THE UNIVERSITY of ADELAIDE

Purine Catabolism in Wheat: Source of Nutrients and Protective Metabolites

By

Alberto Casartelli

Thesis submitted in fulfilment of the requirements for the degree of **Doctorate of Philosophy** in the Faculty of Sciences at The University of Adelaide

School of Agriculture, Food and Wine

February 2018

Purine Catabolism in Wheat: Source of Nutrients and Protective Metabolites

By

Alberto Casartelli

Supervised by:

Dr Mamoru Okamoto
Research Scientist
Plant Genomics Centre
The University of Adelaide, AUS

Dr Vanessa Melino
Research Scientist
Plant Genomics Centre
The University of Adelaide, AUS

Dr Sigrid Heuer
Principle Scientist
Plant Science department
Rothamsted Research, UK

Dr Rainer Höfgen
Research Coordinator
Research Group Leader
Max Plant Institute of Molecular Plant Physiology, DE

Table of Contents

Abstract

Purine catabolism is known to have a dual function in recycling nitrogen (N) and carbon (C) atoms present in the heterocyclic purine ring and participating in stress signalling through the stimulation of ABA metabolism by allantoin, an intermediate in the pathway. However, little information was available of the functions in cereals and, in particular, bread wheat. The aims of this PhD thesis was to investigate the role of purine catabolism and allantoin in Australian bread wheat (*Triticum aestivum*) genotypes (RAC875 and Mace) grown under N deficiency and water deficit.

Firstly, the purine catabolic genes, coding for seven enzymes in total, were identified and annotated in the hexaploid bread wheat genome of cv. Chinese Spring. The analysis revealed 24 loci associated with the enzyme genes. Interestingly, there was a duplication of the xanthine dehydrogenase gene, namely *TaXDH1* and *TaXDH2*. Sequence analysis of the *TaXDH2* homeologs located on chromosome group 6 appeared to be either non-functional (*TaXDH2-6AS/6BS*) or with an inactive xanthine substrate binding site (*TaXDH2-6DS*). Protein structure modelling and a unique expression pattern under stress indicated that, TaXDH2-6DS may have a novel function in wheat.

Characterisation of allantoin and transcription of purine catabolic genes under N and water restrictions, revealed that allantoin levels were reduced (22-fold) when N was limiting, whilst it tended to accumulate in large amounts under drought (up to 30-fold compared to well-watered plants). The latter may suggest allantoin is used as a temporary N sink as the N assimilatory pathway (GS/GOGAT cycle) is likely to have a reduced capacity in plants growing in water deficit conditions. This would prevent the accumulation of ammonium that is toxic at high concentrations and reduce ammonia emissions from plants leaves. The reduction or accumulation of allantoin under drought and N stress appeared transcriptionally regulated by purine catabolic genes. In particular, transcription of *TaALN*, coding for the allantoin-degrading enzyme allantoinase, oppositely reflected the levels of allantoin in the tissue. Further growth studies showed that

wheat seedlings, when re-supplied with xanthine or allantoin as their sole N source after short-term N starvation, had growth rates which were equivalent to plants grown with inorganic nitrogen. This suggests that the N recycled through the purine catabolic pathway can support the growth of wheat. The data also provided evidence that allantoin takes part in N remobilisation during natural senescence.

Sequence analysis of *TaALN* homeologs in a large number of bread wheat accessions highlighted substantial genetic variability when compared to the reference genome of Chinese Spring. Candidate accessions with nucleotide polymorphisms in regulatory elements or in the coding sequence were identified and represent valuable material for future studies.

The outcomes of this PhD project provide the ground work for future fundamental research in wheat focussing on the dual role of purine catabolism in N recycling and abiotic stress. In addition, the candidate accessions identified, besides representing useful experimental material, could ultimately be used for breeding purposes. Additional strategies will include genetic engineering of target genes in the pathway that may lead to improved wheat growth and yields in unfavourable environments.

Thesis Declaration

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

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

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

Alberto Casartelli

October, 2017 30/10/2017

Acknowledgments

There are many people I would like to thank for making this PhD journey memorable.

