canopy stomatal conductance monitoring by thermal remote sensing at the flowering stage could be an effective criterion for the selection of high-yielding rice genotypes (Horie *et al.*, 2006). In the present study, canopy temperature measurements were used to screen 293 rice accessions for stomatal conductance differences at anthesis and to explore correlations between the genotypic variation in canopy temperature and several plant agronomic traits.

Canopy temperature was strongly and positively correlated with plant height, and an equally strong negative correlation was found with grain yield and harvest index, particularly under

stress conditions (Fig. 3 and Supplementary Fig. S8). It is interesting to note that very similar correlations were found between canopy temperature and agronomic traits scored in both years of field trials (Fig. 3), even if canopy temperature was measured one year only. The consistency of these correlations indirectly reinforces the effectiveness of our normalisation procedure in reducing the influence of environmental factors on canopy temperature results. It can therefore be assumed that the effective detection of true genotypic differences in canopy temperature in one year could be similarly correlated with the performance of the accessions for agronomic traits, such as plant height and grain yield, over the two years of field trials.

The stronger correlations we found between canopy temperature and agronomic traits under drought stress support the idea that stomatal conductance is more important for plant performance under stress than under optimal conditions, as reported in previous large field studies in other C3 and C4 cereals. For example, a similar negative correlation between canopy temperature in the reproductive stage and grain yield was described for segregating bi-parental wheat populations (Saint Pierre *et al.*, 2010). Zia *et al.* (2013) also found a negative correlation between canopy temperatures at anthesis, and grain yield in 150 maize single cross-hybrids under water-limited conditions.

The positive correlation between plant height and thermal data found in this study (Fig. 3) is contrary to what has been observed for other crops including wheat (Giunta *et al.*, 2008; Rebetzke *et al.*, 2012) and potato (Prashar *et al.*, 2013). The negative correlation between canopy temperature and plant height observed in the latter studies was interpreted in terms of an atmospheric temperature profile where an increased aerodynamic resistance in the shorter genotypes was responsible for their higher leaf temperature (Rebetzke *et al.*, 2012). It is unclear why this does not apply in our rice trial, but it might be that the taller genotypes really do have more closed stomata than the shorter genotypes, with this effect overriding any aerodynamic effect of height. This hypothesis would need to be tested using direct stomatal conductance measurements, for instance using porometer measurements. The tendency for shorter rice genotypes to have more open stomata may be linked to the fact that the new, semi-dwarf and high-yielding varieties were selected under irrigated conditions

(Pingali, 2012; Kumar *et al.*, 2014) without considering water as a limiting factor, unlike in many wheat production environments.

Another aspect that may confound the interpretation of the thermal imaging data is the effect of flowering on the canopy temperature measurements. Our dataset allowed the quantification of the contribution of differences in flowering stage to canopy temperature results. The accessions of the panel were sown and transplanted to the field in a staggered way to synchronize phenology with the aim of imposing stress at 50% flowering in all the varieties. Despite the good synchronization of flowering observed among genotypes (Kadam *et al.*, 2018), a perfect synchronization is difficult to achieve with such a large and diverse panel, grown under varying conditions. Indeed, drought affects flowering time and in many cases accelerates it, a phenomenon referred to as drought escape (Zhang *et al.*, 2016). The significant correlation between canopy temperature and flowering stage in control plants (Fig. 3A) is in agreement with a 2°C higher temperature in unstressed wheat canopies with spikes, compared to unstressed canopies without spikes (Hatfield *et al.*, 1984), and suggests that the quantification of plant canopy temperature can be significantly affected by differences in flowering. However, flowering time differences did not significantly affect the canopy temperature under drought (Fig. 3B). This result may be explained by the fact that under water-limited conditions the rise in leaf temperature may be larger than the increase resulting from the presence of panicles.

Association mapping and QTL identification

In the present study we used the observed phenotypic variation for rice canopy temperature under control and drought conditions in a GWA mapping experiment to identify the genetic factors contributing to this variation. The genotypic differences observed in canopy temperature under the two treatments show that there is substantial genetic variation, which is especially visible under drought. Mapping results indicate a low pseudo-heritability under control conditions both for the single replicates and their averaged canopy temperature (Supplementary Table S2). This may indicate that either the environmental noise masks the genetic factors under control conditions, or that the absence of stress did not trigger their

expression. The absence of significant marker-trait associations for mean canopy temperature values under control conditions (Supplementary Fig. S10) indirectly confirm this.

In accordance with other reports (Jones *et al.*, 2009), thermal data under drought stress showed good heritability, substantially higher than under control conditions (Supplementary Table S2). This suggests that stress maximises the genotypic differences in canopy temperature, which therefore can be more effectively detected by thermal imaging. A main marker-trait association was identified on chromosome 3 using the mean temperature values of the single stress replicates and the 45K SNP (GBS) map (Fig. 4). Mapping of chromosome 3 with the imputed high-density SNP map increased the signal strength of the marker-trait association and even more the mapping resolution (Fig. 5), supporting the idea that imputation is a quick and cost-effective tool for adding value to existing genotyped panels (Wang *et al.*, 2018). The increased resolution of the imputed map allowed the identification of two distinct, neighbouring QTLs (Fig. 5B). This distinction could not be resolved using the GBS map because of the low density of markers across the region (Supplementary Fig. S11). Haplotype analysis of recombinant accessions carrying either major or minor SNP alleles across the two QTL regions revealed that only one of them, *qCT1*, is responsible for canopy temperature variation (Table 1 and Fig. 6A). This suggests that the detection of *qCT2* is only due to genotypes simultaneously carrying the minor alleles of two QTLs (e.g. *Haplotype III* in Table 1). The minor alleles for the SNPs defining *qCT1* are associated with a higher canopy temperature and occur in taller, low-yielding genotypes (Fig. 6 and Supplementary Table S5) whereas the major alleles across *qCT1* are almost completely fixed in the shorter (plant height < 100 cm) genotypes of this panel (Supplementary Table S6). This finding may support the hypothesis that rice genetic variation for stomatal conductance (here indirectly determined by canopy temperature measurements) was reduced as a result of selection for short-statured, high tillering and productive genotypes for flooded environments. The taller accessions of the population are mostly low tillering and low yielding landraces that were selected for drought-prone, rain-fed environments (Kumar *et al.*, 2014) where alleles responsible for reduced transpiration are preferred, even if it negatively affects grain yield under non-stressed conditions (Passioura, 2012). Fixation of the major alleles at *qCT1* in short-statured genotypes suggests that, in rice breeding, selection for high grain yield under flooded

conditions reduced the genetic variation available for traits related to more conservative water use.

The marker-trait association detected on chromosome 3 using the mean temperature values of stress replicates was also identified using the temperature values of replicate F03 alone (Fig. 4B) but not using F04 (Supplementary Fig. S10), the other drought replicate that shared the field location with F03 (Supplementary Fig. S1). This difference may be explained by the different severity levels of the imposed stress between the two days of imaging, in combination with other environmental variation. F04 was imaged at almost the same time of day, but during a time window characterised by higher mean solar radiation and wind speed (Supplementary Table S1B). Furthermore, F04 was imaged one day later than F03, such that the water limitation further increased (Supplementary Fig. S6), resulting in higher canopy temperature values than observed in F03 (Fig. 2). It is likely that, due to the increased severity of the stress in F04 (Fig. 3), the stomata closed in a larger number of genotypes, thus reducing the variability and sensitivity to detect genetic differences in stomatal closure that were still detectable the day before. This confirms the very dynamic response of stomata to changing environmental conditions (Vico *et al.*, 2011; Drake *et al.*, 2013) and suggests the need of reducing the time window during which the thermal imaging is performed to increase the number of replicates imaged per day. This target can be achieved by assembling thermal cameras on unmanned aerial vehicles (Shi *et al.*, 2016).

Candidate genes

A region of ~42 kbp was identified inside the *qCT1* LD block by considering the most significant SNPs associated with mean canopy temperature under drought (Fig. 7B). This region contains seven genes (Supplementary Table S7), of which two were targeted as interesting *a priori* candidate genes, a mitochondrial fumarase (LOC_Os03g21980) and a plasma membrane high affinity potassium (HAK) transporter (LOC_Os03g21950). Plant guard cells accumulate solutes like K^+ and malate during stomata opening and release/metabolize them during stomata closure. During these processes, solute flux through the plasma membrane of guard cells is highly active with K^+ intake driving stomata opening with the involvement of different types of transporters, including HAK-type transporters (Jezek and Blatt, 2017). Gago et al. (2016)

and Santelia and Lawson (2016) recently reviewed the role of guard cell and adjacent mesophyll cell metabolism in stomatal movement, highlighting the importance of malate (negatively charged) intake as a counter ion for K^+ during stomatal opening. Fumarase is a mitochondrial enzyme involved in the production of malate, through the hydration of fumarate, in a critical step of the tricarboxylic acid (TCA) cycle (Sweetlove *et al.*, 2010). In transgenic tomato plants, inhibition of fumarase resulted in a reduction in TCA cycle activity. This reduction had little effect on leaf metabolism but markedly reduced plant biomass because of a deficiency in stomatal function that resulted in reduced stomatal conductance (Nunes-Nesi *et al.*, 2007). The co-location of fumarase and of a HAK transporter, both important for the mechanism of stomatal opening, in the QTL region of highest interest associated with canopy temperature/stomatal conductance variation reinforces our mapping results.

Finally, our functional analysis of SNPs within the ~42 kbp region of highest significance within *qCT1* did not highlight variants responsible for amino acid changes (Supplementary Table S8) but identified many upstream gene variants suggesting that changes in gene regulation may explain the phenotypic variation associated with this region. This hypothesis should be further investigated by sequencing the candidate genes and their promoter regions in a subset of contrasting lines carrying the major and minor alleles at the significant markers within *qCT1,* and testing gene expression differences in response to stress. This would help to pinpoint the functional nucleotide polymorphisms and to assess how they impact stomatal conductance and response to water stress.

The significant SNPs across *qCT1* may also be of interest to breeding programs aiming to develop drought resistant varieties. These SNPs represent a useful target for either marker-assisted selection or genome editing using CRISPR *Cas9*. This should help to determine whether introducing minor alleles at *qCT1* into semi-dwarf and high yielding varieties equips them with new genetic potential for a more conservative water-use strategy under stress. This trait is currently not available in this germplasm. It will be of great interest to see whether improved water use efficiency in these high-yielding varieties can be accomplished without negatively impacting their productivity when water is abundant.

CONCLUSIONS

Physiological profiling of plant traits combined with genetic analysis has the potential to greatly accelerate crop improvement (Reynolds and Langridge, 2016). The present study shows that changes in stomatal conductance, an important physiological response to water-limitation, can be indirectly measured by thermal imaging and that the latter technique can be used to quantitatively screen a large panel of rice accessions. Canopy temperature during stress is a good predictor of grain yield performance and, therefore, thermal imaging represents an effective tool that can be used to accelerate physiological selection in plant breeding. In addition, association mapping of thermal data revealed the presence of genetic variation controlling canopy temperature under stress. The *a priori* candidate genes that were identified as underlying this genetic variation suggest that differences in the regulation of genes involved in guard cell solute intake affect stomatal behavior, which we detected as canopy temperature differences. Finally, our analysis shows that the major donors of genetic variation for canopy temperature/stomatal conductance are the tall landraces of rice. These old varieties and landraces present in crop germplasm collections represent a strategic reserve of genetic variation that can be tapped for developing new understanding of stress response and new varieties that are physiologically adapted to highly variable, water-limited environments.

ACKNOWLEDGEMENTS

This work is part of the 'Growing Rice like Wheat' research programme financially supported by an anonymous private donor, via Wageningen University Fund, for the first author's PhD fellowship.

SUPPLEMENTARY DATA

Supplementary data are available at:

https://drive.google.com/open?id=1t9Bn8RFIDTjwmXHs02OBtD1VAnY65Vts

Fig. S1. Field trial at IRRI during the dry season 2013-2014.

Fig. S2. Density of distribution of the standardized residuals for all the imaged field replicates.

Fig. S3. Principal component analysis plots for the 271 *indica* rice accessions.

Fig. S4. Mean temperature of thermal pictures relative to E03 drought field replicate.

Fig. S5. Correlation matrices of non-normalised and normalised field replicates.

Fig. S6. Soil water potential of the drought field.

Fig. S7. Boxplots representing the range of variation for the recorded agronomic traits.

Fig. S8. Scatterplots between drought stress mean temperature values and plant height, grain yield and harvest index scored during the 2013 and 2014 dry seasons.

Fig. S9. Quantile-Quantile plots of expected versus observed *p*-values for the GWA mapping results.

Fig. S10. Manhattan plots of the GWA mapping results of control replicate B mean, B03 and B04 and of the separate drought replicates E02, E03 and F04.

Fig.S11. Manhattan plots displaying the level of significance over genomic positions of the chromosome 3 region of the two markers' loci significantly associated with canopy temperature of 'EF mean'.

Fig. S12. Manhattan plots of the GWA mapping results for the QTL region of chromosome 3 using the imputed map for 'EF mean' canopy temperature.

Table S1A. Detailed information on imaged field replicates.

Table S1B. Weather station data of the days of imaging.

Table S2. Pseudo-heritability (h^2) of temperature results for all the field replicates.

Table S3. Significant SNPs identified by GWA mapping of canopy temperature under drought stress using the GBS SNP map.

Table S4. Allelic effect on canopy temperature for markers of the most significant SNPs, located on chromosome 3, in the different drought field replicates and for their mean values.

Table S5. Significant SNPs identified by GWA mapping of 'EF mean' canopy temperature using the imputed map.

Table S6. List of the 246 accessions carrying one of the three haplotype groups, identified considering the most significant SNPs of qCT1 and qCT2, and their phenotypic performance.

Table S7. Genes included in the localized region delimited by the most significantly associated SNPs with canopy temperature of 'EF mean' and located inside *qCT1*.

Table S8. Effect of the 20 SNPs present in the *qCT1* region and significantly associated with 'EF mean' canopy temperature.

CHAPTER

5

Genome wide association mapping reveals the genetic basis of metabolic and enzymatic biomarkers for rice grain yield stability under drought

Giovanni Melandri, Susan R. McCouch, Harro J. Bouwmeester

In preparation for submission

ABSTRACT

The elucidation of the genetic basis of biochemical components predictive of grain yield represents an alternative breeding strategy for increasing crop yield and contributing to global food security. In this study, flag leaf central metabolism and oxidative stress status of 271 *indica* rice (*Oryza sativa*) accessions, grown in the field under well-watered and reproductive stage drought conditions, were used to predict grain yield performance using a multivariate statistical model. The resulting models for grain yield under well-watered conditions and drought displayed a higher predictability than multivariate models using genome data as explanatory variables, especially for the prediction of grain yield under drought and for stress-induced grain yield loss. The best predictive variables of the metabolome/oxidative stress-based models represent metabolic and enzymatic biomarkers that can be used in breeding for grain yield performance in rice. For these biomarkers, the fraction of their genetic basis associated with grain yield differences among accessions was determined by genome-wide association mapping. In this way, we identified genomic regions, and underlying candidate genes, that potentially represent breeding targets to improve rice grain yield under optimal conditions and grain yield/yield stability under drought stress.

KEYWORDS

Oryza sativa, metabolome, oxidative stress, PLSR, grain yield, drought, GWA mapping.

INTRODUCTION

A deeper understanding of the genetic basis of yield in response to environmental stress is key to improving the sustainability and productivity of major agricultural crops. In rice, as in most crops, grain yield is a complex trait with low heritability. It is controlled by many genes of small-effect, and the gene networks that ultimately control grain yield are influenced by genotype-environment interactions, as well as pleiotropic and epistatic effects (Xing and Zhang, 2010). These characteristics make it difficult to identify "yield genes" using genome-wide association (GWA) mapping, because the multiple layers of interaction among variables are not easily incorporated into the linear-mixed models used in GWA mapping (Liu and Yan, 2019). An alternative strategy for identifying "yield loci" would be to dissect yield *per se* into smaller, component traits that could be measured with greater accuracy and precision, and to identify genes and molecular variants associated with rate-limiting yield-component traits as the basis for improving crop yields in applied breeding programs (Reynolds and Langridge, 2016).

Metabolites are related to the biochemical and physiological status of the plant and may be considered intermediate phenotypes. They can also be yield-component traits and therefore, represent a promising alternative target (Luo, 2015). Metabolites are end products of cellular regulatory processes and they inherently incorporate the effect of genetic (i.e. pleiotropy and epistasis) and environmental factors, as well as their interactions (Fiehn, 2002; Herrmann and Schauer, 2013). In many ways, metabolite levels are more closely linked to phenotype than are gene transcripts, and in recent years, an expanding number of studies in diverse plant species have focused on large-scale metabolite profiling of bi-parental populations and association mapping panels (Carreno-Quintero *et al.*, 2013; Luo, 2015; Fernie and Tohge, 2017), including rice (Gong *et al.*, 2013; Chen *et al.*, 2014; Dong *et al.*, 2015; Matsuda *et al.*, 2015). In most of these studies, metabolite levels were considered to be quantitatively inherited phenotypes and their genetic control was successfully explored through QTL or GWA mapping, but their association with traits of agronomic interest was not demonstrated. In different studies, the plant metabolome was used to predict complex traits such as biomass in *Arabidopsis* and maize (Meyer *et al.*, 2007; Sulpice *et al.*, 2009, 2013; Steinfath *et al.*, 2010; Riedelsheimer *et al.*, 2012) or grain yield in rice (Matsuda *et al.*, 2015). These studies

highlighted the value of metabolic traits as either functional intermediates or correlated biomarkers for predicting yield-related traits, leading to the idea that metabolomics could complement genomics- and genetics-based breeding for crop yield improvement (Herrmann and Schauer, 2013; Valluru *et al.*, 2014; Luo, 2015; Kumar *et al.*, 2017).

Recently, Melandri *et al.* (2019) showed that a multivariate model based on flag leaf central metabolites and oxidative stress-related markers/enzymes accurately predicted drought-induced grain yield loss variation in a population of 292 landrace and modern *indica* rice varieties. Here we investigate whether the multivariate model – successfully used to predict grain yield loss under drought – could also accurately predict grain yield *per se* under well-watered and under drought stress conditions. Using 271 (out of 292) accessions genotyped with 81,347 SNP markers, we also determine the predictive power of the partial least squares regression (PLSR) model for the same traits. The best predictors of each model were subsequently analyzed as quantitative traits in a GWA study using the 81,347 SNPs. The resulting associations allowed us to identify genetic markers that can be economically used in breeding to improve rice grain yield under optimal conditions and/or grain yield/yield stability under drought stress.

MATERIALS AND METHODS

Genetic resources and plant growth

The two-hundred seventy-one accessions of *Oryza sativa* subsp. *indica* (Supplementary Table S1) were part of a larger panel (~300) used in a field experiment at the International Rice Research Institute (IRRI), Los Baños, Philippines during the 2013 dry season. The accessions are largely those of the PRAY-*indica* panel (http://ricephenonetwork.irri.org) which includes traditional and improved *indica* rice varieties originating from rice-growing countries in tropical and sub-tropical regions around the world. The same panel was recently used in studies where a number of diverse traits were phenotyped as the basis for GWA mapping (Qiu *et al., 2015;* Al-Tamimi *et al.*, 2016; Rebolledo *et al.*, 2016; Kadam *et al.*, 2017, 2018; Kikuchi *et al.*, 2017). The experiment comprised a control field and a drought stress field, each with three replicates of the population arranged in a serpentine design. To synchronize flowering, the accessions were divided into six groups according to days required to flower (previously

collected data), and progressively sown and transplanted with intervals of 10 days between each group. Drought stress consisted of 14 consecutive days of water withholding applied only to the stress field at the reproductive stage (targeting 50% flowering). At the end of stress, the field was re-watered until all the accessions reached maturity for harvest. The same experiment was conducted again during the 2014 dry season. Further details on the field experiments conducted in the two years can be found in Kadam *et al.* (2018).

Phenotyping

Grain yield (grams/m^2) under control (GY-con) and drought (GY-dro) conditions was calculated considering only the replicates (two for GY-con and three for GY-dro) sampled for metabolomics and oxidative stress status analyses. Percentage of grain yield loss (GY-loss) of each accession was calculated as 100*(GY-con – GY-dro)/(GY-con). Plant height under control (PH-con) and drought (PH-dro) was calculated as for GY-con and GY-dro and expressed in cm. Flowering time differences (expressed in days) among the accessions under control (FT-con) and drought (FT-dro) were calculated as the date of leaf sampling, minus the date of 50% flowering considering the same field replicates as for GY and PH. For the 2014 field trial, GY-con, GY-dro and GY-loss were calculated considering all three field replicates. The values of these traits for all accessions for the 2013 and 2014 field experiments are shown in Supplementary Table S1.

Leaf sampling, metabolite profiling and determination of oxidative stress-related marker values and enzyme activities

Flag/top leaves of each accession were sampled and immediately frozen in liquid nitrogen as described in Melandri *et al.* (2019). Leaf samples of the 271 accessions were analyzed by untargeted GC-MS-based metabolite profiling to assess the variation in polar metabolites as described by Riewe *et al.* (2012) and Riewe *et al.* (2016). For each accession and treatment, equal amounts of replicates (two for GY-con and three for GY-dro) were pooled prior to GC-MS analysis and metabolite identification. A total of 88 metabolites were identified, predominantly primary metabolites (amino acids, sugars and organic acids). Glucose, fructose and sucrose were quantified spectrophotometrically (Riewe *et al.*, 2008). Further details on metabolite profiling can be found in Melandri *et al.* (2019). The same leaf materials were

analyzed for the oxidative stress status. For this, the level of molecular antioxidants (2), oxidative stress markers (2) and the activity of enzymes (16) involved in ROS scavenging mechanisms and photorespiration were quantified using high-throughput colorimetric assays (Melandri *et al.*, 2019). The 111 metabolites and oxidative stress markers/enzymes considered in this study, and their variation among the accessions, are shown in Supplementary Table S2.