I wish to acknowledge the Australian Centre for Plant Functional Genomics (ACPFG) for accepting my PhD application and providing a scholarship that allowed me to undertake a PhD at the University of Adelaide. I also thank the Australian Research Council (ARC) and DuPont-Pioneer for their funding that supported both my scholarship and the research project.

The successful completion of this PhD would have not been possible without the exceptional contribution of my three supervisors. They created a pleasant and productive work environment that allowed me to thrive academically and personally. I would like to express my deepest gratitude to Dr Vanessa Melino; thanks for her patience, guidance and invaluable comments throughout the project but also for the fun and laughs we had along with it! Her passion and dedication as an early career scientist truly inspired and motivated me. I am thankful to Dr Sigrid Heuer for assigning me with an engaging research project, her many helpful advices and discussions and for her friendly encouragement. I am sincerely grateful to Dr Mamoru Okamoto for his invaluable contributions and brilliant insights on plant N nutrition and for being a constant presence throughout my candidature.

I would like to thank my external advisor, Prof Ute Roessner, and her team for the considerable contribution to this project and for hosting me in their lab, which gave me the chance to learn a big deal about mass-spectrometry. A special thank goes to my external supervisor, Dr Rainer Höfgen, and his team at Max Planck Institute (Golm), having the change to spend several weeks at their institution was one of the highlights of my PhD.

I would like to take the opportunity to thank all the current and past members of the NUE group and, in particular, Dr Jessey George and Ms Akiko Enju for their support and assistance during my PhD. Thanks to those who shared the office space with me, it was great fun! I also wish to thank the admin team and

research staff at the Plant Genomics Centre for their help and support. In particular, I would like to thank the Bioinformatics group, led by Dr. Ute Baumann. Their excellent knowledge and assistance in relation to the wheat genome was critical in many occasions. I also thank Dr Melissa Pitman for her assistance in the *in-silico* protein modelling.

I was lucky to have many friends at the Waite Campus with whom I shared many fun moments of the daily routine and planned exciting activities and trips in and around Adelaide. Thank you to the Italian gang: Fabio, Matteo and Caterina; the beloved Frenchies: Allan, Mathieu, Pauline, Amelie, Marie, Abdel; and the many interns; the Dutchies: Yonina, Marianne and Tom; the locals: Nick, Will, Jess and Steph; and the many others from different countries: Juan Carlos, Vahid, Margaret, Judith, Gaffar, Jessica, Jose, Judith, Bart, Simon, Pualina, Kara, Mariam, Pia. It has been great being part of such an international group.

Lastly, I would like to express my deepest gratitude to my family and friends in Italy. Thanks to my mum, my dad and my twin brother for their never-ending support despite the distance and for believing in me.

List of Abbreviations

AAH	Allantoate Amidohydrolase
ABA	Abscisic Acid
ALN	Allantoinase
ANOVA	Analysis of Variance
AS	Allantoin Synthase
C	Carbon
cDNA	complementary DNA
CDS	Coding Sequence
CNRQ	Calibrated Normalised Relative Quantity
CO_2	Carbon Dioxide
D	Drought
DAA	Days After Anthesis
DAS	Days After Sowing
DNA	Deoxyribonucleic Acid
DW	Dry Weight
ER	Endoplasmic Reticulum
FW	Fresh Weight
GC-MS	Gas Chromatography - Mass Spectrometry
GOGAT	Glutamate Synthase
GPC	Grain Protein Content
GS	Glutamine Synthetase
GY	Grain Yield
HI	Harvest Index
HN	High Nitrogen
IC	Ion Chromatography
INDEL	Insertion/Deletion
IWGSC	International Wheat Genome Sequencing Consortium
JA	Jasmonic Acid
KO	Knock Out
LN	Low Nitrogen
M&M	Materials and Methods
mQTL	metabolic Quantitative Trait Loci