Data analysis and PLSR models

Results of all metabolites and oxidative stress markers/enzymes were $\log_{10}$-transformed to improve normality. Statistical analyses and graphical representations were performed using R (version 3.4.3; The R Foundation for Statistical Computing). Imputation of missing values, prior statistical analyses, was performed by the *knnImputation* function in the 'DMwR' R package. Correlation analysis and graphical matrices were produced using a modified function of the 'corrplot' R package. SNP-based principal component analysis (PCA) was performed using the *prcomp* function in the 'stats' R package. To predict GY-con, GY-dro and GY-loss, cross-validated partial least squares regression (PLSR, 'pls' R package) models were used (Mevik and Wehrens, 2007; Mumm *et al.*, 2016). Observations were auto-scaled in the PLSR procedure. The number of latent variables to include in each model was selected by testing the predictability value (Q^2) using an increasing number of latent variables from 1 to 10. Considering the PLSR model based on metabolites and oxidative stress markers/enzymes, the relative importance of the model variables was summarized by multiplying the ten ranks of each variable in the single sub-models generated by the cross-validating procedure (rank-products).

Genome-wide association (GWA) mapping

Genome-wide association (GWA) mapping was performed using a linear-mixed model in EMMAX (Kang *et al.*, 2010), which corrects for population structure by including a kinship matrix (IBS matrix) as covariate. EMMAX also provides an estimate of the phenotypic variance (pseudo-heritability, h^2) explained by the IBS matrix. The 271 accessions of this study are a subgroup of the original panel (PRAY-*indica* panel, n=339 accessions) that was used to generate a 88,753 SNP map using Genotyping-By-Sequencing (GBS) (23.26% missing data imputed by Fast Phase Hidden Markov Model, Scheet and Stephens 2006). The reduced

number of accessions (271) altered the minor allele frequency (MAF) threshold of the 339 accessions panel, originally set at 0.05. To exclude rare alleles from the present study (n=271 accessions), the 88,753 SNP map was re-filtered for MAF > 0.05, resulting in 81,347 SNPs available for GWA mapping. PCA based on the 81,347 SNPs was conducted to quantify subpopulation structure. The main component (PC1) explained only 8.30% of the genetic variation but it showed association with the geographical area of origin of the accessions (Supplementary Fig. S1) and, therefore, it was included as covariate in all the GWA mapping runs. GWA mapping results are presented as Manhattan and Quantile-Quantile plots using the 'qqman' R package. To avoid Type 1 error, a suggestive significance threshold of $p < 0.0001$ (i.e. $-\log_{10} p > 4.0$) was used to identify significant marker-trait associations. This significance threshold was commonly used in other GWA mapping studies on rice based on SNP maps of similar density (Zhao *et al.*, 2011*b*; Dimkpa *et al.*, 2016; Rebolledo *et al.*, 2016; Kadam *et al.*, 2017, 2018). Quantitative trait loci (QTLs) were defined by the presence of a SNP with significance of $-\log_{10} p > 4.0$ and at least two additional SNPs showing significance of -log10 $p \geq 3.0$ and distance $\leq$ 150 kbp from the most significant marker. The specific SNP distance interval represents the double of the ~75 kbp linkage disequilibrium (LD) decay estimated in *indica* rice by Mather *et al.* (2007). For each significant QTL, the phenotypic effect of the minor allele of the most significant (top) SNP marker was calculated for the traits used in GWA mapping and for the other agronomic traits. Welch's t-tests were conducted to evaluate the significance of the phenotypic differences between the groups of accessions carrying the minor and major allele of each top marker. Furthermore, considering that the genome-wide LD decay of this specific rice panel was estimated to be ~65 kbp (Kadam *et al.*, 2017), we applied this interval before and after the top markers of each QTL to identify the specific LD blocks.

RESULTS and DISCUSSION

Genotypic variation in phenotypic traits and their correlations

In this study, the grain yield (GY) performance of 271 traditional and improved *indica* rice varieties (Supplementary Table S1), assembled from major tropical and subtropical rice-growing regions around the world, was assessed in a field experiment under irrigated (control)

and reproductive-stage drought conditions (Kadam *et al.*, 2017). Drought stress significantly (Paired t-test: $P \leq 0.001$) reduced GY by an average of 29.4% (GY-loss) (Supplementary Table S3).

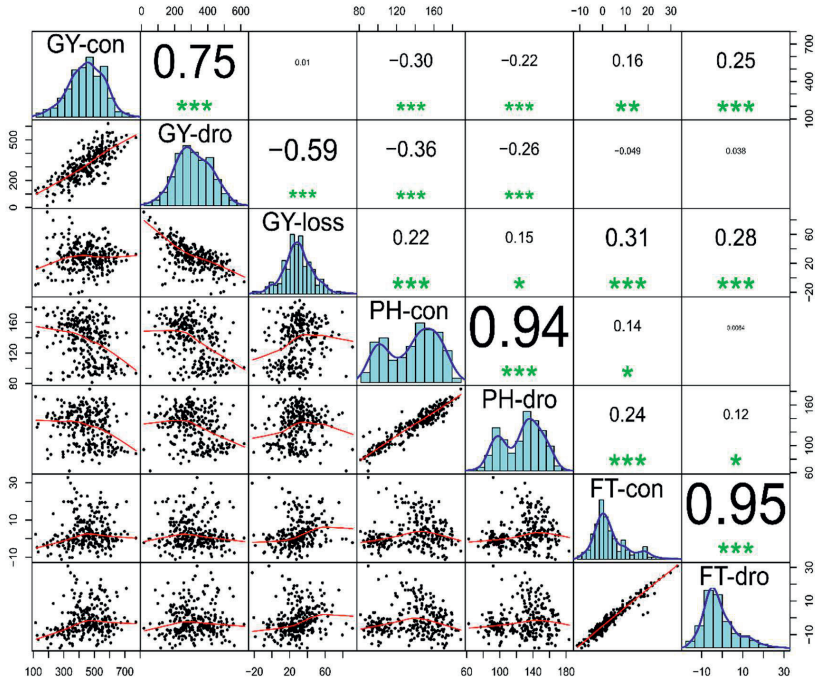


Fig. 1. Trait correlations and distributions.
This figure summarizes the correlation matrix between grain yield (GY), grain yield loss (GY-loss), plant height (PH) and flowering time difference (FT) under control (-con) and drought (-dro) conditions. GY units are expressed in grams/m^2, GY-loss in percentage, PH in centimetres and FT in days between flowering and sampling date. In the upper-right portion of the matrix are reported correlation values determined by Spearman's correlation (stronger correlations are represented by larger numbers) and significance levels (in green, '***'= $P \leq 0.001$, '**'= $P \leq 0.01$, '*'= $P \leq 0.05$). In the bottom-left portion of the matrix are shown scatterplots of the pairwise combinations between traits (trendline in red). Trait distributions are represented along the diagonal of the matrix (trendline in blue).

GY under control (GY-con) and drought (GY-dro) conditions were highly correlated ($r_s = 0.75$, $P \leq 0.001$), suggesting strong genotypic control of the trait under both treatments. However, GY-loss was significantly ($P \leq 0.001$) and negatively correlated ($r_s = -0.59$) with GY-dro only (Fig. 1). This observation suggests that the GY-dro performance is also largely influenced by genotype-by-treatment interactions. Considering that this rice panel includes traditional landraces and improved modern varieties, and that plant height (PH) and flowering time (FT)

have been significantly impacted by modern breeding efforts, we assessed the relationships between the GY-related traits and these two important agronomic traits, which both displayed high genotypic diversity (Fig. 1 and Supplementary Table S1). We also synchronized FT by sowing and transplanting the accessions on different dates such that drought stress was imposed on all genotypes at the same stage of development (Kadam *et al.*, 2017). Flowering was indeed largely synchronized, but not entirely. Melandri *et al.* (2019) showed a prevalent influence of genotype on FT under both control and drought, and that drought stress results in an almost identical delay in flowering for all accessions. A similar genotype-induced correlation (r_s = 0.95, $P \leq 0.001$) on the delay of flowering under drought (FT-dro) versus control (FT-con) was confirmed in the 271 genotypes (Fig. 1 and Supplementary Table S3). FT-dro was positively correlated with GY-loss (r_s = 0.28, $P \leq 0.001$; Fig. 1) confirming that, in rice, drought-induced yield loss is influenced by FT differences during stress imposition (Liu *et al.*, 2006), though this effect is not very large. Specifically, the correlation shows that accessions that flowered before the onset of stress (< 10% of the total) had relatively greater GY-loss than those that flowered during stress (booting stage ~60% and heading stage ~30% of the total) (Kadam *et al.*, 2017). Intriguingly, however, FT-dro does not correlate with GY-dro (Fig. 1), showing that, unlike for GY-loss, the genotype effect on GY-dro totally masked the influence of FT differences. Similar to FT, genotypic differences almost exclusively determined PH under both control (PH-con) and drought (PH-dro) conditions, consistent with the high correlation (r_s = 0.94, $P \leq 0.001$) observed in the two environments, with the drought treatment causing a significant ($P \leq 0.001$) mean reduction of 8.1 cm (Supplementary Table S3). Interestingly and different from GY and FT, the PH distribution under both treatments is bi-modal, with two distinct normal distributions around two different peaks (Fig. 1). This bi-modal distribution is caused by the composition of the rice panel which includes tall, pre-green revolution traditional varieties and shorter post-green revolution modern varieties. In rice *indica* germplasm, the shorter stature of modern varieties (semi-dwarf phenotype) is mainly determined by the introduction of a recessive allele (*sd1*) of the gibberellin 20-oxidase biosynthetic gene (*OsGA20ox2*, LOC_Os01g66100), while the functional wild-type allele (*SD1*) is present in the taller, traditional varieties (Monna *et al.*, 2002; Spielmeyer *et al.*, 2002; Sasaki *et al.*, 2002). Confirming our hypothesis, Kadam *et al.* (2017) identified a QTL, and the underlying *OsGA20ox2* gene, associated with PH differences in this panel. The semi-dwarf

stature in modern, high-yielding varieties is associated with higher harvest index, with more assimilates partitioned to the grains than to the leaves and stems. This translates into a higher yield for semi-dwarf varieties compared to the taller landraces and old varieties (Hedden, 2003). This stature-associated yield difference is present in our rice panel and evidenced by the significant ($P \leq 0.001$) negative correlations between PH and GY under both control (r_s = -0.30) and drought (r_s = -0.26). PH differences are also correlated with drought-induced yield losses (GY-loss and PH-dro; r_s = 0.15, $P \leq 0.05$; Fig. 1) although the low level of significance and strength of correlation indicate that the stress treatment reduced the influence of the height-yield relationship (Fig. 1).

Metabolome/oxidative stress-based multivariate models display higher prediction power for grain yield variation than genomic-based models

Multivariate, metabolome-based models of complex plant traits (i.e. biomass or yield) display higher predictive power than univariate statistical methods (like pairwise correlation analysis), and simultaneously allow identification of the most important predictive metabolites for use as biomarkers in breeding (Meyer *et al.*, 2007; Sulpice *et al.*, 2009, 2013; Steinfath *et al.*, 2010; Fernandez *et al.*, 2016). In rice, these models were successfully employed to predict the yield of hybrids by directly using the hybrid's metabolite profiles (Xu *et al.*, 2016) or those of the parents (Dan *et al.*, 2016). Despite the value of these findings for hybrid breeding programs, the narrow genetic background of the materials used in these studies did not explore the large, qualitative and quantitative genetic diversity available for rice metabolism (Chen *et al.*, 2014; Fernie and Tohge, 2017). In addition, most of the metabolomics studies in crop species, including rice, have been conducted under control conditions while, in *Arabidopsis,* natural variation in metabolic plasticity was shown to be an important factor contributing to phenotypic plasticity (Kleessen *et al.*, 2014). Recently, Melandri *et al.* (2019) showed that a cross-validated partial least squares regression (PLSR) model, based on the levels of 111 flag leaf metabolites and oxidative stress markers/enzymes (hereafter MetabOxi), efficiently predicted stress-induced GY-loss in 292 genetically diverse *indica* rice genotypes.

Here, we expand the PLSR modelling approach to predict GY-con and GY-dro, in addition to, GY-loss in a subset of 271 accessions from the same experiment using the MetabOxi dataset.

The 271 accessions were genotyped with 81,347 SNP markers, and the genomic dataset was also used – instead of the MetabOxi data – to build PLSR models for prediction of the above-mentioned traits. Fig. 2 shows the observed versus predicted values and predictability (Q^2) of the best (based on increasing latent variables, Supplementary Table S4) cross-validated MetabOxi- and genomic-based PLSR models for the three traits.

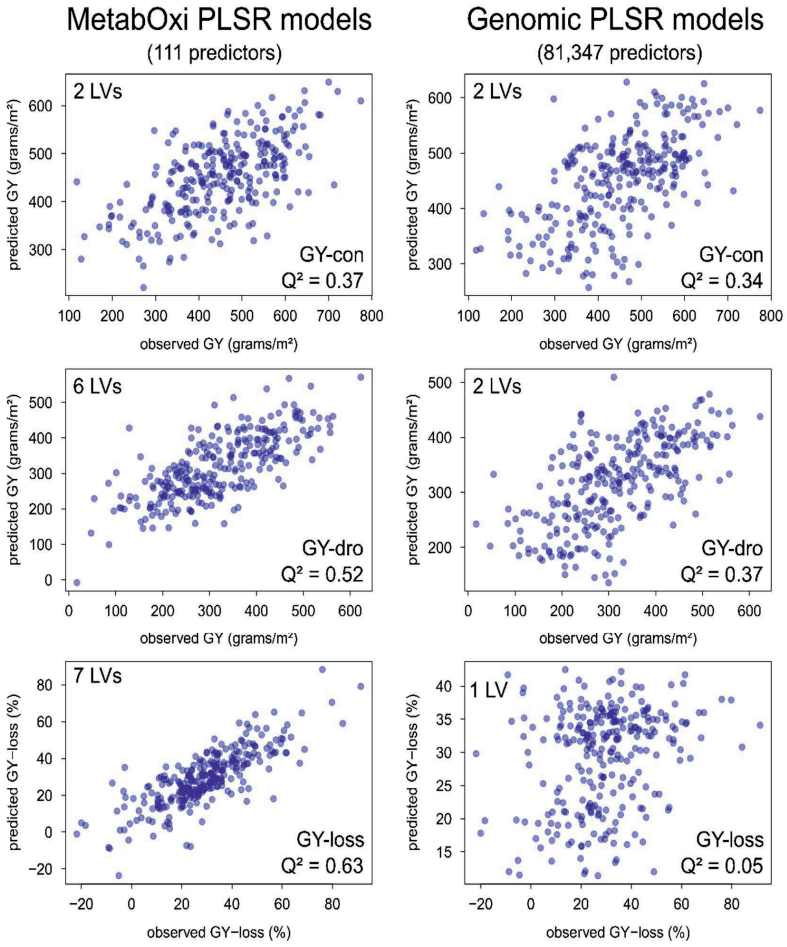


Fig. 2. Predicted versus observed values of the MetabOxi-and genomic- based PLSR models for prediction of grain yield and grain yield loss under control and drought conditions.
PLSR plot of the cross-validated models for GY-con (top), GY-dro (middle) and GY-loss (bottom) based on values of the 111 MetabOxi variables (left) and on 81,347 SNPs (right). Predictability (Q^2) and linear latent variables (LVs) of the model are displayed in each plot. Each MetabOxi-based model shows the mean results of the model replicated five times.

Trait predictability is always higher for the MetabOxi- than for the genomics-based models. Particularly for GY-dro and GY-loss (Q^2 = 0.52 and 0.63), large differences were observed between the MetabOxi-based models (both built using the values of the MetabOxi variables under drought; Supplementary Table S2) and their respective genomic-based models (Q^2 = 0.37 and 0.05). Overall, the higher predictive power of the MetabOxi dataset compared to the genomic dataset shows that metabolite values and enzyme activities are more closely aligned to the phenotype than are single genetic determinants (Herrmann and Schauer, 2013; Xu *et al.*, 2016). This closer alignment can be explained by the fact that metabolites/enzyme activities represent complex biological processes whose measured value or activity incorporates the effect of a dynamic network comprised of multiple layers of regulation (DNA, RNA, protein) in response to dynamic internal and external stimuli (Sulpice and McKeown, 2015). This is especially true in the context of environmental stress-driven perturbations, including those caused by drought. Indeed, the ability of plants to dynamically respond to environmental change (phenotypic plasticity) is often achieved by adjusting metabolite levels and enzyme activities through rapid responses to external signals from the environment mediated via post-translational and/or transcriptional regulation (Stitt *et al.*, 2010). In support of this hypothesis, we observed that the MetabOxi-based model was the best in predicting GY-loss, while the genomic-based model was the least predictive (almost null) (Fig. 2). Under control conditions, the predictive power of the MetabOxi- (based on control values of the variables, Supplementary Table S2) and genomic-based prediction models for GY-con were virtually the same (Q^2 = 0.37 and 0.34, respectively). This observation supports our interpretation that the reason for the improved predictive accuracy of the MetabOxi-based model under drought is due to the fact that the variables in the model (basal flag leaf levels of central metabolites and oxidative stress markers/enzymes) integrate the effect of post-translational and transcriptional regulatory processes that differentiate the ability of plant genotypes to respond to stress, while these stress-responsive physiological and biochemical processes are not significant determinants of plant performance under control conditions.

Best model predictors and their relationships with grain yield, flowering and plant height

Each MetabOxi-based PLSR model provided a ranking of importance for the predictive variables. Table 1 shows the top three variables for the GY-con, GY-dro and GY-loss models,

together with their rank-product values. The top three MetabOxi variables for each model, those with the lowest rank-products (lower rank-product implies higher importance) (Supplementary Table S5) are potentially useful biomarkers for GY. We next performed a correlation analysis. Table 1 summarizes the correlations of the top-ranked variables with the predicted GY trait, as well as with FT and PH (correlations between all the MetabOxi variables and traits are reported in Supplementary Table S6 and S7). The top-ranked MetabOxi variable for GY-con (chlorogenic acid) was significantly correlated with both FT and PH, while the second- and third-ranked variables for GY-con (isocitric acid and citric acid) were correlated with PH, but not with FT. None of the top-ranked variables for GY-dro or GY-loss were significantly correlated with either PH or FT. Table 1 provides additional information about the nature (positive or negative) and strength (r_s) of the associations.

Among the 111 MetabOxi variables evaluated as predictors of grain yield under control conditions (Supplementary Table S6), the top-ranked variables were all organic acids (Table 1). Among them, chlorogenic acid (3-caffeoyl-quinic acid) is by far the most important predictor (rank-prod.=1). Its negative correlation with GY-con (r_s= -0.40) is surprising, considering that chlorogenic acid is noted in the literature for its beneficial antioxidant and anti-herbivore activity in plants (Takahama and Oniki, 1997; Niggeweg *et al.*, 2004; Ferreres *et al.*, 2011; Kundu and Vadassery, 2019). A possible explanation lies in the fact that chlorogenic acid is positively correlated with both PH-con and, to a lesser extent, with FT-con (Table 1), and that GY-con is strongly and negatively associated with PH-con (Fig. 1). In general, tall, late-flowering, low-yielding traditional varieties/landraces of rice are characterized by a higher degree of environmental robustness, derived from their adaptation to environmentally variable low-input production systems (Lempe *et al.*, 2013; Dwivedi *et al.*, 2016), in comparison to shorter, early-flowering, higher-yielding modern varieties that have been bred for relatively stable, high-input systems. It might be that in this field trial, under irrigated conditions and with the application of fertilizers and weed, insect and disease control, a constitutively higher activity of the chlorogenic acid pathway in the traditional, tall accessions produced no advantage in terms of GY and, on the contrary, was associated with lower GY performance.

Table 1. Best predictive variables of the MetabOxi-based PLSR models for grain yield under control (GY-con) and drought (GY-dro) conditions and for grain yield loss (GY-loss).

Top three ranked predictive variables of the cross-validated MetabOxi-based PLSR models for prediction of grain yield under control (GY-con) and drought (GY-dro) conditions and for grain yield loss (GY-loss). Variables are ranked based on their rank-product value (Rank-prod.). Correlations between the variables and GY-related traits, flowering (FT) and plant height (PH) under control (-con) and drought (-dro) are reported. r_s: Spearman's rho of correlation. Correlation significance (sign.): '***'= $P < 0.001$, n.s.= not significant. R^2: variance explained by the linear model created between the variable and each trait.

Model	Rank	Variable	Rank-prod.	GY-con			FT-con			PH-con		
				r_s	sign.	R^2	r_s	sign.	R^2	r_s	sign.	R^2
GY-con	1	Chlorogenic acid	1	-0.40	***	0.13	0.33	***	0.12	0.39	***	0.14
	2	Isocitric acid	1,024	0.39	***	0.15	0.13	n.s.	0.04	-0.42	***	0.15
	3	Citric acid	527,018	0.35	***	0.13	0.04	n.s.	0.01	-0.46	***	0.20
Model	**Rank**	**Variable**	**Rank-prod.**	**GY-dro**			**FT-dro**			**PH-dro**		
				r_s	sign.	R^2	r_s	sign.	R^2	r_s	sign.	R^2
GY-dro	1	DHAR	1	0.59	***	0.34	-0.06	n.s.	0.00	-0.17	n.s.	0.01
	2	MDA	1,126	-0.40	***	0.17	0.22	n.s.	0.03	0.16	n.s.	0.02
	3	MDHAR	186,332	0.28	n.s.	0.04	0.18	n.s.	0.03	-0.10	n.s.	0.01
Model	**Rank**	**Variable**	**Rank-prod.**	**GY-loss**			**FT-dro**			**PH-dro**		
				r_s	sign.	R^2	r_s	sign.	R^2	r_s	sign.	R^2
GY-loss	1	DHAR	1	-0.57	***	0.38	-0.06	n.s.	0.00	-0.17	n.s.	0.01
	2	MDA	1,024	0.62	***	0.37	0.22	n.s.	0.03	0.16	n.s.	0.02
	3	MDHAR	59,049	-0.04	n.s.	0.00	0.18	n.s.	0.03	-0.10	n.s.	0.01

As noted above, chlorogenic acid also shows a significant, positive correlation (r_s= 0.33) with FT-con (Table 1), but the overall influence of flowering time on GY-con is low (Fig. 1), and thus we regard this association as having limited importance in understanding the selection of chlorogenic acid as one of the top predictors in the model. Indeed, the control values of many other MetabOxi variables display stronger correlations with FT-con than chlorogenic acid (Supplementary Table S6).

The second and third highest ranking predictive variables of the GY-con model, isocitric and citric acid (rank-prod.=1,024 and 527,018 respectively), correspond to the two first intermediates of the tricarboxylic acid (TCA) cycle, a fundamental pathway that provides energy and carbon skeletons for many plant biosynthetic processes (Sweetlove *et al.*, 2010; Araujo *et al.*, 2012). The fact that the two organic acids are positively correlated with GY-con (r_s= 0.39 and 0.35, respectively, Table 1) suggests that their high rank-values represent a signature of sustained metabolic activity in accordance with higher GY performance. In contrast to the positive correlation with GY-con, isocitric and citric acid are strongly and negatively associated with PH-con (r_s= -0.42 and -0.46, respectively). These negative correlations suggest the presence of higher TCA/biosynthetic activity in the short, high-yielding varieties of the panel compared to the tall, lower-yielding traditional accessions. This suggests that the ability to translate increased radiation- and nitrogen-use efficiency into higher yields in modern rice varieties (Zhu *et al.*, 2016) is related to metabolic adaptations of central metabolism. Overall, the top predictive MetabOxi variables of the GY-con model are also individually strongly correlated with grain yield under control conditions, but their respective linear models explain only a low percentage of the trait variance (R^2 = 0.13~0.15) (Table 1). In addition, these variables are strongly associated with FT-con (chlorogenic acid) and PH-con (all three), two traits with a confounding effect on yield (Fig. 1). For these reasons, the use of chlorogenic acid, isocitric and citric acid as biomarkers for GY performance in rice under well-watered conditions should be treated with caution.

In contrast to the GY-con model, the top three predictive variables of the GY-dro and GY-loss models are not significantly associated with variation in either FT or PH (Table 1) and, therefore, likely represent more reliable trait biomarkers. Interestingly, the GY-dro and GY-loss models display the same three top predictors, in the same rank order and with similar

ranking-products (Table 1). Dehydroascorbate reductase (DHAR) is the most important predictor (rank-prod.=1 in both models), and is on its own also positively correlated with GY-dro (r_s= 0.59) and negatively with GY-loss (r_s= -0.57). The correlations between DHAR and GY-dro/GY-loss are the most significant among the 111 MetabOxi variables (Supplementary Table S7). The third highest ranking predictor, monodehydroascorbate reductase (MDHAR), shows a marked reduction of importance compared with DHAR (rank-prod.=186,332 and 59,049 in the GY-dro and GY-loss models, respectively), and is not significantly correlated with GY-dro or GY-loss. However, DHAR and MDHAR are both involved in the ascorbate-glutathione cycle, the central redox-hub in plants, where they regenerate oxidized ascorbate into its reduced form which, in turn, can be utilized for ROS scavenging (Foyer and Shigeoka, 2010; Foyer and Noctor, 2011). The presence of these two enzymes as top predictors of the GY-dro and GY-loss models highlights the importance of antioxidant defenses, primarily of the ascorbate-glutathione cycle, in preventing drought-induced oxidative damage at the flowering stage that negatively impacts rice GY (Melandri *et al.*, 2019). The non-significant correlation between MDHAR and GY-dro/GY-loss demonstrates that the multivariate modelling approach does not select variables only based on their simple linear association with the trait being predicted. Nevertheless, because of the low correlation of MDHAR with the GY traits, only DHAR can be regarded as an effective trait biomarker, especially considering that alone, it already explains a high percentage of the trait variance (R^2 = 0.34 and 0.38 for GY-dro and GY-loss, respectively).

The second top ranked variable of the two models (Table 1), malondialdehyde (MDA), is a lipid peroxidation product indicative of oxidative damage to the cellular lipid membranes (Møller *et al.*, 2007). MDA is relatively important for the models (rank-prod.= 1,126 and 1,024 in the GY-dro and GY-loss models, respectively) and shows a negative correlation with GY-dro (r_s= -0.40, R^2= 0.17) and an even stronger positive correlation with GY-loss (r_s= 0.62, R^2= 0.37). These strong correlations support the role of MDA as reliable biomarker of GY in rice under drought conditions, especially for GY-loss (Melandri *et al.*, 2019). The presence of DHAR, MDA and MDHAR as top-ranked predicting variables of the GY-dro and GY-loss models indicates that, during drought imposition, the flag leaf oxidative stress status of the accessions is more predictive of GY performance than their flag leaf central metabolism. This insight is supported

by the fact that the model for predicting GY-loss selected two additional oxidative stress-related variables, ascorbate oxidase (AO) and total antioxidant capacity (TAC), as fourth and fifth highest ranked variables (Supplementary Table S5). AO and TAC, similar to the other oxidative stress-related variables, are not significantly correlated with FT-dro and PH-dro (Supplementary Table S7) (Fig. 1). In contrast, the fourth and fifth highest ranked predictors for GY-dro, a trait highly correlated with GY-con (Fig. 1), were α-ketoglutaric acid and isocitric acid. Similar to the predictors selected by the GY-con model, these two predictors were both highly and negatively correlated with PH-dro (r_S= -0.37 and -0.43, respectively) (Supplementary Table S7), an observation that indirectly confirms the genotypic signature on GY-dro from a metabolic perspective.

GWA mapping of metabolic and enzymatic grain yield predictors

Metabolic and enzymatic biomarkers originate from and are fully exploited by researchers in the medical field, especially as diagnostic and predictive markers for diseases, but their application in plant breeding remains very limited (Herrmann and Schauer, 2013; Fernandez *et al.*, 2016). In this study we have identified a set of metabolic and enzymatic predictors for grain yield in a large panel of genetically diverse *indica* rice accessions, and found that these metabolic and enzymatic biomarkers are particularly valuable for predicting GY under drought. We next investigated the possibility of identifying genomic regions (QTLs) associated with quantitative variation in levels/activities of these predictors as the basis for converting them into high-throughput, low cost DNA markers (SNPs), greatly facilitating their application in breeding (Valluru *et al.*, 2014).

To achieve this target, we conducted GWA mapping using levels of the top predictors of the GY PLSR models (Table 1) as phenotypes, and the 81,347 SNP markers as genotypes for the 271 accessions of the panel. We used the first SNP-based principal component (PC1) as a covariate in the linear-mixed model to control for variation associated with geographical origin of the accessions (Supplementary Fig. S1), and because of the correlations between model predictors, and PH and FT (particularly in the GY-con model) (Table 1), we also included these traits as covariates in the model. GWA mapping was performed using the values/activities of each predictor estimated under control and independently under drought conditions to

determine whether the detected QTLs were identified in one or both of the environments (Wu *et al.*, 2018). As summarized in Fig. 3, the phenotypic distributions of predictor values/activities were approximately normal in both environments and therefore suitable for genetic analysis.

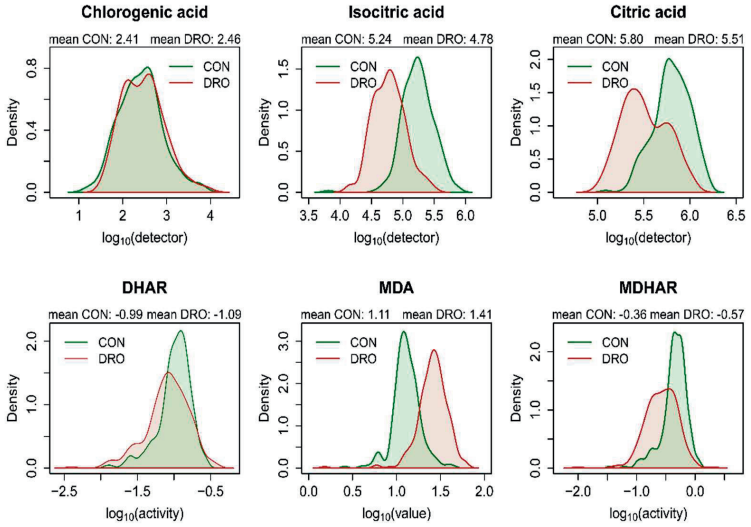


Fig. 3. Phenotypic distribution of control and drought stress values/activities of the best predictors of the PLSR models.

Phenotypic distributions of log_{10}-tranformed values/activities for the best GY-con predictors (chlorogenic acid, isocitric acid, citric acid) and GY-dro/GY-loss predictors (DHAR, MDA, MDHAR) in the 271 accessions of the panel. Each plot displays the distribution of predictor values/activities under control (green) and drought (brown) conditions, together with the condition-specific mean values (mean CON and mean DRO). Drought stress marginally affected the flag leaf levels of chlorogenic acid, which displayed similar mean values in both conditions. In contrast, drought reduced the mean levels of isocitric and citric acids (more marked) and the mean activity of DHAR and MDHAR (less marked) while MDA displayed a stress-induced increase of the population mean value. Additionally, the shape of the distribution curves of chlorogenic acid, isocitric acid and MDA was similar in both conditions whereas, citric acid, DHAR and MDHAR displayed curves with different shapes, a narrow distribution around the mean under control conditions and a wider distribution under drought.

GWAS of the best GY-con predictors: chlorogenic acid, isocitric acid and citric acid

Using the control values of the top GY-con predictors, a total of 6 QTLs were detected (Fig. 4), 3 for chlorogenic acid, 2 for isocitric acid and 1 for citric acid (Quantile-Quantile plots available in Supplementary Fig. S2). The best GY-con predictors displayed a high marker-based heritability (or pseudo-heritability, h^2), varying between 0.34 and 0.66 (Supplementary Table S8), which suggests strong genetic control of their levels. This holds particularly true for

chlorogenic acid and isocitric acids (h^2 = 0.62 and 0.66, respectively) for which the same QTLs were detected under control and drought conditions. In contrast, citric acid showed a lower pseudo-heritability (h^2 = 0.34) and the single QTL identified under control conditions was not detected under drought (Supplementary Fig. S3).

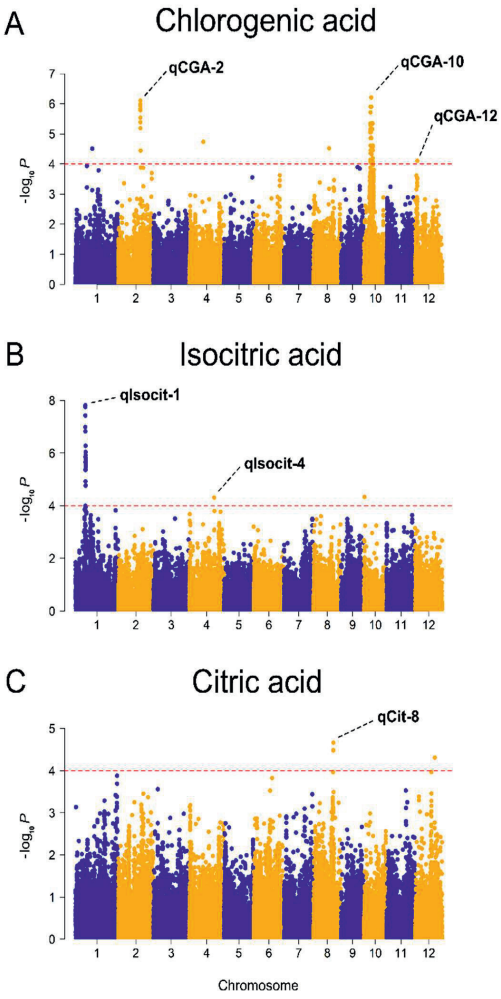


Fig. 4. Manhattan plots of GWA mapping results using control values of the best GY-con PLSR model predictors.

Manhattan plots of the GWA mapping results using 81,347 SNP markers for the control values of chlorogenic acid (A), isocitric acid (B) and citric acid (C). The red dashed line indicates the genome-wide significance threshold for QTLs (-$\log_{10}$ P > 4.0). Significant QTLs are indicated by name in the plots, where q=QTL, followed by a 3-6 letter abbreviation indicating the MetabOxi-predictor, and a numerical suffix indicating the chromosome location.

Among the three QTLs identified for chlorogenic acid, those on chromosomes 2 (*qCGA-2*) and 10 (*qCGA-10*) were the most significant. The top SNP for *qCGA-2* (LOD = 6.11) had a minor allele frequency (MAF) of 0.395, and was associated with lower chlorogenic acid values (Table 2). This minor allele is significantly ($P < 0.001$) associated with higher GY-con performance (+68.1 grams/m^2), in accordance with the negative correlation between chlorogenic acid and GY-con (Table 1). *qCGA-2* is also significantly associated with PH-con ($P < 0.001$) and FT-con ($P < 0.01$); accessions carrying the minor allele have lower levels of chlorogenic acid, and are both shorter in stature and earlier in flowering than accessions carrying the major allele at this locus. This observation is consistent with the positive correlation observed between chlorogenic acid levels, and PH-con and FT-con (Table 1) and identifies this QTL as a key determinant (possibly a rate-limiting factor) of the relationship between levels of chlorogenic acid, GY-con, PH and FT. The fact that there is substantial residual variation in PH-con and FT-con within both genotypic groups (Supplementary Fig. S4) underscores the value of *qCGA-2* as a target for rice GY improvement under agronomically optimal conditions.

qCGA-10 is associated with a large (5-10 Mbp) QTL region on chromosome 10 (Supplementary Fig. S5). The most significant SNP, *qCGA-10* (LOD = 6.21) has a minor allele (MAF = 0.280) associated with higher values of chlorogenic acid and lower GY-con performance (-65.6 grams/m^2), but it was not associated with significant differences in PH-con and FT-con (Table 2 and Supplementary Fig. S4). The absence of a significant PH/FT correlation makes this QTL an optimal target for altering leaf levels of chlorogenic acid and improving GY-con in rice under agronomically optimal conditions.

We next examined the *qCGA-2* and *qCGA-10* QTL regions for *a priori* candidate genes with known functions potentially associated with the observed variation in chlorogenic acid leaf levels. The *qCGA-2* marker maps within a putative UDP-glucosyltransferase (LOC_Os02g37690). Glucosylation increases the solubility and stability of phenolic compounds (Cui *et al.*, 2016) such as chlorogenic acid which accumulates to high levels in the cell vacuole where it exerts its antioxidative activity (Nakabayashi and Saito, 2015). Additionally, in *Medicago truncatula* and *Arabidopsis*, glucosylated polyphenols have been shown to be transported into the vacuole by MATE efflux transporter proteins (Marinova *et al.*, 2007; Pang *et al.*, 2008; Zhao and Dixon, 2009; Zhao *et al.*, 2011*a*). A MATE efflux protein

(LOC_Os10g13940) maps within the *qCGA-10* QTL region, localized ~59 kbp away from one of the most significant QTL markers, and a second MATE efflux protein (LOC_Os10g11860) is localized ~900 kbp upstream of the start of the QTL (Table 2 and Supplementary Table S9). While just outside the QTL detected in this study, MATE efflux protein "LOC_Os10g11860" was indicated as a candidate gene underlying a chlorogenic acid QTL in a previous GWA mapping study in rice (Chen *et al.*, 2014). The third significant QTL identified for chlorogenic acid (*qCGA-12*) (Table 2) is located on chromosome 12 and has a minor allele (MAF = 0.118) that is strongly associated with higher chlorogenic acid values, but not with differences in GY-con. *qCGA-12* is therefore of limited interest for GY improvement.

Two QTLs associated with variation in control values of isocitric acid were identified, one on chromosome 1 (*qIsocit-1*) and the second on chromosome 4 (*qIsocit-4*) (Fig. 4). Minor alleles at the two QTLs had opposite effects on organic acid values. The top *qIsocit-1* marker was highly significant (LOD = 7.82) and its minor allele (MAF = 0.247) was associated with higher values of isocitric acid (Table 2). The top *qIsocit-4* marker was less significant (LOD = 4.31) and its minor allele (MAF = 0.203) was associated with lower values of the metabolite (Supplementary Fig. S6). An isocitrate dehydrogenase (*OsIDHa*; LOC_Os01g16900) gene which catalyzes the production of α-ketoglutaric acid using isocitric acid as substrate in the TCA cycle (Table 2) (Abiko *et al.*, 2005) was identified just ~273 kbp upstream of the top *qIsocit-1* marker (Supplementary Table S9), and a second isocitrate dehydrogenase (LOC_Os04g42920) was localized inside the *qIsocit-4* QTL, approximately 28 kbp from the most significant marker (Table 2 and Supplementary Table S9). Despite the significant association with levels of isocitric acid, neither of these QTLs was associated with differences in GY-con (Table 2). We next examined the relationship between these QTLs as determinants of isocitric acid levels, given the opposing effects of their minor alleles. To do this, we included the most significant *qIsocit-1* marker as a covariate in the GWA model and re-ran the analysis using the control values of isocitric acid. We again detected a significant signal from *qIsocit-4*, confirming that the two QTLs are independent and explain different components of variation (Table 3 and Supplementary fig. S7). Of greater interest was the fact that, with *qIsocit-1* as a co-variate in the model, three additional QTLs were detected: *qIsocit-4+*, *qIsocit-7+* and *qIsocit-11+* (LOD = 4.77, 4.51 and 5.00, respectively). Signal from these three QTLs was also visible in the previous

GWAS but just below the threshold (Supplementary Fig. S7). When we examined the accessions carrying minor alleles at *qIsocit-4+*, *qIsocit-7+* and *qIsocit-11+,* (MAF = 0.269, 0.114, 0.133, respectively), we observed that they had lower values of isocitric acid and a strong reduction in GY-con (between -122.7 and -79.7 grams/m^2) (Table 3). The effects of the three additional QTLs help explain the positive correlation observed between control values of isocitric acid and GY-con in the 271 accessions of the panel (Table 1). Further, accessions carrying minor alleles at *qIsocit-4+*, *qIsocit-7+* and *qIsocit-11+* were also significantly taller than accessions carrying the major alleles (Table 3), consistent with the negative correlation between isocitric acid values and PH-con (Table 1). This is similar to the situation described for chlorogenic acid (*qCGA-2)*, but in contrast to *Isocit-4*, where individuals carrying the minor allele had lower isocitric acid levels but showed no significant association with PH or GY-con.

A single QTL for citric acid was detected on chromosome 8 (*qCit-8*) (Fig. 4). The minor allele of the most significant *qCit-8* marker was extremely rare (MAF = 0.054; present in only 15 individuals in the panel) and was associated with higher values of citric acid ($P < 0.01$) and higher GY-con ($P < 0.05$; +56.7 grams/m^2). No obvious candidate genes involved in the biosynthetic pathway of citric acid (i.e., citrate synthase or aconitase; Sweetlove *et al.*, 2010) were detected within the *qCit-8* QTL region (Supplementary Table S9) or its proximity. Nevertheless, because of its association with GY-con, we regard this QTL as a possible breeding target of interest for GY improvement under well-watered conditions.

Table 2. QTLs identified by GWA mapping of the best GY-con PLSR model predictors (from Table 1).

Trait	QTL name	Chr	MS SNP (pos)[a]	LOD score[b]	Alleles[c]	MAF[d]	Allele effects[e]							Candidate genes
							Trait (log_{10})	*P*[f]	GY-con ($gr./m^2$)	*P*	PH-con (cm)	*P*	FT-con (days)	*P*	
Chlorogenic acid	qCGA-2	2	22,742,146	6.11	T:C	0.395	-0.51	***	68.1	***	-11.5	***	-3.0	**	UDP-glucosyltranferase (LOC_Os02g37690)
	qCGA-10	10	7,476,500 7,515,294 7,515,295	6.21	T:C G:T A:C	0.280	0.31	***	-65.6	***	-1.7	n.s.	-0.9	n.s.	MATE efflux protein (LOC_Os10g11860) MATE efflux protein (LOC_Os10g13940)
	qCGA-12	12	2,182,479 2,182,516	4.11	C:G G:C	0.118	0.52	***	-3.5	n.s	9.7	*	1.3	n.s	not determined
Isocitric acid	qIsocit-1	1	9,943,382	7.82	A:G	0.247	0.17	***	-12.1	n.s	7.8	*	0.6	n.s	Isocitrate dehydrogenase (*OsIDHa*; LOC_Os01g16900)
	qIsocit-4	4	25,422,291	4.31	A:G	0.203	-0.10	*	16.2	n.s.	16	***	3.4	**	Isocitrate dehydrogenase (LOC_Os04g42920)
Citric acid	qCit-8	8	20,451,092	4.67	G:A	0.054	0.28	**	56.7	*	-4.3	n.s	-4.6	**	not determined

[a] Chromosomal position (bp) of the most significant (MS) SNP in the QTL region.

[b] Significance ($-log_{10}P$) of the MS SNP in the QTL region.

[c] MS SNP variants: major: minor allele.

[d] Minor allele frequency at the MS SNP.

[e] Allele effect for the MS SNP calculated as [mean trait value in accessions carrying the major allele] - [mean trait value in accessions carrying the minor allele].

[f] *P*-value of allele effect based on Welch's t-test for each trait: $P < 0.001$(***); $P < 0.01$(**); $P < 0.05$(*).

Table 3. QTLs identified by GWA mapping using the control values of isocitric acid and the most significant *qIsocit-1* marker as covariate.

Trait	QTL name	Chr	MS SNP (pos)[a]	LOD score[b]	Alleles[c]	MAF[d]	Allele effects[e]								Candidate genes
							Trait (log_{10})	*P*[f]	GY-con ($gr./m^2$)	*P*	PH-con (cm)	*P*	FT-con (days)	*P*	
Isocitric acid	qIsocit-4+	4	608,198	4.77	G:A	0.269	-0.17	***	-91.7	***	11	***	-0.8	n.s.	not determined
	qIsocit-4	*4*	*25,422,291*	*4.51*	*A:G*	*0.203*	*-0.10*	*	*16.2*	*n.s.*	*16*	***	*3.4*	**	see Table 2
	qIsocit-7+	7	28,647,934 28,648,012 28,648,013 28,648,016 28,652,483 28,652,495 28,652,644 28,669,634 28,673,761 28,674,817 28,679,613 28,679,648 28,679,741 28,679,815 28,679,937 28,683,809 28,684,459 28,684,877	4.70	T:C C:T G:C C:T C:T T:C A:C C:G G:A G:C T:C T:C T:C G:A A:G C:T T:G T:C	0.114	-0.12	*	-122.7	***	21.5	***	0.5	n.s.	not determined
	qIsocit-11+	11	25,906,572	5.00	G:A	0.133	-0.10	n.s.	-79.7	**	19.2	**	2.0	n.s.	not determined

[a] Chromosomal position (bp) of the most significant (MS) SNP in the QTL region.

[b] Significance ($-log_{10}P$) of the MS SNP in the QTL region.

[c] MS SNP variants: major: minor allele.

[d] Minor allele frequency at the MS SNP.

[e] Allele effect for the MS SNP calculated as [mean trait value in accessions carrying the major allele] - [mean trait value in accessions carrying the minor allele].

[f] *P*-value of allele effect based on Welch's t-test for each trait: $P < 0.001$(***); $P < 0.01$(**); $P < 0.05$(*).

GWAS of the best GY-dro and GY-loss predictors: DHAR, MDHAR and MDA

A total of ten QTLs were detected using the drought values/activities of the top GY-dro and GY-loss model predictors (Fig. 5), one for DHAR, five for MDHAR, and four for MDA (Quantile-Quantile plots available in Supplementary Fig. S9).

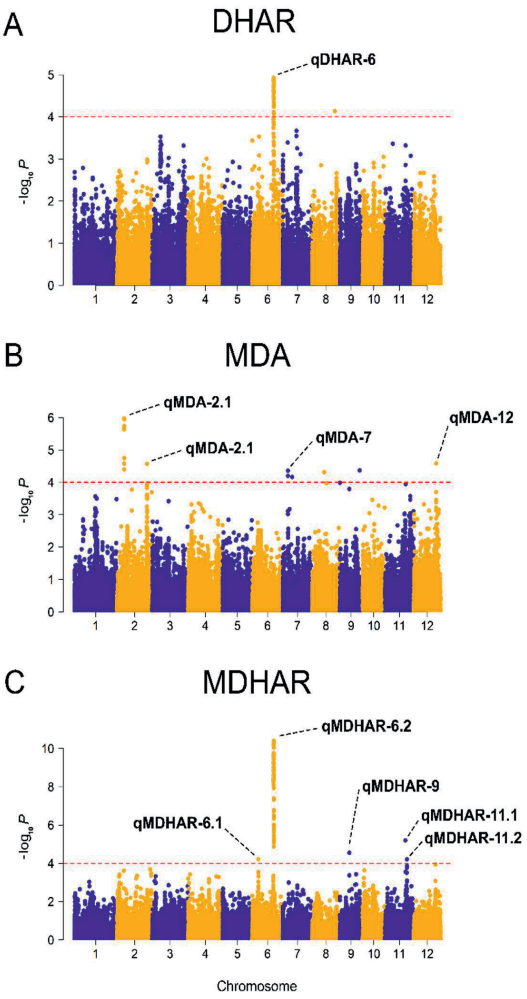


Fig. 5. Manhattan plots of GWA mapping results using drought values of the best GY-dro and GY-loss PLSR model predictors.

Manhattan plots of GWA mapping results using 81,347 SNP markers for the control values of DHAR (A), MDA (B) and MDHAR (C). Red dashed lines indicate the genome-wide significance threshold for QTLs ($-\log_{10} P > 4.0$). Significant QTLs are indicated by name in the plots, where q=QTL, followed by a 3-6 letter abbreviation indicating the MetabOxi-predictor, and a numerical suffix indicating the chromosome location.

The best GY-dro and GY-loss predictors showed h^2 values between 0.17 and 0.48, lower than the pseudo-heritability of the predictors for GY-con (Supplementary Table S8). The QTLs mapped using values/activities of DHAR, MDHAR and MDA estimated under drought conditions were entirely different than those mapped using control values/activities (Supplementary Fig. S10). This suggests that the variation associated with these ten QTLs is specifically associated with the stress environment, and could be exploited to improve rice GY performance under drought, potentially without impacting GY under control conditions.

A single QTL was identified on chromosome 6 (*qDHAR-6*) for the activity of DHAR under drought (Fig. 5). The top SNP of *qDHAR-6* (LOD = 4.93) had a MAF of 0.284, and the minor allele was associated with higher enzyme activity (Table 4), but not with GY differences (Table 4 and Supplementary Fig. S11). An ascorbate oxidase gene, *OsAAO1* (LOC_Os06g37080), was localized in the *qDHAR-6* QTL region ~10 kbp from the most significant QTL marker (Supplementary Table S10), and a homolog, *OsAAO2* (LOC_Os06g37150), was located in close proximity, ~68 kbp outside of the QTL region (Table 4). Both genes are annotated as responsive to salinity and drought stress in rice, with *OsAAO1* being expressed in both shoots and roots, and *OsAAO2* only in shoots (Batth *et al.*, 2017). Additionally, Wu *et al.* (2017) described *OsAAO2* as responsive to Fe-toxicity in rice shoots, with higher expression of the gene associated with tolerance. Ascorbate oxidases are cell-wall localized enzymes that oxidize reduced ascorbate (the active form utilized for ROS scavenging) to monodehydroascorbate, an unstable compound that spontaneously and rapidly converts to dehydroascorbate (Smirnoff, 2000). In turn, dehydroascorbate can be converted back to reduced ascorbate by the action of DHAR (Foyer and Noctor, 2011). Thus, considering the role of ascorbate oxidase in regulating the redox state of the apoplastic ascorbate pool, which is key in the response to abiotic stress in plants (Pignocchi and Foyer, 2003; Fotopoulos *et al.*, 2006; De Tullio *et al.*, 2013), *qDHAR-6* may indicate that variation in *OsAAO1* and/or *OsAAO2* plays a pivotal role in determining the activity of DHAR, without being directly responsible for the final activity values measured in the flag leaves of the accessions. We tested this hypothesis by re-running the association analysis for the activity of DHAR under drought, adding one of the top *qDHAR-6* markers as a covariate in the linear-mixed model.

Table 4. QTLs identified by GWA mapping using the drought values of the best GY-con and GY-loss PLSR model predictors.

Trait	QTL name	Chr	MS SNP (pos)[a]	LOD score[b]	Alleles[c]	MAF[d]	Allele effects[e]									Candidate genes
							Trait (log_{10})	*P*[f]	GY-dro (gr./m²)	*P*	GY-loss (%)	*P*	PH-dro (cm)	*P*	FT-dro (days)	*P*	
DHAR	qDHAR-6	6	21,883,335	4.93	A:G	0.284	0.14	***	-25.9	n.s.	2.6	n.s.	3.0	n.s.	1.9	n.s.	Ascorbate oxidase (*OsAAO1*; LOC_Os06g37080)
			21,883,406		G:A												Ascorbate oxidase (*OsAAO2*; LOC_Os06g37150)
MDA	qMDA-2.1	2	7,464,624	5.97	G:T	0.063	-0.25	**	107.9	**	-15.5	***	-6.5	n.s.	-2.3	n.s.	HSF-type protein (*OsHsfC2a*; LOC_Os02g13800)
			7,501,924		T:C												
			7,513,188		G:A												
			7,525,501		C:G												
			7,542,770		A:G												
	qMDA-2.2	2	30,699,332	4.57	T:C	0.295	-0.09	**	22.8	n.s	-5.1	*	5.8	n.s.	0.5	n.s.	not determined
	qMDA-7	7	5,263,699	4.36	C:T	0.065	-0.21	*	80.6	*	-12.9	**	-5.6	n.s.	-4.3	*	not determined
	qMDA-12	12	23,111,019	4.58	C:T	0.055	-0.26	*	110.8	**	-16.5	**	-26.8	***	-5.0	**	Phospholipase C (*OsPLC3*; LOC_Os12g37560)
MDHAR	qMDHAR-6.1	6	6,233,116	4.22	T:C	0.125	-0.19	**	7.8	n.s.	1.1	n.s.	13.1	**	1.6	n.s.	not determined
	qMDHAR-6.2	6	21,933,565	10.40	G:A	0.280	0.26	***	-25.1	n.s.	2.4	n.s.	2.9	n.s.	2.0	n.s.	Ascorbate oxidase (*OsAAO1*; LOC_Os06g37080)
			21,933,612		C:G												
			21,954,237		G:A												Ascorbate oxidase (*OsAAO2*; LOC_Os06g37150)
	qMDHAR-9	9	9,725,768	4.55	T:C	0.085	-0.14	n.s.	33.8	n.s.	1	n.s.	-17.3	**	3.6	n.s.	not determined
	qMDHAR-11.1	11	20,632,173	5.20	G:A	0.052	-0.27	*	12.9	n.s.	-7.1	n.s	-4.5	n.s.	-0.9	n.s.	not determined
	qMDHAR-11.2	11	22,178,608	4.21	C:A	0.183	-0.19	***	-31.4	n.s.	0.7	n.s.	4.2	n.s.	-3.9	***	not determined

[a] Chromosomal position (bp) of the most significant (MS) SNP in the QTL region.

[b] Significance ($-log_{10}P$) of the MS SNP in the QTL region.

[c] MS SNP variants: major: minor allele.

[d] Minor allele frequency at the MS SNP.

[e] Allele effect for the MS SNP calculated as [mean trait value in accessions carrying the major allele] - [mean trait value in accessions carrying the minor allele].

[f] *P*-value of allele effect based on Welch's t-test for each trait: $P < 0.001$(***); $P < 0.01$(**); $P < 0.05$(*).

Table 5. QTLs identified by GWA mapping using the drought activity values of DHAR and the most significant *qDHAR-6* marker as covariate.

Trait	QTL name	Chr	MS SNP (pos)[a]	LOD score[b]	Alleles[c]	MAF[d]	Allele effects[e]								Candidate genes
							Trait (log_{10})	*P*[f]	GY-dro ($gr./m^2$)	*P*	GY-loss (%)	*P*	PH-dro (cm)	*P*	FT-dro (days)	*P*	
DHAR	qDHAR-6+	6	7,119,281	4.47	G:C	0.081	-0.32	***	-94.3	***	13.4	**	-0.3	n.s.	-1.6	n.s.	Glutathione S-transferase (*OsDHAR2*; LOC_Os06g12630)

[a] Chromosomal position (bp) of the most significant (MS) SNP in the QTL region.

[b] Significance ($-log_{10}P$) of the MS SNP in the QTL region.

[c] MS SNP variants: major: minor allele.

[d] Minor allele frequency at the MS SNP.

[e] Allele effect for the MS SNP calculated as [mean trait value in accessions carrying the major allele] - [mean trait value in accessions carrying the minor allele].

[f] *P*-value of allele effect based on Welch's t-test for each trait: $P < 0.001$(***); $P < 0.01$(**); $P < 0.05$(*).

This analysis detected an additional QTL on chromosome 6 (*qDHAR-6+*) (Supplementary Fig. S12), similar to what we observed for *qIsocit-1*. The *qDHAR-6+* locus is characterized by a rare minor allele (MAF = 0.081) associated with a significant ($P < 0.001$) reduction in DHAR activity, a reduction in GY-dro (-94.3 grams/m^2), an increase in GY-loss (13.4%), and non-significant differences in PH-dro and FT-dro (Table 5 and Supplementary Fig. S13). The detection of *qDHAR-6+* helps explain the positive correlation of DHAR activity with GY-dro and the negative correlation with GY-loss (Table 1). Interestingly, one of the two DHAR isoforms (glutathione S-transferases) present in rice (Jain *et al.*, 2010), *OsDHAR2* (LOC_Os06g12630), is localized ~224 kbp from the top *qDHAR-6+* marker (Table 5). For these reasons, the *qDHAR-6+* QTL, and specifically the underlying *OsDHAR2* candidate gene, represent promising targets for manipulating DHAR activity and improving the GY performance of rice under drought stress. It is worth noting that 21 of the 22 accessions carrying the deleterious minor allele at the most significant *qDHAR-6+* marker originate from China (Supplementary Table S11 and Fig. S1), including the variety Minghui 63, one of the most successful male parents adopted for hybrid rice production in China (Xie and Zhang, 2018). It would be of interest to introduce the favorable allele of *OsDHAR2* (LOC_Os06g12630) into Minghui 63 and/or related hybrid varieties to determine the impact on GY performance under drought.

Five QTLs were identified for MDHAR under drought (Fig. 5), two on chromosome 6 (*qMDHAR-6.1* and *qMDHAR-6.2*), one on chromosome 9 (*qMDHAR-9*) and two on chromosome 11 (*qMDHAR-11.1* and *qMDHAR-11.2*). The minor alleles at the most significant markers for the five QTLs were not significantly associated with GY-dro or GY-loss differences (Table 4 and Supplementary Fig. S14) and, therefore, the QTLs are not of immediate interest as breeding targets for improving GY performance of rice under drought. This observation is consistent with the lack of correlation observed between the drought values of MDHAR and GY-dro/GY-loss (Table 1). Nevertheless, it is noteworthy that the most significant of the five QTLs, *qMDHAR-6.2* (LOD = 10.40), is closely linked to *qDHAR-6* at a distance of 50 - 70 kbp (Table 4). The *qMDHAR-6.2* QTL region (Supplementary Table S10) includes both *OsAAO1* (LOC_Os06g37080) and *OsAAO2* (LOC_Os06g37150), discussed above as *a priori* candidate genes for *qDHAR-6*. It is possible that genetic variation in the two ascorbate oxidases may be crucial to determining the different levels of MDHAR activity under drought, considering that

MDHAR directly reduces monodehydroascorbate to ascorbate before it spontaneously converts to dehydroascorbate (Smirnoff, 2000). In addition, by its activity, MDHAR may reduce the DHAR workload, thus increasing the efficiency of ascorbate reduction and, therefore, *qMDHAR-6.2* helps to explain, from a genetic perspective, why the activity of this enzyme under drought has been selected as an important predictor by the GY-dro and GY-loss PLSR models (Table 1).

Four QTLs were detected for the lipid peroxidation product MDA (Fig. 5), two on chromosome 2 (*qMDA-2.1* and *qMDA-2.2*), one on chromosome 7 (*qMDA-7*) and one on chromosome 12 (*qMDA-12*). The top marker for *qMDA-2.2* was not associated with significant differences in GY-dro or GY-loss (Table 4) and thus, this QTL is of marginal interest for breeding applications. In contrast, the other three QTLs were significantly associated with MDA values under drought, GY-dro and GY-loss. The minor alleles at all three loci were rare (MAF ranged from 0.055 to 0.063) but conferred a favorable effect, enhancing GY-dro between +80.6 to +110.8 grams/m^2, and decreasing GY-loss from -12.9% to -16.5% (Table 4 and Supplementary Fig. S11). Thus, these loci represent interesting breeding targets for improving rice grain yield stability under drought, and are consistent with the negative correlation between drought values of MDA and GY-dro and the positive correlation with GY-loss (Table 1). The most significant of the three QTLs is *qMDA-2.1* (LOD = 5.97), and one of the most significant SNP markers within this QTL is localized within a DNA-binding heat shock transcription factor, *OsHsfC2a* (LOC_Os02g13800) (Table 4 and Supplementary Table S10). *OsHsfC2a* is described as a primary player in the pathways involving ROS accumulation and sensing in rice plants exposed to oxidative and heat stresses (Mittal *et al.*, 2009). The second most significant MDA QTL is *qMDA-12* (LOD = 4.58), and a candidate gene encoding a phospholipase C (*OsPLC3*; LOC_Os12g37560) is located ~58 kbp from the most significant marker for this QTL (Table 4 and Supplementary Table S10). Phospholipases are enzymes that hydrolyze phospholipids and lead to the generation of lipid-derived messengers involved in many plant physiological processes, including stress responses (Chen *et al.*, 2012). It is particularly noteworthy that *OsPLC3* is specifically and strongly up-regulated in leaf tissue at the reproductive stage (panicle development) in rice plants exposed to drought stress (Singh *et al.*, 2013). The identification of two *a priori* candidate genes with stress-induced regulatory and signaling functions, such as

OsHsfC2a and *OsPLC3*, underlying two different MDA QTLs suggests that flag leaf levels of MDA under drought are determined by the action of several, possibly independent, genetic pathways, with genetic variation occurring in genes that regulate key, rate-limiting steps that impact GY performance under drought.

Validation of selected QTLs with an effect on GY performance in a second field trial

Although functional validation of the genetic associations (QTLs and underlying *a priori* candidate genes) disclosed and discussed above is beyond the scope of this study, we undertook a second field trial to determine whether the QTLs identified in the first experiment (2013) were significantly associated with GY differences the following year (2014). We used the same rice panel, and performed the experiment under the same conditions (control and drought stress) and in the same location (Kadam *et al.*, 2017). In the 2014 field trial, 268 accessions (out of the 271 from 2013) were used to evaluate GY performance under well-watered and drought conditions (Supplementary Table S1). The reduced number of accessions in 2014 only marginally altered the allele frequencies at the QTL markers identified in 2013. This allowed a fair comparison of the effect of these markers between years. Under control conditions, all six QTLs associated with GY-con identified in 2013 also displayed significant GY-con differences in 2014 (Table 6). The GY-con variation associated with these QTLs was, in fact, greater in 2014 than in the 2013 experiment, with the single exception of *qCGA-10*. Under drought conditions, the four QTLs associated with GY-dro/GY-loss identified in 2013 had a less marked effect in the 2014 experiment, and were not always significant (Table 7). These results are not entirely surprising, given the challenges involved in precisely duplicating the timing and intensity of drought stress, and the dynamic nature of plant response to stress. These results also suggest that the GY-associated QTLs determined by GWA mapping of the best GY-con biomarkers are less influenced by genotype-by-year interactions than the ones mapped for the best GY-dro/GY-loss biomarkers. This is likely dependent on the more stable environment present under well-watered conditions than under drought stress over the two years. Indeed, the stronger influence of environmental factors in determining the levels of the best GY-dro/GY-loss biomarkers than the GY-con ones is confirmed by their lower pseudo-heritability values (Supplementary Table 8). However, it is noteworthy that the effect of all the top markers of the GY-associated QTLs in 2013 was consistent in the 2014 (Table 6 and 7),

confirming their robustness as breeding targets to improve GY in rice. Interestingly, the GY-associated QTLs described in this study were not detected when the field trials (2013 and 2014) were used for the direct GWA mapping of GY and its components (Kadam *et al.*, 2018). This supports the validity of our approach which allowed to dissect the complex genetics of GY performance into relevant biochemical components (biomarkers) and to identify the fraction of the genetic variation of these components associated with GY only.

Table 6. Effect and significance of QTLs associated with GY-con in 2013 compared to the 2014 field trial.

QTL name	Chr	2013		2014	
		GY-con (gr./m²)	*P*	GY-con (gr./m²)	*P*
qCGA-2	2	68.1	***	126.1	***
qCGA-10	10	-65.6	***	-48.7	**
qIsocit-4+	4	-91.7	***	-116.4	***
qIsocit-7+	7	-122.7	***	-167.4	***
qIsocit-11+	11	-79.7	**	-131.9	***
qCit-8	8	56.7	*	104.0	**

Table 7. Effect and significance of QTLs associated with GY-dro/GY-loss in 2013 compared to the 2014 field trial.

QTL name	Chr	2013				2014			
		GY-dro (gr./m²)	*P*	GY-loss (%)	*P*	GY-dro (gr./m²)	*P*	GY-loss (%)	*P*
qDHAR-6+	6	-94.3	***	13.4	**	-78.0	**	14.4	**
qMDA-2.1	2	107.9	**	-15.5	***	76.8	n.s.	-10.2	n.s.
qMDA-7	7	80.6	*	-12.9	**	67.4	n.s.	-4.8	n.s.
qMDA-12	12	110.8	**	-16.5	**	112.3	**	-12.8	n.s.

CONCLUSIONS

This study has shown that in a large panel of phenotypically and genetically diverse rice accessions the flag leaf central metabolism and oxidative stress status during the reproductive stage represent a proxy for the plant's physiological status that is highly informative for predicting GY performance under both well-watered conditions and under drought stress. We showed that the predictive power of metabolites and oxidative stress markers/enzymes was equal to genetic markers for predicting GY-con, and significantly better than genetic markers

for predicting GY-dro and GY-loss. This underscores the importance of plant metabolism and oxidative stress status in contributing to phenotype and as the basis for improving crop GY, especially in the attempt to create drought resilient crop varieties. Our metabolome-oxidative stress-based multivariate models identified a set of biomarkers predictive of GY-con, GY-dro and GY-loss. Despite their value, using metabolites and enzyme activities as GY biomarkers is challenging because of their responsiveness to environmental changes, developmental stages and even diurnal variation. The efforts of collecting a large number of plant tissues in the field, in a limited time window, and synchronizing the developmental stage of many hundreds of accessions, as we did in this study, will likely remain a job for the fundamental research community rather than for breeders. However, in this study, we demonstrated that it is possible to convert GY biomarkers into genetic markers associated with GY performance. These DNA markers, and the underlying candidate genes they tag, represent easy and cost-effective tools that are readily available to the breeding community to improve rice GY under non-stress and drought-stress conditions.

Finally, the complex and polygenic nature of agronomic traits, such as GY, is often determined by many small effect loci that are difficult to identify by direct GWA mapping of the trait. The research approach described in the present study represents an effective strategy to begin to disentangle the dynamic complexity of grain yield and to facilitate the discovery of trait heritability hidden into its metabolic and enzymatic components.

ACKNOWLEDGEMENTS

This work is part of the 'Growing Rice like Wheat' research programme financially supported by an anonymous private donor, via Wageningen University Fund, for the first author's PhD fellowship (GM). We acknowledge financial support from the 'Rapid Mobilization of Alleles for Rice Cultivar Improvement in Sub-Saharan Africa' project at Cornell (GM, SMc) funded (through the Africa Rice Foundation) by the Bill & Melinda Gates Foundation.

SUPPLEMENTARY DATA

Supplementary data are available at:

https://drive.google.com/drive/folders/1t7R7GDWJja4KkugdDFtRLC963QfACgdL?usp

Fig. S1. Principal component analysis (PCA) based on 81,347 SNPs for the 271 *indica* rice accessions.

Fig. S2. Quantile-Quantile plots of GWA mapping results using values of the best GY-con PLSR model predictors.

Fig. S3. Manhattan plots showing GWA mapping results for control and drought values of the best GY-con PLSR model predictors.

Fig. S4. Box-plots displaying trait variation among accessions carrying major or minor alleles at top SNP markers detecting GWA-QTLs for control values of chlorogenic acid.

Fig. S5. Manhattan plot displaying a GWA peak on chromosome 10 for control values of chlorogenic acid.

Fig. S6. Box-plots displaying trait variation among accessions carrying major or minor alleles at top SNP markers detecting GWA-QTLs for control values of isocitric acid and citric acid.

Fig. S7. Manhattan plots of GWA mapping results for control values of isocitric acid.

Fig. S8. Box-plots displaying trait variation among accessions carrying major or minor alleles at top SNP markers detecting GWA-QTLs for control values of isocitric acid adding the top *qIsocit-1* marker as a covariate.

Fig. S9. Quantile-Quantile plots of GWA mapping results for values of the best GY-dro and GY-loss PLSR model predictors.

Fig. S10. Manhattan plots showing GWA mapping results for drought and control values of the best GY-dro and GY-loss PLSR model predictors.

Fig. S11. Box-plots displaying trait variation among accessions carrying major or minor alleles at top SNP markers detecting GWA-QTLs for drought values/activities of DHAR and MDA.

Fig. S12. Manhattan plots showing GWA mapping results for drought activity values of DHAR.

Fig. S13. Box-plots displaying trait variation among accessions carrying major or minor alleles at top SNP markers detecting GWA-QTLs for drought activity values of DHAR and using the top *qDHAR-6* marker as a covariate.

Fig. S14. Box-plots displaying trait variation among accessions carrying major or minor alleles at top SNP markers detecting GWA-QTLs for drought activity values of MDHAR.

Table S1. Grain yield, flowering and plant height for the 271 indica rice accessions.

Table S2. Flag leaf values of the 111 MetabOxi variables in the 271 indica rice accessions under control and drought stress conditions.

Table S3. Phenotypic trait performance of the panel.

Table S4. Predictability of the cross-validated PLSR models based on increasing latent variables.

Table S5. Ranking of the variables of the best MetabOxi-based PLSR models for the prediction of grain yield under control and drought conditions and for grain yield loss.

Table S6. Correlations between control values of the 111 MetabOxi variables and grain yield, flowering and plant height under control conditions.

Table S7. Correlations between drought values of the 111 MetabOxi variables and grain yield, flowering and plant height under control conditions and drought-induced grain yield loss.

Table S8. Phenotypic variance explained by the genetic markers.

Table S9. List of genes mapping within the QTL regions determined by GWA mapping of the best GY-con predictors.

Table S10. List of genes mapping within the LD blocks of the QTLs determined by GWA mapping of the best GY-dro/GY-loss predictors.

Table S11. List of accessions carrying the major and minor allele of the top qDHAR-6+ marker.

CHAPTER

6

General discussion

Plants are often exposed to conditions of limited water supply and this reduces their productivity. Ray *et al.* (2015) showed that the variation in temperature, precipitation and their interaction explains 32-39% of the globally average year-to-year yield variability of maize, rice, wheat and soybean during the period 1979-2008. In the coming decades, drought episodes associated with global climate change are projected to become more frequent and erratic. Recently, Leng and Hall (2019) estimated that by the end of the 21st century the risk of drought-induced grain yield loss in the major Asian rice producing countries will increase by 18-19% compared to present conditions. In this context, improving drought resilience in rice will be critical to meet the growing global food demand, particularly considering that, in Asia, ~40% of the total crop area is cultivated in rainfed agroecosystems (FAO, 2014) which are prone to droughts.

In this thesis I identified mechanisms of tolerance to drought, by investigating how stress-induced changes in the physiology, central metabolism, and oxidative stress status of rice impact its growth and yield. In Chapter 2, I show that a number of physiological, metabolic and antioxidative responses to drought are associated with stress-tolerance during the vegetative and reproductive stages in three selected rice varieties. In Chapter 3, I identify the coordinated activity of antioxidant enzymes from the ascorbate-glutathione cycle as an essential mechanism to prevent drought-induced yield loss in a population of rice accessions exposed to drought at the reproductive stage. I also studied drought-induced leaf metabolic reprogramming of the accessions, and found that photorespiration, protein degradation and nitrogen recycling were the main processes induced by drought. Of these three processes, photorespiration is caused by limitation of CO_2 diffusion into the leaf due to drought-induced stomata closure. Therefore, in Chapter 4, the differences in transpiration among the same accessions used in Chapter 3 were quantified by measuring canopy temperature, a proxy for stomatal conductance. Just as for Chapter 3, the accessions were exposed to drought at the reproductive stage and their canopy temperature during stress showed a significant negative correlation with grain yield, proving that leaf temperature under stress is a good predictor of plant performance. In addition, temperature data were used in a genome-wide association (GWA) study and a QTL for canopy temperature under drought was detected. In Chapter 5, I show that models based on the set of metabolites and oxidative stress markers/enzymes that

in Chapter 3 predicted drought-induced grain yield loss, accurately predict grain yield *per se* under well-watered and drought conditions in the ~270 accessions of the population I used. These metabolic- and oxidative stress-based models always predicted grain yield more accurately than genomic-based models. Finally, the best metabolic and stress-enzyme model predictors of grain yield were used as traits in a GWA study and the resulting associations allowed to identify genetic markers that can be used in breeding to improve rice grain yield under optimal conditions and/or grain yield/yield stability under drought stress. Here, in Chapter 6, I will discuss the main highlights of this thesis by connecting the results of the different experimental chapters. I will also discuss the possible follow-up research based on my findings and the prospects for their use in agriculture and breeding.

The developmental stage influences the effectiveness of drought tolerance mechanisms

Among the different plant developmental stages, the reproductive stage is the most sensitive to drought in cereals (Passioura, 2012; Biswal and Kohli, 2013; Reynolds *et al.*, 2016) and particularly in rice, in which even a moderate stress during flowering can induce a marked reduction in grain yield (Liu *et al.*, 2006; Venuprasad *et al.*, 2007). As a consequence, breeders have directed the majority of their efforts towards the improvement of rice drought tolerance at the reproductive stage, often using direct selection for grain yield under stress (Lanceras *et al.*, 2004; Venuprasad *et al.*, 2007; Kumar *et al.*, 2008, 2014; Vikram *et al.*, 2011). Nevertheless, when grown in rainfed agroecosystems, rice can be exposed to periods of drought between major rainfall events during its entire life cycle, including the vegetative stage (Kamoshita *et al.*, 2008). Additionally, considering the increased uncertainty and variability of rainfall patterns due to global warming (Reynolds *et al.*, 2016), in the coming years, the occurrence of drought stress will likely increase at every plant stage. In this perspective, a better understanding of drought tolerance mechanisms at different developmental stages might help in the selection of drought-resilient rice genotypes with an improved tolerance throughout the entire crop cycle.

In Chapter 2, I investigated the drought-induced changes in leaf central metabolism and oxidative stress status in three *indica* rice varieties exposed to drought during the vegetative and reproductive stage. These three varieties, a lowland (IR64), an aerobic (Apo) and an

upland (UPL Ri-7) one, were selected based on their good yield potential under well-watered conditions and their contrasting levels of tolerance to drought (low in IR64 and higher in Apo and UPL Ri-7). Overall, Chapter 2 clearly showed that, when drought is applied at the reproductive stage, the leaves of the three varieties maintained a stable sugar export (low levels of glucose and fructose in the leaves) with this highlighting the importance of the flag leaf as source of assimilates for the developing panicles and thus for yield (Yoshida, 1972; Biswal and Kohli, 2013). In contrast, when exposed to mild-to-severe drought at the vegetative stage, the leaves of the three varieties accumulated sugars (high levels of glucose and fructose), which provides osmotic protection against dehydration (Luquet *et al.*, 2008). The fact that this sugar-mediated osmotic adjustment did not occur at the reproductive stage, caused much stronger drought-induced oxidative stress (higher levels of lipid peroxidation and protein oxidation) at this stage. Under these conditions, only Apo, the variety displaying the strongest antioxidative response (enzymatic antioxidant activities), exhibited limited drought-induced grain yield loss. The importance of antioxidant enzymes for grain yield stability under drought at the reproductive stage was also discovered in Chapters 3 and 5 by multivariate modelling of leaf metabolome and oxidative stress status in a large association panel of phenotypically and genetically diverse *indica* rice accessions.

Focusing on the vegetative stage, Chapter 2 results highlighted that the accumulation of sugars (glucose and fructose) in the leaves of the improved upland variety UPL-Ri7 was a genotype-specific acclimation response that occurred early during drought. The sugar-mediated osmotic acclimation in UPL-Ri7 was equally effective in limiting stress-induced biomass loss as the strong antioxidative response displayed by Apo at a more severe intensity of stress. Lilley and Ludlow (1996) have shown the presence of wide genotypic variation in osmotic adjustment and dehydration avoidance in rice lines exposed to water-limited conditions. Particularly, they observed a greater osmotic adjustment capacity under stress among *indica* lines compared to *japonica* ones. Chapter 2 results confirm these genotypic differences in the three *indica* lines and highlight the importance of osmotic adjustment for drought tolerance at the vegetative stage. For these reasons, it would be of great interest to perform, at the vegetative stage, the same metabolic profiling that I conducted during flowering in Chapter 3 on the association mapping panel. This would allow to map possibly new genetic control mechanisms of sugar-

mediated osmotic adjustment, a prime drought stress adaptive trait (Blum, 2017), for which QTLs were already detected in rice using bi-parental populations (Robin *et al.*, 2003; Lilley *et al.*, 2007).

Metabolic and enzymatic biomarkers for grain yield stability under drought can complement genomic and genetic tools for the improvement of the trait

Despite genomics-assisted crop breeding has greatly contributed to the improvement of complex traits (Varshney *et al.*, 2014), closing the gap between genotype and phenotype remains a difficult breeding target. This is particularly true for a trait like grain yield under drought conditions which shows low heritability, complex polygenic control and, most importantly, a strong genotype-by-environment interaction (Cattivelli *et al.*, 2008). Plants are highly plastic in their response to dynamic environments and physiological trait plasticity is particularly important for the adaptation to extremely adverse environmental conditions (Dalal *et al.*, 2017), such as periods of drought. Plant metabolism is a readout of the plant physiological status and, thus, metabolites represent functional intermediate phenotypes linking the genotype to complex traits (Luo, 2015). For these reasons, targeting metabolic traits might represent a successful breeding strategy for improving crop yield under drought stress.

In Chapter 5, I found that a set of flag leaf central metabolites and oxidative stress markers/enzymes displayed a higher accuracy than DNA markers (SNPs) for the prediction of grain yield performance under drought in a large panel of *indica* rice accessions. Under well-watered conditions, the grain yield prediction accuracy of the metabolites and enzymes was equal to the one of the SNPs. These observations support the idea that metabolites and enzymes - because the effect of stress-induced dynamic regulatory processes is integrated in their levels and activities, respectively (Fiehn, 2002; Herrmann and Schauer, 2013) - are closer to plant grain yield performance under drought than the 'fixed' genetic markers. It is, therefore, under stressful conditions, more than under control, that the physiological information carried by the leaf metabolites and oxidative stress markers/enzymes can effectively complement genetics and genomics to improve drought tolerance in rice varieties.

In Chapter 3, I analyzed flag leaf primary metabolism under drought in the accessions of the panel and identified that photorespiration, protein degradation and nitrogen recycling were the main processes involved in the drought-induced leaf metabolic reprogramming. Additionally, by considering the oxidative stress status under drought of the same accessions, I showed that high levels of photorespiration and protein degradation were strongly linked with high values of malondialdehyde (MDA), a lipid peroxidation product indicative of membrane oxidative damage (Møller *et al.*, 2007). This link suggests that under drought, the high rate of peroxisomal H_2O_2 production, induced by enhanced photorespiration, is responsible for the oxidation of proteins and membrane lipids (Noctor, 2002; Avramova *et al.*, 2017). Plants counteract the enhanced generation of ROS and their oxidative action by enzymatic and non-enzymatic antioxidants (Miller *et al.*, 2010; Gill and Tuteja, 2010; You and Chan, 2015). In Chapter 3, I found that the activity of the antioxidant enzymes ascorbate peroxidase (APX), catalase (CAT) and dehydroascorbate reductase (DHAR) displays a negative correlation with the metabolic processes associated with drought stress. These results highlight how drought stress deeply alters cellular ROS homeostasis and that this has profound effects on leaf metabolism. In addition, even if genotypic variation in antioxidative responses to drought in different rice varietie was observed before (Selote and Khanna-Chopra, 2004; Basu *et al.*, 2009), the work conducted in Chapter 3 represents, to the best of my knowledge, the only large screening of drought-induced changes in the oxidative stress status of a crop species.

Besides representing the plant physiological status, metabolites and enzymes are also interesting breeding targets because they may serve as 'biomarkers' predictive of trait performance (Fernandez *et al.*, 2016). In Chapter 3 and 5, multivariate modelling and correlation analysis of flag leaf metabolome and oxidative stress status of the rice accessions revealed that DHAR and MDA were the best biomarkers for grain yield and grain yield loss under drought at the reproductive stage. These results further confirm that stress-induced alteration of leaf oxidative stress status is crucial for grain yield performance under drought at flowering and it is more predictive of trait variation than metabolic alterations.

The identification of DHAR and MDA as biomarkers for rice grain yield stability under drought is one of the main achievements of this thesis work. However, it is important to note that in

addition to being predictive, biomarkers need to be easy and cheap to score in order to be successfully applied in plant breeding (Zabotina, 2013; Fernandez *et al.*, 2016). As previously discussed, metabolites and enzymes are responsive to environmental and developmental changes (Florian *et al.*, 2014; Sulpice *et al.*, 2014; Sulpice and McKeown, 2015) and their analysis as biomarker would imply a huge effort in collecting a large number of samples in the field, during a limited time window and with all plants synchronized at the same plant developmental stage. In addition, these samples would need to undergo expensive laboratory procedures for the quantification of metabolites/enzyme activities. The time and costs associated with sample collection and laboratory work currently limits their use to fundamental research rather than to breeding. Nevertheless, in recent years, the development of field high-throughput phenotyping has allowed rapid and low-cost measurements of many phenotypes across time and space (Pauli *et al.*, 2016). Furthermore, the use of hyperspectral sensors has demonstrated the possibility to infer leaf chemical properties, including nitrogen and lignin content, with enormous throughput in different plant species (Martin *et al.*, 2008; Ustin *et al.*, 2009; Asner *et al.*, 2011; Feilhauer *et al.*, 2015). Thus, it can be conceived that, in the near future, the routine use in agriculture of high-throughput phenotyping devices (e.g gantries, drones and satellites) equipped with different types of sensors will enable the kind of biochemical markers described in my thesis in a time- and cost-effective manner. An alternative solution to translate the information carried by metabolic and enzymatic biomarkers into a breeder-friendly format is described in Chapter 5. By using the drought values of DHAR and MDA as quantitative traits in GWA mapping, I identified loci (SNPs) that represent the genetics that are underpinning these biomarkers associated with grain yield differences among the accessions of the panel. DNA markers are easy and cost-effective tools readily available to breeders. In addition, the loci identified in Chapter 5, were not determined by GWA mapping of grain yield under drought and, therefore, they represent a new source of genetic variation for improving rice grain yield under drought that was hidden in its metabolic and enzymatic components.

The importance of selecting the right trait for the right drought scenario

The thought-provoking idea that any trait or trait-related allele can increase or decrease drought tolerance depending on the drought scenario (Tardieu, 2012), highlights the

fundamental importance of genotype-by-environment interactions in determining plant productivity under water-limited conditions. I have already discussed above and in Chapter 1, how different mechanisms of drought tolerance in rice can have a positive or negative effect on yield based on length, intensity and timing (relative to plant developmental stage) of the drought event. Thus, it is of great importance to consider the prevailing water supply of a specific agroecosystem in order to identify the right combination of alleles and traits that can confer drought tolerance under those specific conditions (Tardieu *et al.*, 2018; Varshney *et al.*, 2018). In rice, most of the tall upland varieties and landraces have a good level of drought tolerance when grown in rainfed upland fields, where conditions of mild-to-severe drought often prevail, but perform poorly under irrigated lowland conditions (Bernier *et al.*, 2008). This is caused by the fact that tolerance to drought of upland rice varieties mainly relies on mechanisms of drought avoidance (e.g. early flowering, reduced leaf growth, limited tillering, deep rooting, and reduced stomatal conductance) which, by promoting a constitutive moderation of water use, hamper high plant productivity when water is available (Blum, 2005).

An example of the contrast between mechanisms of drought avoidance and yield maximization is the negative correlation between stomatal conductance (indirectly measured by canopy temperature) and plant height, observed in Chapter 4, among the accessions of the already mentioned association panel of *indica* rice lines. This negative correlation was present under both well-watered (less marked) and drought stress (more marked) conditions and confirms that reduced stomatal conductance is a constitutive water-saving trait associated with the tall and low-yielding upland rice varieties. In the same study, I also found that stomatal conductance was positively correlated with grain yield of the accessions under both treatments and, particularly, under drought stress. Considering that the closure of stomata to conserve water status contrasts with carbon assimilation and productivity (Luo, 2010), Chapter 4 results contribute to explain, from a physiological perspective, the negative link between drought avoidance mechanisms and high yield in rice (Price *et al.*, 2002). In addition, it is noteworthy that the maintenance of high stomatal conductance under stress of the high-yielding, semi-dwarf varieties of the panel did not represent a negative factor for their grain yield performance. This might be due to the advantageous evaporative cooling effect under

drought determined by the high transpiration of the short, high-yielding varieties that avoided heat stress (Jagadish *et al.*, 2015). Another possible explanation might be due to the fact that higher transpiration reduces photorespiration that, in Chapter 3, was shown to be associated with drought-induced premature leaf senescence and oxidation of membrane lipids, two processes that have strong negative effects on grain yield. However, it is important to remark that, in the experiment conducted in Chapter 4, the accessions of the stressed field were grown in paddy field conditions except for 14 consecutive days of drought stress treatment during the flowering stage. It is therefore likely that, under more prolonged and severe drought conditions, the sustained transpiration of the short, high-yielding varieties of the panel would have exposed their leaves to strong dehydration resulting in a drawback for their yield performance. The latter consideration reinforces the idea that each trait can have a positive or negative effect on yield depending on the drought scenario (Tardieu, 2012). Interestingly, in Chapter 4, a QTL associated with stomatal conductance differences among the accessions under drought was mapped on chromosome 3 and genetic variation for the significant markers of this QTL was present only within the tall, low-yielding upland varieties of the panel. The absence of genetic variation for this QTL in the high-yielding, semi-dwarf accessions suggests that selection for high grain yield under paddy conditions has resulted in reduced genetic variation for their stomatal response under drought. Thus, only the tall upland rice varieties and landraces possess the genetic variation associated with reduced stomatal conductance which could be exploited to induce a more conservative water use in the modern high-yielding varieties allowing their survival in growing environments exposed to prolonged water-limited conditions.

Under persistent water-limited conditions, the disruption of cellular homeostasis enhances the generation of reactive oxygen species (ROS), which, if they exceed the cellular ROS scavenging capacity, leads to oxidative damage to proteins, lipids and DNA (Halliwell, 2006; Suzuki *et al.*, 2012; Noctor *et al.*, 2014). As previously discussed, in Chapter 3 and 5, the lipid peroxidation product MDA was identified as an accurate biomarker of grain yield loss and grain yield *per se* under drought, suggesting that drought-induced oxidative damage is a major factor limiting rice productivity under stress. Chapter 2, 3 and 5 results show that the higher activity of antioxidant enzymes, particularly the ones involved in the ascorbate-glutathione

cycle, has a fundamental effect in improving rice grain yield stability under drought at the reproductive stage. Furthermore, in Chapter 2, I also found that, when exposed to severe drought at the vegetative stage, the strong antioxidant response displayed by the aerobic rice variety Apo was effective in limiting stress-induced biomass reduction. This indicates that a strong antioxidant system confers tolerance to drought at both the vegetative and reproductive stage. Different from constitutive drought avoidance mechanisms, which conflict with yield maximization when water is available, the ROS scavenging antioxidant defenses are only triggered by stress-induced ROS accumulation (Gill and Tuteja, 2010; You and Chan, 2015) and, therefore, should not result in yield penalty under sufficient water supply. Supporting this hypothesis, in Chapter 5, the QTLs mapped using the drought activities of the two ascorbate-glutathione antioxidative enzymes, DHAR and MDHAR, identified as highly predictive of grain yield and grain yield loss under drought, were not detected using their control activity values, suggesting that the genetic expression of these QTLs is stress-triggered only. In addition, the drought activities of DHAR and MDHAR did not show significant correlations with plant height differences among the accessions of the *indica* rice panel (Chapter 5). This further suggests that, in contrast to the reduced stomatal conductance, the higher activity of these two enzymes does not conflict with the better yield performance of the semi-dwarf rice varieties. For all these reasons, increasing the (drought-induced) ROS scavenging capacity of the modern high-yielding rice varieties might represent an effective and robust trait to improve their tolerance to periods of severe drought, and possibly during the entire crop cycle. The bio- and genetic markers I have discovered in my thesis may make an important contribution to this.

Concluding remarks

The results of this thesis work contribute to a deeper understanding on how drought-induced changes in stomatal regulation, central metabolism, and oxidative stress status impact rice physiology and, ultimately, biomass and grain yield. The identification of DHAR and MDA as biomarkers for grain yield stability under drought provide two important breeding targets for improving rice tolerance to drought. I also identified a set of genetic markers (SNPs) associated with these two biomarkers as well as with differences in stomatal conductance in a large panel of *indica* rice accessions. These SNPs are readily available to the breeders for improving rice productivity under drought or for increasing survival to the stress, through marker-assisted

selection. Finally, further analysis is needed to validate the *a priori* candidate genes that are underlying these markers. The confirmation of these genes as functional targets would allow their rapid modification through genome-editing technologies which would greatly accelerate the improvement of drought tolerance in rice.

References

AbdElgawad H, Zinta G, Hegab MM, Pandey R, Asard H, Abuelsoud W. 2016. High Salinity Induces Different Oxidative Stress and Antioxidant Responses in Maize Seedlings Organs. Frontiers in Plant Science **7**, 1–11.

Abiko T, Obara M, Ushioda A, Hayakawa T, Hodges M, Yamaya T. 2005. Localization of NAD-isocitrate dehydrogenase and glutamate dehydrogenase in rice roots: Candidates for providing carbon skeletons to NADH-glutamate synthase. Plant and Cell Physiology **46**, 1724–1734.

Aebi H. 1984. [13] Catalase in vitro. Methods in Enzymology.121–126.

Al-Tamimi N, Brien C, Oakey H, Berger B, Saade S, Ho YS, Schmöckel SM, Tester M, Negrão S. 2016. Salinity tolerance loci revealed in rice using high-throughput non-invasive phenotyping. Nature Communications **7**, 13342.

Apel K, Hirt H. 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annual Review of Plant Biology **55**, 373–99.

Arai-Sanoh Y, Takai T, Yoshinaga S, Nakano H, Kojima M, Sakakibara H, Kondo M, Uga Y. 2014. Deep rooting conferred by DEEPER ROOTING 1 enhances rice yield in paddy fields. Scientific Reports **4**, 1–6.

Araujo WL, Nunes-Nesi A, Nikoloski Z, Sweetlove LJ, Fernie AR. 2012. Metabolic control and regulation of the tricarboxylic acid cycle in photosynthetic and heterotrophic plant tissues. Plant, Cell & Environment **35**, 1–21.

Araújo WL, Tohge T, Ishizaki K, Leaver CJ, Fernie AR. 2011. Protein degradation - an alternative respiratory substrate for stressed plants. Trends in Plant Science **16**, 489–498.

Asada K. 2006. Production and scavenging of reactive oxygen species in chloroplasts and their functions. Plant Physiology **141**, 391–396.

Asner GP, Martin RE, Knapp DE, Tupayachi R, Anderson C, Carranza L, Martinez P, Houcheime M, Sinca F, Weiss P. 2011. Spectroscopy of canopy chemicals in humid tropical forests. Remote Sensing of Environment **115**, 3587–3598.

Atkin OK, Macherel D. 2008. The crucial role of plant mitochondria in orchestrating drought tolerance. Annals of Botany **103**, 581–597.

Atlin GN, Lafitte HR, Tao D, Laza M, Amante M, Courtois B. 2006. Developing rice cultivars for high-fertility upland systems in the Asian tropics. Field Crops Research **97**, 43–52.

Avila-Ospina L, Moison M, Yoshimoto K, Masclaux-Daubresse C. 2014. Autophagy, plant senescence, and nutrient recycling. Journal of Experimental Botany **65**, 3799–3811.

Avramova V, AbdElgawad H, Vasileva I, Petrova AS, Holek A, Mariën J, Asard H, Beemster GTS. 2017. High Antioxidant Activity Facilitates Maintenance of Cell Division in Leaves of Drought Tolerant Maize Hybrids. Frontiers in Plant Science **8**.

Bailey-Serres J, Fukao T, Ronald P, Ismail A, Heuer S, Mackill D. 2010. Submergence Tolerant Rice: SUB1's Journey from Landrace to Modern Cultivar. Rice **3**, 138–147.

Bañuelos MA, Garciadeblas B, Cubero B, Rodríguez-Navarro A. 2002. Inventory and functional characterization of the HAK potassium transporters of rice. Plant Physiology **130**, 784–95.

Basu S, Roychoudhury A, Saha PP, Sengupta DN. 2009. Differential antioxidative responses of indica rice cultivars to drought stress. Plant Growth Regulation **60**, 51–59.

Batth R, Singh K, Kumari S, Mustafiz A. 2017. Transcript Profiling Reveals the Presence of Abiotic Stress and Developmental Stage Specific Ascorbate Oxidase Genes in Plants. Frontiers in Plant Science **8**, 1–15.

Bauwe H, Hagemann M, Fernie AR. 2010. Photorespiration: players, partners and origin. Trends in Plant Science **15**, 330–336.

Baxter A, Mittler R, Suzuki N. 2014. ROS as key players in plant stress signalling. Journal of Experimental Botany **65**, 1229–1240.

Bentsink L, Alonso-Blanco C, Vreugdenhil D, Tesnier K, Groot SP, Koornneef M. 2000. Genetic analysis of seed-soluble oligosaccharides in relation to seed storability of Arabidopsis. Plant Physiology **124**, 1595–1604.

Benzie IFF, Strain JJ. 1999. [2] Ferric reducing/antioxidant power assay: Direct measure of total antioxidant activity of biological fluids and modified version for simultaneous measurement of total antioxidant power and ascorbic acid concentration. Methods in Enzymology.15–27.

Bernier J, Atlin GN, Serraj R, Kumar A, Spaner D. 2008. Breeding upland rice for drought resistance. Journal of the Science of Food and Agriculture **88**, 927–939.

Betti M, Bauwe H, Busch FA, *et al.* 2016. Manipulating photorespiration to increase plant productivity: Recent advances and perspectives for crop improvement. Journal of Experimental Botany **67**, 2977–2988.

Biswal AK, Kohli A. 2013. Cereal flag leaf adaptations for grain yield under drought: knowledge status and gaps. Molecular Breeding **31**, 749–766.

Blum A. 2005. Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive? Australian Journal of Agricultural Research **56**, 1159.

Blum A. 2017. Osmotic adjustment is a prime drought stress adaptive engine in support of plant production. Plant, Cell & Environment **40**, 4–10.

Borrell A, Garside A, Fukai S. 1997. Improving efficiency of water use for irrigated rice in a semi-arid tropical environment. Field Crops Research **52**, 231–248.

Bouman BA., Tuong T. 2001. Field water management to save water and increase its productivity in irrigated lowland rice. Agricultural Water Management **49**, 11–30.

Van Breusegem F, Dat JF. 2006. Reactive oxygen species in plant cell death. Plant Physiology **141**, 384–90.

Browning SR, Browning BL. 2007. Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies By Use of Localized Haplotype Clustering. The American Journal of Human Genetics **81**, 1084–1097.

Cabuslay GS, Ito O, Alejar AA. 2002. Physiological evaluation of responses of rice (Oryza sativa L.) to water deficit. Plant Science **163**, 815–827.

Carreno-Quintero N, Acharjee A, Maliepaard C, Bachem CWB, Mumm R, Bouwmeester H, Visser RGF, Keurentjes JJB. 2012. Untargeted Metabolic Quantitative Trait Loci Analyses Reveal a Relationship between Primary Metabolism and Potato Tuber Quality. Plant Physiology **158**, 1306–1318.

Carreno-Quintero N, Bouwmeester HJ, Keurentjes JJB. 2013. Genetic analysis of metabolome-phenotype interactions: From model to crop species. Trends in Genetics **29**, 41–50.

Carreno-Quintero N, Undas A, Bachem CWB, Mumm R, Visser RRGF, Bouwmeester HHJ, Keurentjes JJJB. 2014. Cross-platform comparative analyses of genetic variation in amino acid content in potato tubers. Metabolomics **10**, 1239–1257.

Cattivelli L, Rizza F, Badeck FW, Mazzucotelli E, Mastrangelo AM, Francia E, Marè C, Tondelli A, Stanca AM. 2008. Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. Field Crops Research **105**, 1–14.

Centritto M, Lauteri M, Monteverdi MC, Serraj R. 2009. Leaf gas exchange, carbon isotope discrimination, and grain yield in contrasting rice genotypes subjected to water deficits during the reproductive stage. Journal of Experimental Botany **60**, 2325–39.

Chaves MM, Flexas J, Pinheiro C. 2009. Photosynthesis under drought and salt stress: Regulation mechanisms from whole plant to cell. Annals of Botany **103**, 551–560.

Chen W, Gao Y, Xie W, *et al.* 2014. Genome-wide association analyses provide genetic and biochemical insights into natural variation in rice metabolism. Nature genetics **46**, 714–21.

Chen G, Greer MS, Lager I, Lindberg Yilmaz J, Mietkiewska E, Carlsson AS, Stymne S, Weselake RJ. 2012. Identification and characterization of an LCAT-like Arabidopsis thaliana gene encoding a novel phospholipase A. FEBS Letters **586**, 373–377.

Chrobok D, Law SR, Brouwer B, *et al.* 2016. Dissecting the metabolic role of mitochondria during developmental leaf senescence. Plant Physiology **172**, pp.01463.2016.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly **6**, 80–92.

Claeys H, Inze D. 2013. The Agony of Choice: How Plants Balance Growth and Survival under Water-Limiting Conditions. Plant Physiology **162**, 1768–1779.

Cobb JN, DeClerck G, Greenberg A, Clark R, McCouch S. 2013. Next-generation phenotyping: Requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. Theoretical and Applied Genetics **126**, 867–887.

Cuadros-Inostroza Á, Caldana C, Redestig H, Kusano M, Lisec J, Peña-Cortés H, Willmitzer L, Hannah MA. 2009. TargetSearch - a Bioconductor package for the efficient preprocessing of GC-MS metabolite profiling data. BMC Bioinformatics **10**, 428.

Cui L, Yao S, Dai X, *et al.* 2016. Identification of UDP-glycosyltransferases involved in the biosynthesis of astringent taste compounds in tea (Camellia sinensis). Journal of Experimental Botany **67**, 2285–2297.

Dalal A, Attia Z, Moshelion M. 2017. To Produce or to Survive: How Plastic Is Your Crop Stress Physiology? Frontiers in Plant Science **8**, 1–8.

Dan Z, Hu J, Zhou W, Yao G, Zhu R, Zhu Y, Huang W. 2016. Metabolic prediction of important agronomic traits in hybrid rice (Oryza sativa L.). Scientific Reports **6**, 21732.

Das K, Roychoudhury A. 2014. Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. Frontiers in Environmental Science **2**, 53 (1-13).

Degenkolbe T, Do PT, Kopka J, Zuther E, Hincha DK, Köhl KI. 2013. Identification of Drought Tolerance Markers in a Diverse Population of Rice Cultivars by Expression and Metabolite Profiling. PLoS ONE **8**, e63637.

Dhindsa RS, Plumb-Dhindsa PL, Reid DM. 1982. Leaf senescence and lipid peroxidation: Effects of some phytohormones, and scavengers of free radicals and singlet oxygen. Physiologia Plantarum **56**, 453–457.

Dimkpa SON, Lahari Z, Shrestha R, Douglas A, Gheysen G, Price AH. 2016. A genome-wide association study of a global rice panel reveals resistance in Oryza sativa to root-knot nematodes. Journal of Experimental Botany **67**, 1191–1200.

Dong X, Gao Y, Chen W, Wang W, Gong L, Liu X, Luo J. 2015. Spatiotemporal distribution of phenolamides and the genetics of natural variation of hydroxycinnamoyl spermidine in rice. Molecular Plant **8**, 111–121.

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.

Drotar A, Phelps P, Fall R. 1985. Evidence for glutathione peroxidase activities in cultured plant cells. Plant Science **42**, 35–40.

Dwivedi SL, Ceccarelli S, Blair MW, Upadhyaya HD, Are AK, Ortiz R. 2016. Landrace Germplasm for Improving Yield and Abiotic Stress Adaptation. Trends in Plant Science **21**, 31–42.

Fàbregas N, Fernie AR. 2019. The metabolic response to drought. Journal of Experimental Botany **33**, 585–588.

FAO. 2014. A Regional Rice Strategy for Sustainable Food Security in Asia and the Pacific. http://www.fao.org/3/a-i3643e.pdf.

Feierabend J, Beevers H. 1972. Developmental Studies on Microbodies in Wheat Leaves : I. Conditions Influencing Enzyme Development. Plant Physiology **49**, 28–32.

Feilhauer H, Asner GP, Martin RE. 2015. Multi-method ensemble selection of spectral bands related to leaf biochemistry. Remote Sensing of Environment **164**, 57–65.

Fernandez O, Urrutia M, Bernillon S, Giauffret C, Tardieu F, Le Gouis J, Langlade N, Charcosset A, Moing A, Gibon Y. 2016. Fortune telling: metabolic markers of plant performance. Metabolomics **12**, 158.

Fernie AR, Bauwe H, Eisenhut M, *et al.* 2013. Perspectives on plant photorespiratory metabolism. Plant Biology **15**, 748–753.

Fernie AR, Tohge T. 2017. The Genetics of Plant Metabolism. Annual Review of Genetics **51**, 287–310.

Ferreres F, Figueiredo R, Bettencourt S, *et al.* 2011. Identification of phenolic compounds in isolated vacuoles of the medicinal plant Catharanthus roseus and their interaction with vacuolar class III peroxidase: An H2O2 affair? Journal of Experimental Botany **62**, 2841–2854.

Fiehn O. 2002. Metabolomics - the link between genotyopes and phenotypes. Plant Molecular Biology **48**, 155–171.

Florian A, Nikoloski Z, Sulpice R, Timm S, Araújo WL, Tohge T, Bauwe H, Fernie AR. 2014. Analysis of short-term metabolic alterations in arabidopsis following changes in the prevailing environmental conditions. Molecular Plant **7**, 893–911.

Fotopoulos V, Sanmartin M, Kanellis AK. 2006. Effect of ascorbate oxidase over-expression on ascorbate recycling gene expression in response to agents imposing oxidative stress (A Kloss-Brandstaetter, Ed.). Journal of Experimental Botany **57**, 3933–3943.

Foyer CH, Noctor G. 2003. Redox sensing and signalling associated with reactive oxygen in chloroplasts, peroxisomes and mitochondria. Physiologia Plantarum **119**, 355–364.

Foyer CH, Noctor G. 2011. Ascorbate and glutathione: the heart of the redox hub. Plant Physiology **155**, 2–18.

Foyer CH, Shigeoka S. 2010. Understanding Oxidative Stress and Antioxidant Functions to Enhance Photosynthesis. Plant Physiology **155**, 93–100.

Fuller DQ, Sato YI, Castillo C, Qin L, Weisskopf AR, Kingwell-Banham EJ, Song J, Ahn SM, van Etten J. 2010. Consilience of genetics and archaeobotany in the entangled history of rice. Archaeological and Anthropological Sciences **2**, 115–131.

Furbank RT, Tester M. 2011. Phenomics--technologies to relieve the phenotyping bottleneck. Trends in Plant Science **16**, 635–44.

Gago J, Daloso D de M, Figueroa CM, Flexas J, Fernie AR, Nikoloski Z. 2016. Relationships of Leaf Net Photosynthesis, Stomatal Conductance, and Mesophyll Conductance to Primary Metabolism: A Multispecies Meta-Analysis Approach. Plant Physiology **171**, 265–279.

Galili G, Amir R, Fernie AR. 2015. The Regulation of Essential Amino Acid Synthesis and Accumulation in Plants. Annual Review of Plant Biology **67**, annurev-arplant-043015-112213.

Gill SS, Tuteja N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. Plant Physiology and Biochemistry **48**, 909–930.

Giunta F, Motzo R, Pruneddu G. 2008. Has long-term selection for yield in durum wheat also induced changes in leaf and canopy traits? Field Crops Research **106**, 68–76.

Gong L, Chen W, Gao Y, Liu X, Zhang H, Xu C, Yu S, Zhang Q, Luo J. 2013. Genetic analysis of the metabolome exemplified using a rice population. Proceedings of the National Academy of Sciences **110**, 20320–20325.

Habig WH, Pabst MJ, Jakoby WB. 1974. Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. The Journal of Biological Chemistry **249**, 7130–9.

Halliwell B. 2006. Reactive species and antioxidants. Redox biology is a fundamental theme of aerobic life. Plant Physiology **141**, 312–322.

Hare PD, Cress WA. 1997. Metabolic implications of stress-induced proline accumulation in plants. Plant Growth Regulation **21**, 79–102.

Hatfield J., Pinter P., Chasseray E, Ezra C., Reginato R., Idso S., Jackson R. 1984. Effects of panicles on infrared thermometer measurements of canopy temperature in wheat. Agricultural and Forest Meteorology **32**, 97–105.

He H, de Souza Vidigal D, Snoek LB, Schnabel S, Nijveen H, Hilhorst H, Bentsink L. 2014. Interaction between parental environment and genotype affects plant and seed performance in Arabidopsis. Journal of Experimental Botany **65**, 6603–6615.

Hedden P. 2003. The genes of the Green Revolution. Trends in Genetics **19**, 5–9.

Herrmann A, Schauer N. 2013. Metabolomics-Assisted Plant Breeding. In: Weckwerth W,, In: Kahl G, eds. The Handbook of Plant Metabolomics. Wiley-VCH Verlag GmbH & Co. KGaA.

Hildebrandt TM, Nunes Nesi A, Araújo WL, Braun H-P. 2015. Amino Acid Catabolism in Plants. Molecular Plant **8**, 1563–1579.

Hodges M, Dellero Y, Keech O, Betti M, Raghavendra AS, Sage R, Zhu XG, Allen DK, Weber APM. 2016. Perspectives for a better understanding of the metabolic integration of photorespiration within a complex plant primary metabolism network. Journal of Experimental Botany **67**, 3015–3026.

Hodges DM, Delong JM, Forney CF, Prange RK. 1999. Improving the thiobarbituric acid-reactive-substances assay for estimating lipid peroxidation in plant tissues containing anthocyanin and other interfering compounds. Planta **207**, 604–611.

Horie T, Matsuura S, Takai T, Kuwasaki K, Ohsumi A, Shiraiwa T. 2006. Genotypic difference in canopy diffusive conductance measured by a new remote-sensing method and its association with the difference in rice yield potential. Plant, Cell & Environment **29**, 653–660.

Horling F, Lamkemeyer P, König J, Finkemeier I, Kandlbinder A, Baier M, Dietz K. 2003. Dependent Regulation of Expression of the Peroxiredoxin Gene Family in Arabidopsis 1. Plant Physiology **131**, 317–325.

Hu H, Xiong L. 2014. Genetic engineering and breeding of drought-resistant crops. Annual Review of Plant Biology **65**, 715–41.

Huang X, Han B. 2014. Natural Variations and Genome-Wide Association Studies in Crop Plants. Annual Review of Plant Biology **65**, 531–551.

Hummel I, Pantin F, Sulpice R, *et al.* 2010. Arabidopsis Plants Acclimate to Water Deficit at Low Cost through Changes of Carbon Usage: An Integrated Perspective Using Growth, Metabolite, Enzyme, and Gene Expression Analysis. Plant Physiology **154**, 357–372.

Jagadish S, Craufurd P, Wheeler T. 2007. High temperature stress and spikelet fertility in rice (Oryza sativa L.). Journal of Experimental Botany **58**, 1627–1635.

Jagadish SVK, Murty MVR, Quick WP. 2015. Rice responses to rising temperatures - challenges, perspectives and future directions. Plant, Cell & Environment **38**, 1686–1698.

Jain M, Ghanashyam C, Bhattacharjee A. 2010. Comprehensive expression analysis suggests overlapping and specific roles of rice glutathione S-transferase genes during development and stress responses. BMC Genomics **11**, 73.

Jezek M, Blatt MR. 2017. The Membrane Transport System of the Guard Cell and Its Integration for Stomatal Dynamics. Plant Physiology. **174**, 487–519.

Ji K, Wang Y, Sun W, Lou Q, Mei H, Shen S, Chen H. 2012. Drought-responsive mechanisms in rice genotypes with contrasting drought tolerance during reproductive stage. Journal of Plant Physiology **169**, 336–44.

Jones HG. 1999. Use of thermography for quantitative studies of spatial and temporal variation of stomatal conductance over leaf surfaces. Plant, Cell & Environment **22**, 1043–1055.

Jones HG. 2013. *Plants and Microclimate: A Quantitative Approach to Environmental Plant Physiology*. Cambridge: Cambridge University Press.

Jones HG. 2014. The use of indirect or proxy markers in plant physiology. Plant, Cell & Environment **37**, 1270–1272.

Jones HG, Archer N, Rotenberg E, Casa R. 2003. Radiation measurement for plant ecophysiology. Journal of Experimental Botany **54**, 879–889.

Jones HG, Serraj R, Loveys BR, Xiong L, Wheaton A, Price AH. 2009. Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. Functional Plant Biology **36**, 978.

Kadam NN, Tamilselvan A, Lawas LMF, *et al.* 2017. Genetic Control of Plasticity in Root Morphology and Anatomy of Rice in Response to Water Deficit. Plant Physiology **174**, 2302–2315.

Kadam NN, Struik PC, Rebolledo MC, Yin X, Jagadish SVK. 2018. Genome-wide association reveals novel genomic loci controlling rice grain yield and its component traits under water-deficit stress during the reproductive stage. Journal of Experimental Botany **69**, 4017–4032.

Kamoshita A, Babu RC, Boopathi NM, Fukai S. 2008. Phenotypic and genotypic analysis of drought-resistance traits for development of rice cultivars adapted to rainfed environments. Field Crops Research **109**, 1–23.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong S-Y, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. Nature Genetics **42**, 348–354.

Keurentjes JJB, Sulpice R, Gibon Y, Steinhauser MC, Fu J, Koornneef M, Stitt M, Vreugdenhil D. 2008. Integrative analyses of genetic variation in enzyme activities of primary carbohydrate metabolism reveal distinct modes of regulation in Arabidopsis thaliana. Genome Biology **9**.

Kikuchi S, Bheemanahalli R, Jagadish KSV, Kumagai E, Masuya Y, Kuroda E, Raghavan C, Dingkuhn M, Abe A, Shimono H. 2017. Genome-wide association mapping for phenotypic plasticity in rice. Plant, Cell & Environment **40**, 1565–1575.

Kleessen S, Laitinen R, Fusari CM, Antonio C, Sulpice R, Fernie AR, Stitt M, Nikoloski Z. 2014. Metabolic efficiency underpins performance trade-offs in growth of Arabidopsis thaliana. Nature Communications **5**, 1–10.

Kooke R, Kruijer W, Bours R, *et al.* 2016. Genome-wide association mapping and genomic prediction elucidate the genetic architecture of morphological traits in Arabidopsis thaliana. Plant Physiology **170**, pp.00997.2015.

Krasensky J, Jonak C. 2012. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. Journal of Experimental Botany **63**, 1593–1608.

Kumar A, Bernier J, Verulkar S, Lafitte HR, Atlin GN. 2008. Breeding for drought tolerance: Direct selection for yield, response to selection and use of drought-tolerant donors in upland and lowland-adapted populations. Field Crops Research **107**, 221–231.

Kumar R, Bohra A, Pandey AK, Pandey MK, Kumar A. 2017. Metabolomics for Plant Improvement: Status and Prospects. Frontiers in Plant Science **8**, 1–27.

Kumar A, Dixit S, Ram T, Yadaw RB, Mishra KK, Mandal NP. 2014. Breeding high-yielding drought-tolerant rice: genetic variations and conventional and molecular approaches. Journal of Experimental Botany **65**, 6265–6278.

Kumar KB, Khan PA. 1983. Age-related changes in catalase and peroxidase activities in the excised leaves of Eleusine coracana Gaertin. cv PR 202 during senescence. Experimental Gerontology **18**, 409–17.

Kundu A, Vadassery J. 2019. Chlorogenic acid-mediated chemical defence of plants against insect herbivores. Plant Biology **21**, 185–189.

Lafitte H, Yongsheng G, Yan S, Li Z-K. 2006. Whole plant responses, key processes, and adaptation to drought stress: the case of rice. Journal of Experimental Botany **58**, 169–175.

Lanceras JC, Pantuwan G, Jongdee B, Toojinda T. 2004. Quantitative Trait Loci Associated with Drought Tolerance at Reproductive Stage in Rice. Plant Physiology **135**, 384–399.

Lauteri M, Haworth M, Serraj R, Monteverdi MC, Centritto M. 2014. Photosynthetic diffusional constraints affect yield in drought stressed rice cultivars during flowering. PLoS ONE **9**.

Leinonen I, Grant OM, Tagliavia CPP, Chaves MM, Jones HG. 2006. Estimating stomatal conductance with thermal imagery. Plant, Cell & Environment **29**, 1508–1518.

Lempe J, Lachowiec J, Sullivan AM, Queitsch C. 2013. Molecular mechanisms of robustness in plants. Current Opinion in Plant Biology **16**, 62–69.

Leng G, Hall J. 2019. Crop yield sensitivity of global major agricultural countries to droughts and the projected changes in the future. Science of the Total Environment **654**, 811–821.

Levine RL, Williams JA, Stadtman EP, Shacter E. 1994. [37] Carbonyl assays for determination of oxidatively modified proteins. Methods in Enzymology.346–357.

Lilley JM, Ludlow MM. 1996. Expression of osmotic adjustment and dehydration tolerance in diverse rice lines. Field Crops Research **48**, 185–197.

Lilley JM, Ludlow MM, McCouch SR, O'Toole JC. 2007. Locating QTL for osmotic adjustment and dehydration tolerance in rice. Journal of Experimental Botany **47**, 1427–1436.

Liu JX, Liao DQ, Oane R, Estenor L, Yang XE, Li ZC, Bennett J. 2006. Genetic variation in the sensitivity of anther dehiscence to drought stress in rice. Field Crops Research **97**, 87–100.

Liu HJ, Yan J. 2019. Crop genome-wide association study: a harvest of biological relevance. Plant Journal **97**, 8–18.

Lowry OH, Rosebrough NJ, Farr AL, Randall RJ. 1951. Protein measurement with the Folin-Phenol Reagent. The Journal of Biological Cemistry.

Luna A, Nicodemus KK. 2007. snp.plotter: An R-based SNP/haplotype association and linkage disequilibrium plotting package. Bioinformatics **23**, 774–776.

Lundberg M, Johansson C, Chandra J, Enoksson M, Jacobsson G, Ljung J, Johansson M, Holmgren A. 2001. Cloning and Expression of a Novel Human Glutaredoxin (Grx2) with Mitochondrial and Nuclear Isoforms. Journal of Biological Chemistry **276**, 26269–26275.

Luo LJ. 2010. Breeding for water-saving and drought-resistance rice (WDR) in China. Journal of Experimental Botany **61**, 3509–3517.

Luo J. 2015. Metabolite-based genome-wide association studies in plants. Current Opinion in Plant Biology **24**, 31–38.

Luquet D, Clément-Vidal A, Fabre D, This D, Sonderegger N, Dingkuhn M. 2008. Orchestration of transpiration, growth and carbohydrate dynamics in rice during a dry-down cycle. Functional Plant Biology **35**, 689–704.

Mackill DJ, Khush GS. 2018. IR64: a high-quality and high-yielding mega variety. Rice 2018 11:1 **11**, 18.

Maes WH, Steppe K. 2012. Estimating evapotranspiration and drought stress with ground-based thermal remote sensing in agriculture: A review. Journal of Experimental Botany **63**, 4671–4712.

Marinova K, Pourcel L, Weder B, Schwarz M, Barron D, Routaboul J-M, Debeaujon I, Klein M. 2007. The Arabidopsis MATE Transporter TT12 Acts as a Vacuolar Flavonoid/H + -Antiporter Active in Proanthocyanidin-Accumulating Cells of the Seed Coat . The Plant Cell **19**, 2023–2038.

Martin ME, Plourde LC, Ollinger S V., Smith ML, McNeil BE. 2008. A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. Remote Sensing of Environment **112**, 3511–3519.

Mather KA, Caicedo AL, Polato NR, Olsen KM, McCouch S, Purugganan MD. 2007. The extent of linkage disequilibrium in rice (Oryza sativa L.). Genetics **177**, 2223–2232.

Matsuda F, Nakabayashi R, Yang Z, Okazaki Y, Yonemaru JI, Ebana K, Yano M, Saito K. 2015. Metabolome-genome-wide association study dissects genetic architecture for generating natural variation in rice secondary metabolism. Plant Journal **81**, 13–23.

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

McCouch SR, Wright MH, Tung C-W, *et al.* 2016. Open access resources for genome-wide association mapping in rice. Nature Communications **7**, 10532.

Melandri G, AbdElgawad H, Riewe D, *et al.* 2019. Biomarkers for grain yield stability in rice under drought stress. Journal of Experimental Botany, 10.1093/jxb/erz221.

Mevik B-H, Wehrens R. 2007. The pls Package: Principle Component and Partial Least Squares Regression in R. Journal of Statistical Software **18**, 1–24.

Meyer RC, Steinfath M, Lisec J, *et al.* 2007. The metabolic signature related to high plant growth rate in Arabidopsis thaliana. Proceedings of the National Academy of Sciences **104**, 4759–4764.

Mhamdi A, Queval G, Chaouch S, Vanderauwera S, Van Breusegem F, Noctor G. 2010. Catalase function in plants: A focus on Arabidopsis mutants as stress-mimic models. Journal of Experimental Botany **61**, 4197–4220.

Michaeli S, Fromm H. 2015. Closing the loop on the GABA shunt in plants: are GABA metabolism and signaling entwined? Frontiers in Plant Science **6**, 419.

Miller G, Suzuki N, Ciftci-Yilmaz S, Mittler R. 2010. Reactive oxygen species homeostasis and signalling during drought and salinity stresses. Plant, Cell & Environment **33**, 453–467.

Mittal D, Chakrabarti S, Sarkar A, Singh A, Grover A. 2009. Heat shock factor gene family in rice: Genomic organization and transcript expression profiling in response to high temperature, low temperature and oxidative stresses. Plant Physiology and Biochemistry **47**, 785–795.

Mittler R. 2017. ROS Are Good. Trends in Plant Science **22**, 11–19.

Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V, Van Breusegem F. 2011. ROS signaling: The new wave? Trends in Plant Science **16**, 300–309.

Molina J, Sikora M, Garud N, *et al.* 2011. Molecular evidence for a single evolutionary origin of domesticated rice. Proceedings of the National Academy of Sciences **108**, 8351–8356.

Møller IM, Jensen PE, Hansson A. 2007. Oxidative Modifications to Cellular Components in Plants. Annual Review of Plant Biology **58**, 459–481.

Monna L, Kitazawa N, Yoshino R, Suzuki J, Masuda H, Maehara Y. 2002. Positional Cloning of Rice Semidwarfing Gene , sd1 : Rice " Green Revolution Gene " Encodes a Mutant Enzyme Involved in Gibberellin Synthesis. DNA Research, **17**, 11–17.

Morita S, Nakano H. 2011. Nonstructural Carbohydrate Content in the Stem at Full Heading Contributes to High Performance of Ripening in Heat-Tolerant Rice Cultivar Nikomaru. Crop Science **51**, 818.

Muller B, Pantin F, Génard M, Turc O, Freixes S, Piques M, Gibon Y. 2011. Water deficits uncouple growth from photosynthesis, increase C content, and modify the relationships between C and growth in sink organs. Journal of Experimental Botany **62**, 1715–1729.

Mumm R, Hageman JA, Calingacion MN, *et al.* 2016. Multi-platform metabolomics analyses of a broad collection of fragrant and non-fragrant rice varieties reveals the high complexity of grain quality characteristics. Metabolomics **12**, 1–19.

Munns R, James R a, Sirault XRR, Furbank RT, Jones HG. 2010. New phenotyping methods for screening wheat and barley for beneficial responses to water deficit. Journal of Experimental Botany **61**, 3499–3507.

Murshed R, Lopez-Lauri F, Sallanon H. 2008. Microplate quantification of enzymes of the plant ascorbate-glutathione cycle. Analytical Biochemistry **383**, 320–322.

Nakabayashi R, Saito K. 2015. Integrated metabolomics for abiotic stress responses in plants. Current Opinion in Plant Biology **24**, 10–16.

Nakashima K, Ito Y, Yamaguchi-Shinozaki K. 2009. Transcriptional Regulatory Networks in Response to Abiotic Stresses in Arabidopsis and Grasses. Plant Physiology **149**, 88–95.

Niggeweg R, Michael AJ, Martin C. 2004. Engineering plants with increased levels of the antioxidant chlorogenic acid. Nature Biotechnology **22**, 746–54.

Nishizawa A, Yabuta Y, Shigeoka S. 2008. Galactinol and Raffinose Constitute a Novel Function to Protect Plants from Oxidative Damage. Plant Physiology **147**, 1251–1263.

Noctor G. 2002. Drought and Oxidative Load in the Leaves of C3 Plants: a Predominant Role for Photorespiration? Annals of Botany **89**, 841–850.

Noctor G, Mhamdi A, Foyer CH. 2014. The Roles of Reactive Oxygen Metabolism in Drought: Not So Cut and Dried. Plant Physiology **164**, 1636–1648.

Nunes-Nesi A, Carrari F, Gibon Y, Sulpice R, Lytovchenko A, Fisahn J, Graham J, Ratcliffe RG, Sweetlove LJ, Fernie AR. 2007. Deficiency of mitochondrial fumarase activity in tomato plants impairs photosynthesis via an effect on stomatal function. Plant Journal **50**, 1093–1106.

Obata T, Fernie AR. 2012. The use of metabolomics to dissect plant responses to abiotic stresses. Cellular and Molecular Life Sciences **69**, 3225–3243.

Obata T, Witt S, Lisec J, Palacios-Rojas N, Florez-Sarasa I, Araus JL, Cairns JE, Yousfi S, Fernie AR. 2015. Metabolite profiles of maize leaves in drought, heat and combined stress field trials reveal the relationship between metabolism and grain yield. Plant Physiology **169**, 2665–2583.

Obidiegwu JE, Bryan G, Jones HG, Prashar A. 2015. Coping with drought: stress and adaptive responses in potato and perspectives for improvement. Frontiers in Plant Science **6**, 542.

Ohsumi A, Hamasaki A, Nakagawa H, Yoshida H, Shiraiwa T, Horie T. 2006. A Model Explaining Genotypic and Ontogenetic Variation of Leaf Photosynthetic Rate in Rice (Oryza sativa) Based on Leaf Nitrogen Content and Stomatal Conductance. Annals of Botany **99**, 265–273.

Ouyang W, Struik PC, Yin X, Yang J. 2017. Stomatal conductance, mesophyll conductance, and transpiration efficiency in relation to leaf anatomy in rice and wheat genotypes under drought. Journal of Experimental Botany **100**, 726–734.

Pandey V, Shukla A. 2015. Acclimation and Tolerance Strategies of Rice under Drought Stress. Rice Science **22**, 147–161.

Pang Y, Peel GJ, Sharma SB, Tang Y, Dixon RA. 2008. A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of Medicago truncatula. Proceedings of the National Academy of Sciences **105**, 14210–14215.

Parent B, Suard B, Serraj R, Tardieu F. 2010. Rice leaf growth and water potential are resilient to evaporative demand and soil water deficit once the effects of root system are neutralized. Plant, Cell & Environment **33**, no-no.

Passioura JB. 2012. Phenotyping for drought tolerance in grain crops: when is it useful to breeders? Functional Plant Biology **39**, 851.

Pauli D, Chapman SC, Bart R, Topp CN, Lawrence-Dill CJ, Poland J, Gore MA. 2016. The quest for understanding phenotypic variation via integrated approaches in the field environment. Plant Physiology **172**, pp.00592.2016.

Peng S, Bouman B, Visperas RM, Castañeda A, Nie L, Park HK. 2006. Comparison between aerobic and flooded rice in the tropics: Agronomic performance in an eight-season experiment. Field Crops Research **96**, 252–259.

Peterhansel C, Maurino VG. 2011. Photorespiration Redesigned. Plant Physiology **155**, 49–55.

Saint Pierre C, Crossa J, Manes Y, Reynolds MP. 2010. Gene action of canopy temperature in bread wheat under diverse environments. Theoretical and Applied Genetics **120**, 1107–1117.

Pignocchi C, Foyer CH. 2003. Apoplastic ascorbate metabolism and its role in the regulation of cell signalling. Current Opinion in Plant Biology **6**, 379–389.

Pingali PL. 2012. Green revolution: impacts, limits, and the path ahead. Proceedings of the National Academy of Sciences of the United States of America **109**, 12302–8.

Pinheiro C, Chaves MM. 2011. Photosynthesis and drought: can we make metabolic connections from available data? Journal of Experimental Botany **62**, 869–882.

Pires M V., Pereira Júnior AA, Medeiros DB, *et al.* 2016. The influence of alternative pathways of respiration that utilize branched-chain amino acids following water shortage in Arabidopsis. Plant, Cell & Environment **39**, 1304–1319.

Prashar A, Jones H. 2014. Infra-Red thermography as a high-throughput tool for field phenotyping. Agronomy **4**, 397–417.

Prashar A, Jones H. 2016. Assessing Drought Responses Using Thermal Infrared Imaging. In *Environmental Responses in Plants* (P Duque, Ed.). New York, NY, NY: Springer New York.

Prashar A, Yildiz J, McNicol JW, Bryan GJ, Jones HG. 2013. Infra-red Thermography for High Throughput Field Phenotyping in Solanum tuberosum. PLoS ONE **8**, e65816.

Price AH, Cairns JE, Horton P, Jones HG, Griffiths H. 2002. Linking drought-resistance mechanisms to drought avoidance in upland rice using a QTL approach: progress and new opportunities to integrate stomatal and mesophyll responses. Journal of Experimental Botany **53**, 989–1004.

Qiu X, Pang Y, Yuan Z, Xing D, Xu J, Dingkuhn M, Li Z, Ye G. 2015. Genome-wide association study of grain appearance and milling quality in a worldwide collection of Indica rice germplasm. PLoS ONE **10**, 1–25.

Ray DK, Gerber JS, MacDonald GK, West PC. 2015. Climate variation explains a third of global crop yield variability. Nature Communications **6**, 5989.

Rebetzke GJ, Rattey AR, Farquhar GD, Richards RA, Condon A (Tony) G. 2012. Genomic regions for canopy temperature and their genetic association with stomatal conductance and grain yield in wheat. Functional Plant Biology **40**, 14–33.

Rebolledo MC, Peña AL, Duitama J, Cruz DF, Dingkuhn M, Grenier C, Tohme J. 2016. Combining Image Analysis, Genome Wide Association Studies and Different Field Trials to Reveal Stable Genetic Regions Related to Panicle Architecture and the Number of Spikelets per Panicle in Rice. Frontiers in Plant Science **7**, 1384.

Reynolds M, Langridge P. 2016. Physiological breeding. Current Opinion in Plant Biology **31**, 162–171.

Reynolds MP, Quilligan E, Aggarwal PK, *et al.* 2016. An integrated approach to maintaining cereal productivity under climate change. Global Food Security **8**, 9–18.

Riedelsheimer C, Brotman Y, Méret M, Melchinger AE, Willmitzer L. 2013. The maize leaf lipidome shows multilevel genetic control and high predictive value for agronomic traits. Scientific Reports **3**, 2479.

Riedelsheimer C, Czedik-Eysenberg A, Grieder C, Lisec J, Technow F, Sulpice R, Altmann T, Stitt M, Willmitzer L, Melchinger AE. 2012. Genomic and metabolic prediction of complex heterotic traits in hybrid maize. Nature Genetics **44**, 217–220.

Riewe D, Grosman L, Zauber H, Wucke C, Fernie AR, Geigenberger P. 2008. Metabolic and Developmental Adaptations of Growing Potato Tubers in Response to Specific Manipulations of the Adenylate Energy Status. Plant Physiology **146**, 1579–1598.

Riewe D, Jeon HJ, Lisec J, Heuermann MC, Schmeichel J, Seyfarth M, Meyer RC, Willmitzer L, Altmann T. 2016. A naturally occurring promoter polymorphism of the Arabidopsis FUM2 gene causes expression variation, and is associated with metabolic and growth traits. Plant Journal **88**, 826–838.

Riewe D, Koohi M, Lisec J, Pfeiffer M, Lippmann R, Schmeichel J, Willmitzer L, Altmann T. 2012. A tyrosine aminotransferase involved in tocopherol synthesis in Arabidopsis. Plant Journal **71**, 850–859.

Rijsberman FR. 2006. Water scarcity: Fact or fiction? Agricultural Water Management **80**, 5–22.

Robin S, Pathan MS, Courtois B, Lafitte R, Carandang S, Lanceras S, Amante M, Nguyen HT, Li Z. 2003. Mapping osmotic adjustment in an advanced back-cross inbred population of rice. Theoretical and Applied Genetics **107**, 1288–1296.

Rodriguez RE, Lodeyro A, Poli HO, *et al.* 2006. Transgenic Tobacco Plants Overexpressing Chloroplastic Ferredoxin-NADP(H) Reductase Display Normal Rates of Photosynthesis and Increased Tolerance to Oxidative Stress. Plant Physiology, **43**, 639–649.

Ros R, Cascales-Miñana B, Segura J, Anoman AD, Toujani W, Flores-Tornero M, Rosa-Tellez S, Muñoz-Bertomeu J. 2013. Serine biosynthesis by photorespiratory and non-photorespiratory pathways: an interesting interplay with unknown regulatory networks. Plant Biology **15**, 707–712.

Rutkoski J, Poland J, Mondal S, Autrique E, Pérez LG, Crossa J, Reynolds M, Singh R. 2016. Canopy Temperature and Vegetation Indices from High-Throughput Phenotyping Improve Accuracy of Pedigree and Genomic Selection for Grain Yield in Wheat. Genes|Genomes|Genetics **6**, 2799–2808.

Sandhu N, Singh A, Dixit S, Sta Cruz M, Maturan P, Jain R, Kumar A. 2014. Identification and mapping of stable QTL with main and epistasis effect on rice grain yield under upland drought stress. BMC Genetics **15**, 63.

Santelia D, Lawson T. 2016. Rethinking Guard Cell Metabolism. Plant Physiology **172**, 1371–1392.

Sasaki A, Ashikari M, Ueguchi-Tanaka M, *et al.* 2002. A mutant gibberellin-synthesis gene in rice. Nature **416**, 701–702.

Scheet P, Stephens M. 2006. A Fast and Flexible Statistical Model for Large-Scale Population Genotype Data: Applications to Inferring Missing Genotypes and Haplotypic Phase. The American Journal of Human Genetics **78**, 629–644.

Schroeder JI, Kwak JM, Allen GJ. 2001. Guard cell abscisic acid signalling and engineering drought hardiness in plants. Nature **410**, 327–330.

Schwitzguebel J-P, Siegenthaler P-A. 1984. Purification of Peroxisomes and Mitochondria from Spinach Leaf by Percoll Gradient Centrifugation. Plant Physiology **75**, 670–674.

Selote DS, Khanna-Chopra R. 2004. Drought-induced spikelet sterility is associated with an inefficient antioxidant defence in rice panicles. Physiologia Plantarum **121**, 462–471.

Sharma P, Dubey RS. 2005. Drought induces oxidative stress and enhances the activities of antioxidant enzymes in growing rice seedlings. Plant Growth Regulation **46**, 209–221.

Sharma P, Jha AB, Dubey RS, Pessarakli M. 2012. Reactive Oxygen Species, Oxidative Damage, and Antioxidative Defense Mechanism in Plants under Stressful Conditions. Journal of Botany **2012**, 1–26.

Shi Y, Thomasson JA, Murray SC, *et al.* 2016. Unmanned Aerial Vehicles for High-Throughput Phenotyping and Agronomic Research. PLOS ONE **11**, e0159781.

Shulaev V, Cortes D, Miller G, Mittler R. 2008. Metabolomics for plant stress response. Physiologia Plantarum **132**, 199–208.

Sinclair TR. 2011. Challenges in breeding for yield increase for drought. Trends in Plant Science **16**, 289–293.

Singh A, Kanwar P, Pandey A, Tyagi AK, Sopory SK, Kapoor S, Pandey GK. 2013. Comprehensive Genomic Analysis and Expression Profiling of Phospholipase C Gene Family during Abiotic Stresses and Development in Rice. PLoS ONE **8**, 1–14.

Skirycz A, De Bodt S, Obata T, *et al.* 2010. Developmental stage specificity and the role of mitochondrial metabolism in the response of Arabidopsis leaves to prolonged mild osmotic stress. Plant Physiology **152**, 226–44.

Skirycz A, Inzé D. 2010. More from less: Plant growth under limited water. Current Opinion in Biotechnology **21**, 197–203.

Smirnoff N. 2000. Ascorbic acid: metabolism and functions of a multi-facetted molecule. Current Opinion in Plant Biology **3**, 229–235.

South PF, Cavanagh AP, Liu HW, Ort DR. 2019. Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. Science **363**, eaat9077.

Spielmeyer W, Ellis MH, Chandler PM. 2002. Semidwarf (sd-1), 'green revolution' rice, contains a defective gibberellin 20-oxidase gene. Proceedings of the National Academy of Sciences **99**, 9043–9048.

Steinfath M, Gärtner T, Lisec J, Meyer RC, Altmann T, Willmitzer L, Selbig J. 2010. Prediction of hybrid biomass in Arabidopsis thaliana by selected parental SNP and metabolic markers. Theoretical and Applied Genetics **120**, 239–247.

Stitt M, Sulpice R, Keurentjes J. 2010. Metabolic Networks: How to Identify Key Components in the Regulation of Metabolism and Growth. Plant Physiology **152**, 428–444.

Stitt M, Zeeman SC. 2012. Starch turnover: pathways, regulation and role in growth. Current Opinion in Plant Biology **15**, 282–292.

Sulpice R, Flis A, Ivakov AA, Apelt F, Krohn N, Encke B, Abel C, Feil R, Lunn JE, Stitt M. 2014. Arabidopsis coordinates the diurnal regulation of carbon allocation and growth across a wide range of Photoperiods. Molecular Plant **7**, 137–155.

Sulpice R, McKeown PC. 2015. Moving Toward a Comprehensive Map of Central Plant Metabolism. Annual Review of Plant Biology **66**, 187–210.

Sulpice R, Nikoloski Z, Tschoep H, *et al.* 2013. Impact of the carbon and nitrogen supply on relationships and connectivity between metabolism and biomass in a broad panel of Arabidopsis accessions. Plant Physiology **162**, 347–63.

Sulpice R, Pyl ET, Ishihara H, *et al.* 2009. Starch as a major integrator in the regulation of plant growth. Proceedings of the National Academy of Sciences of the United States of America **106**, 10348–10353.

Sulpice R, Trenkamp S, Steinfath M, *et al.* 2010. Network analysis of enzyme activities and metabolite levels and their relationship to biomass in a large panel of Arabidopsis accessions. The Plant Cell **22**, 2872–2893.

Suzuki N, Koussevitzky S, Mittler R, Miller G. 2012. ROS and redox signalling in the response of plants to abiotic stress. Plant, Cell & Environment **35**, 259–270.

Sweetlove LJ, Beard KFM, Nunes-Nesi A, Fernie AR, Ratcliffe RG. 2010. Not just a circle: Flux modes in the plant TCA cycle. Trends in Plant Science **15**, 462–470.

Szabados L, Savouré A. 2010. Proline: a multifunctional amino acid. Trends in Plant Science **15**, 89–97.

Tabbal DF, Bouman BAM, Bhuiyan SI, Sibayan EB, Sattar MA. 2002. On-farm strategies for reducing water input in irrigated rice; case studies in the Philippines. Agricultural Water Management **56**, 93–112.

Takahama U, Oniki T. 1997. A peroxidase/phenolics/ascorbate system can scavenge hydrogen peroxide in plant cells. Physiologia Plantarum **101**, 845–852.

Takahashi S, Badger MR. 2011. Photoprotection in plants: a new light on photosystem II damage. Trends in Plant Science **16**, 53–60.

Takahashi S, Murata N. 2008. How do environmental stresses accelerate photoinhibition? Trends in Plant Science **13**, 178–182.

Tardieu F. 2012. Any trait or trait-related allele can confer drought tolerance: just design the right drought scenario. Journal of Experimental Botany **63**, 25–31.

Tardieu F, Simonneau T, Muller B. 2018. The Physiological Basis of Drought Tolerance in Crop Plants: A Scenario-Dependent Probabilistic Approach. Annual Review of Plant Biology **69**.

Tester M, Langridge P. 2010. Breeding Technologies to Increase Crop Production in a Changing World. Science **327**, 818–822.

Todaka D, Zhao Y, Yoshida T, *et al.* 2017. Temporal and spatial changes in gene expression, metabolite accumulation and phytohormone content in rice seedlings grown under drought stress conditions. Plant Journal **90**, 61–78.

De Tullio MC, Guether M, Balestrini R. 2013. Ascorbate oxidase is the potential conductor of a symphony of signaling pathways. Plant Signaling and Behavior **8**, 1–4.

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

Ustin SL, Gitelson AA, Jacquemoud S, Schaepman M, Asner GP, Gamon JA, Zarco-Tejada P. 2009. Retrieval of foliar information about plant pigment systems from high resolution spectroscopy. Remote Sensing of Environment **113**, S67–S77.

Valluru R, Reynolds MP, Salse J. 2014. Genetic and molecular bases of yield-associated traits: A translational biology approach between rice and wheat. Theoretical and Applied Genetics **127**, 1463–1489.

Van den Ende W. 2013. Multifunctional fructans and raffinose family oligosaccharides. Frontiers in Plant Science **4**, 1–11.

Varshney RK, Terauchi R, McCouch SR. 2014. Harvesting the Promising Fruits of Genomics: Applying Genome Sequencing Technologies to Crop Breeding. PLoS Biology **12**, 1–8.

Varshney RK, Tuberosa R, Tardieu F. 2018. Progress in understanding drought tolerance: From alleles to cropping systems. Journal of Experimental Botany **69**, 3175–3179.

Venuprasad R, Bool ME, Quiatchon L, Atlin GN. 2012. A QTL for rice grain yield in aerobic environments with large effects in three genetic backgrounds. Theoretical and Applied Genetics **124**, 323–332.

Venuprasad R, Lafitte HR, Atlin GN. 2007. Response to direct selection for grain yield under drought stress in rice. Crop Science **47**, 285–293.

Verslues PE, Juenger TE. 2011. Drought, metabolites, and Arabidopsis natural variation: A promising combination for understanding adaptation to water-limited environments. Current Opinion in Plant Biology **14**, 240–245.

Vico G, Manzoni S, Palmroth S, Katul G. 2011. Effects of stomatal delays on the economics of leaf gas exchange under intermittent light regimes. New Phytologist **192**, 640–652.

Vikram P, Swamy BPM, Dixit S, Ahmed HU, Teresa M, Cruz S. 2011. Reproductive-Stage Drought Stress With a Consistent Effect in Multiple Elite Genetic Backgrounds. BMC Genetics **12**, 1–15.

Vikram P, Swamy BPM, Dixit S, Singh R, Singh BP, Miro B, Kohli A, Henry A, Singh NK, Kumar A. 2015. Drought susceptibility of modern rice varieties: An effect of linkage of drought tolerance with undesirable traits. Scientific Reports **5**, 1–18.

Vogt T. 2010. Phenylpropanoid biosynthesis. Molecular Plant **3**, 2–20.

Walker BJ, VanLoocke A, Bernacchi CJ, Ort DR. 2016. The Costs of Photorespiration to Food Production Now and in the Future. Annual Review of Plant Biology **67**, 107–129.

Wang DR, Agosto-Pérez FJ, Chebotarov D, Shi Y, Marchini J, Fitzgerald M, McNally KL, Alexandrov N, McCouch SR. 2018. An imputation platform to enhance integration of rice genetic resources. Nature Communications **9**, 3519.

Wang DR, Han R, Wolfrum EJ, McCouch SR. 2017. The buffering capacity of stems: genetic architecture of nonstructural carbohydrates in cultivated Asian rice, Oryza sativa. New Phytologist **215**, 658–671.

Watanabe M, Balazadeh S, Tohge T, Erban A, Giavalisco P, Kopka J, Mueller-Roeber B, Fernie AR, Hoefgen R. 2013. Comprehensive dissection of spatiotemporal metabolic shifts in primary, secondary, and lipid metabolism during developmental senescence in Arabidopsis. Plant Physiology **162**, 1290–1310.

Watanabe S, Matsumoto M, Hakamori Y, Takagi H, Shimada H, Sakamoto A. 2014. The purine metabolite allantoin enhances abiotic stress tolerance through synergistic activation of abscisic acid metabolism. Plant, Cell & Environment **37**, 1022–1036.

Weber APM, Bar-Even A. 2019. Improving the efficiency of photosynthetic carbon reactions. Plant Physiology, pp.01521.2018.

White JW, Andrade-Sanchez P, Gore MA, *et al.* 2012. Field-based phenomics for plant genetics research. Field Crops Research **133**, 101–112.

Wing RA, Purugganan MD, Zhang Q. 2018. The rice genome revolution: From an ancient grain to Green Super Rice. Nature Reviews Genetics **19**, 505–517.

Wingler a, Lea PJ, Quick WP, Leegood RC. 2000. Photorespiration: metabolic pathways and their role in stress protection. Philosophical Transactions of the Royal Society B: Biological Sciences **355**, 1517–1529.

Wingler A, Purdy S, MacLean JA, Pourtau N. 2006. The role of sugars in integrating environmental signals during the regulation of leaf senescence. Journal of Experimental Botany **57**, 391–399.

Wolosiuk RA, Crawford NA, Yee BC, Buchanan BB. 1979. Isolation of three thioredoxins from spinach leaves. The Journal of Biological Chemistry **254**, 1627–32.

Wu S, Tohge T, Cuadros-Inostroza Á, *et al.* 2018. Mapping the Arabidopsis Metabolic Landscape by Untargeted Metabolomics at Different Environmental Conditions. Molecular Plant **11**, 118–134.

Wu L-B, Ueda Y, Lai S-K, Frei M. 2017. Shoot tolerance mechanisms to iron toxicity in rice (Oryza sativa L.). Plant, Cell & Environment **40**, 570–584.

Xia H, Zheng X, Chen L, Gao H, Yang H, Long P, Rong J, Lu B, Li J, Luo L. 2014. Genetic differentiation revealed by selective loci of drought-responding EST-SSRs between upland and lowland rice in China. PLoS ONE **9**.

Xie F, Zhang J. 2018. Shanyou 63: an elite mega rice hybrid in China. Rice **11**.

Xing Y, Zhang Q. 2010. Genetic and Molecular Bases of Rice Yield. Annual Review of Plant Biology **61**, 421–442.

Xu S, Xu Y, Gong L, Zhang Q. 2016. Metabolomic prediction of yield in hybrid rice. The Plant Journal **88**, 219–227.

Yang J, Zhang J, Wang Z, Zhu Q, Wang W. 2001. Remobilization of carbon reserves in response to water deficit during grain filling of rice. Field Crops Research **71**, 47–55.

Yoshida S. 1972. Physiological Aspects of Grain Yield. Annual Review of Plant Physiology **23**, 437–464.

Yoshimura K, Ishikawa T, Nakamura Y, Tamoi M, Takeda T, Tada T, Nishimura K, Shigeoka S. 1998. Comparative study on recombinant chloroplastic and cytosolic ascorbate peroxidase isozymes of spinach. Archives of Biochemistry and Biophysics.

You J, Chan Z. 2015. ROS regulation during abiotic stress responses in crop plants. Frontiers in Plant Science **6**, 1092.

Zabotina O. 2013. Metabolite-Based Biomarkers for Plant Genetics and Breeding. In Diagnostics in plant breeding. T. Lübberstedt and R.K. Varshney editors, 1–521.

Zhang N, Gibon Y, Gur A, *et al.* 2010. Fine Quantitative Trait Loci Mapping of Carbon and Nitrogen Metabolism Enzyme Activities and Seedling Biomass in the Maize IBM Mapping Population. Plant Physiology **154**, 1753–1765.

Zhang C, Liu J, Zhao T, *et al.* 2016. A Drought-Inducible Transcription Factor Delays Reproductive Timing in Rice. Plant Physiology **171**, 334–343.

Zhang Q, Zhang J, Shen J, Silva A, Dennis DA, Barrow CJ. 2006. A Simple 96-Well Microplate Method for Estimation of Total Polyphenol Content in Seaweeds. Journal of Applied Phycology **18**, 445–450.

Zhao J, Dixon RA. 2009. MATE Transporters Facilitate Vacuolar Uptake of Epicatechin 3'-O-Glucoside for Proanthocyanidin Biosynthesis in Medicago truncatula and Arabidopsis. The Plant Cell **21**, 2323–2340.

Zhao J, Huhman D, Shadle G, He X-Z, Sumner LW, Tang Y, Dixon RA. 2011*a*. MATE2 Mediates Vacuolar Sequestration of Flavonoid Glycosides and Glycoside Malonates in Medicago truncatula . The Plant Cell **23**, 1536–1555.

Zhao K, Tung C-W, Eizenga GC, *et al.* 2011*b*. Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. Nature Communications **2**, 467.

Zhu G, Peng S, Huang J, Cui K, Nie L, Wang F. 2016. Genetic Improvements in Rice Yield and Concomitant Increases in Radiation- and Nitrogen-Use Efficiency in Middle Reaches of Yangtze River. Scientific Reports **6**, 1–12.

Zhu X, Tang G, Galili G. 2000. The catabolic function of the alpha-aminoadipic acid pathway in plants is associated with unidirectional activity of lysine-oxoglutarate reductase, but not saccharopine dehydrogenase. The Biochemical Journal **351**, 215–20.

Zia S, Romano G, Spreer W, Sanchez C, Cairns J, Araus JL, Müller J. 2013. Infrared Thermal Imaging as a Rapid Tool for Identifying Water-Stress Tolerant Maize Genotypes of Different Phenology. Journal of Agronomy and Crop Science **199**, 75–84.

Zinta G, AbdElgawad H, Domagalska MA, Vergauwen L, Knapen D, Nijs I, Janssens IA, Beemster GTS, Asard H. 2014. Physiological, biochemical, and genome-wide transcriptional analysis reveals that elevated CO 2 mitigates the impact of combined heat wave and drought stress in Arabidopsis thaliana at multiple organizational levels. Global Change Biology **20**, 3670–3685.

Summary

In **Chapter 1**, I explain that in the coming decades, drought episodes associated with climate change will be more frequent and erratic. Under this scenario, increasing or maintaining crop yields to meet the growing global food demand will become gradually more difficult. Rice (*Oryza sativa*), a staple food for more than half of the world's population, shows the greatest sensitivity to water limitation among the cereal crops. Improving drought tolerance in rice by limiting the stress-induced yield penalties is pivotal for global food security.

Drought stress impacts the physiology of plants and disrupts cellular homeostasis leading to metabolic alterations and increased oxidative stress. In this thesis, I investigate how drought-induced changes in rice physiology, central metabolism and oxidative stress status impact crop growth and yield. I also exploit the genetic diversity of a large panel of *indica* rice accessions to map genes and genomic regions associated with the quantitative variation in metabolic and physiological traits important for drought tolerance.

In **Chapter 2**, I study the physiological, metabolic and antioxidative responses to drought in three *indica* rice varieties selected for their contrasting levels of tolerance/susceptibility to the stress. The analysis was conducted during both the vegetative and reproductive stages and different mechanisms of tolerance to drought were identified between the different tolerant varieties and between stages. This study provides a framework for the exploration of the genetic control of these mechanisms of tolerance to drought.

In **Chapter 3**, I analyse the stress-induced changes in flag leaf central metabolism and oxidative stress status in ~300 *indica* rice accessions exposed to drought in the field at the reproductive stage. Photorespiration, protein degradation and nitrogen recycling were identified as the main flag leaf metabolic processes induced by drought. By integrating the metabolite data and the oxidative stress status of the accessions I showed that the activity of specific enzymatic antioxidants is important to limit the metabolic processes associated with drought stress which have a negative impact on grain yield. Finally, the levels of metabolites and oxidative stress markers/enzymes were also used to generate a multivariate model that accurately

predicts grain yield loss across the accessions. The best predictors of this model can be used as biomarkers for grain yield stability in rice under drought.

In **Chapter 4**, I quantify the differences in transpiration among the same accessions used in Chapter 3 by measuring canopy temperature, a proxy for stomatal conductance, in the field. Canopy temperature under drought at the reproductive stage was negatively correlated with the grain yield performance of the accessions, proving that leaf temperature under stress is a good predictor of drought tolerance that can be used to accelerate physiological selection in plant breeding. In addition, association mapping of canopy temperature data revealed a QTL associated with temperature differences under drought. Genetic variation for the significant markers of the QTL was present only within the tall, low-yielding landraces of rice adapted to drought-prone environments. This study confirms that these old varieties and landraces represent a strategic reservoir of genetic variation that can be tapped into for developing new varieties that are physiologically adapted to environments with unpredictable and variable water availability.

In **Chapter 5**, I show that the multivariate model based on the set of metabolites and oxidative stress markers/enzymes developed in Chapter 3 also accurately predicts grain yield *per se* under well-watered and drought conditions in ~270 accessions of the population. The latter model predicted grain yield more accurately than a genomics-based model that I developed for the same genotypes. Finally, the best metabolic and enzymatic model predictors of grain yield were used as traits in a GWA study and the resulting associations allowed me to identify genetic markers that can be used in breeding to improve rice grain yield under optimal conditions and/or grain yield/yield stability under drought stress.

Finally, in **Chapter 6**, I discuss the main findings of this thesis, connecting the results of the different experimental chapters and highlighting how they can be used to improve drought tolerance in rice.

Acknowledgements

The work presented in this thesis would not have been possible without the generosity of an anonymous private donor who provided financial support for my PhD fellowship and to whom I express, first and foremost, my sincere gratitude and respect.

Among the many people who contributed to this research, I would like to first thank my promotor **Prof. Harro Bouwmeester** and my co-promotor **Dr. Carolien Ruyter-Spira**. Since the beginning of the project at the Laboratory of Plant Physiology, your guidance has always been invaluable. I appreciated the freedom you gave me to explore the different topics of the experimental chapters. I also thank you for your constant encouragement and support in finishing the thesis, especially when I was losing my grip in the last year. Despite the initial struggle, we accomplished a lot and I am pleased to see that our efforts have taken shape in the form of this thesis.

Furthermore, I am grateful to all the members of collaborative 'Growing Rice Like Wheat' (GRLW) programme, to which my PhD project belongs to, and whose coordination was effectively conducted by **Dr. Gerard van der Linden.** Among the GRLW team members, my deepest gratitude goes to **Dr. Krishna Jagadish** and his heat stress team at IRRI, in particular, **Dr. Niteen Kadam**. Niteen's field experiments in the Philippines are at the basis of many of my thesis chapters. Niteen, I really enjoyed our converstions on rice drought stress physiology and quantitative genetics, both during my time at IRRI and later on when you moved to Wageningen.

A special thank goes to **Dr. David Riewe** for the metabolomics work we conducted together at IPK. Thanks to **Dr. Hamada Abdelgawad** and **Prof. Han Asard** for the oxidative stress analyses that they conducted at the University of Antwerpen. The extensive metabolic and enzymatic phenotyping presented in this thesis would not have been possible without their involvement. Equally important was the contribution of **Dr. Jos Hageman** for the statistical analysis of data. I would also like to thank **Dr. Ankush Prashar** for his help and guidance on the infrared imaging phenotyping.

My sincerest gratitude goes to **Prof. Susan McCouch** who gave me the opportunity to join her lab at Cornell University as a visiting scholar during the last two years of my PhD. This experience was crucial in expanding my knowledge of rice genetics. Susan, I would also like to thank you for letting me use part of the time at Cornell to work on finishing my PhD thesis and for your important contribution in two thesis chapters.

During my PhD years in Waegeningen and Ithaca, I have met so many people with whom I enjoyed my time. I have great memories of all the staff members, students and friends from the Laboratory of Plant Physiology and at the McCouch lab.

Finally, I want to thank my paranimphs **Dr. Beatriz Andreo Jimenez** and **Dr. Esmer Jongedijk**. Bea and Esmer, in these many years, you were not just colleagues but dear friends. I am really happy you are going to be here with me during my defense.

For further information on the author please visit:

https://www.linkedin.com/in/giovanni-melandri-02850312/

https://www.researchgate.net/profile/Giovanni_Melandri2

This work was performed at the Laboratory of Plant Physiology, Wageningen University. This work is part of the 'Growing Rice like Wheat' research programme financially supported by an anonymous private donor, via Wageningen University Fund.

Cover design and thesis layout by Giovanni Melandri.

Printed by Proefschriftmaken.nl – Digiforce Vianen.