mRNA	messenger RNA
N	Nitrogen
N:C	Nitrogen to Carbon ratio
N_2	Dinitrogen
NAD^+/NADH	Nicotinamide Adenine Dinucleotide
NH_3	Ammonia
NH_4^+	Ammonium
NiR	Nitrite Reductase
NO_2^-	Nitrite
NO_3^-	Nitrate
NR	Nitrate Reductase
NRQ	Normalised Relative Quantity
NT	Nulli-Tetrasomic
NUE	Nitrogen Use Efficiency
P	Phosphorus
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
qRT-PCR	quantitative Reverse Transcription -PCR
RNA	Ribonucleic Acid
RNAi	RNA interference
ROS	Reactive Oxygen Species
S:R	Shoot to Root ratio
SEM	Standard Error of the Mean
SWC	Soil Water Content
TF	Transcription Factor
UAH	Ureidoglycolate Amidohydrolase
UGAH	Ureidoglycine Aminohydrolase
UOX	Urate Oxidase
UPLC	Ultra Performance Liquid Chromatography
UPS	Ureide Permease
WW	Well-Watered
XDH	Xanthine Dehydrogenase
YFEL	Youngest Fully Emerged Leaf

Chapter 1: Introduction

Increasing world population and climate changes are currently threatening global food production. Future agriculture systems will be required to be more productive and, at the same time, more sustainable. Currently, cereals are the major staple food worldwide and wheat (*Triticum aestivum* L.) alone provides more than the 20% of the daily recommended intake of calories and proteins (Hawkesford *et al.*, 2013). Nitrogen (N) fertilisers are required at high rates in order to reach wheat yield potentials and, for example, it is the most expensive input for growers in Australia. In drought prone regions of South Australia and other dry areas wheat yields are limited by the interaction of water and N, rather than being affected by a single factor alone (Sadras, 2005; Cossani *et al.*, 2010). Drought in southern Australian environments usually occurs in the field during terminal growth of spring wheat, which corresponds to the reproductive phase. By this time, up to 80% of the total plant N is already taken up (McDonald and Hooper, 2013), suggesting that terminal drought more likely affects processes of N recycling and loading to the grain rather than N uptake from the soil.

In recent time, metabolomics has gained popularity among plant scientists as it allows the simultaneous detection of a large number of metabolites and can provide a "snapshot" of the metabolic status of a plant/tissue (Beckles and Roessner, 2012). This approach can be used to study how N homeostasis is affected by abiotic stresses such as drought. Small N-rich metabolites, such as amino acids, are easily detected and quantified by several metabolomics techniques. A recent study used gas chromatography coupled with mass spectrometry (GC-MS) to analyse the metabolomic response of highly tolerant traditional aus-type rice to drought (Casartelli *et al.*, 2018; Appendix 1). When compared to drought sensitive cultivars, tolerant varieties showed higher accumulation of specific N-rich metabolites in their shoots. In particular, allantoin was the metabolite with highest magnitude of accumulation. Allantoin is an intermediate of purine catabolism and its oxidation generates glyoxylate, three molecules of CO_2 and four molecules of ammonium (Werner and Witte, 2011). There are a number of studies reporting allantoin to

differentially accumulate under different stress conditions in a wide range of plant species. In addition, allantoin was shown to enhance the metabolism of stress-related phytohormones abscisic acid (ABA) and jasmonic acid (JA) in the model plant *Arabidopsis thaliana* (Watanabe *et al.*, 2014; Takagi *et al.*, 2016). Purine catabolism has been extensively studied in tropical legumes because of its pivotal role in N translocation between plant compartments (Herridge *et al.*, 1978; Pate *et al.*, 1980). However, the role of allantoin and purine catabolism in relation to N homeostasis has been neglected in plants that uses amino acids as their main forms for long-distance transport of N, such as monocot grasses.

The overall aim of this dissertation was to investigate the role of purine catabolism and specifically allantoin degradation and their contribution to the N economy of bread wheat exposed to water or N stress. In brief, the specific aims of this study were to 1) identify purine catabolic genes in the bread wheat hexaploid genome; 2) assess the response of both allantoin accumulation and transcription of purine catabolic genes to drought or N deficiency; 3) investigate the role of purine catabolism in N recycling and remobilisation in wheat; 4) identify bread wheat accessions with allelic variants in key purine catabolic genes.

This thesis consists of six chapters:

Chapter 1 (*Introduction*) provides a broad overview of the thesis background. Research gaps and specific aims of this thesis are briefly discussed.

Chapter 2 (*Literature review*) provides a comprehensive review of the available literature setting the background to plant N homeostasis under stress and the function of N-rich metabolites in nutrient recycling. The possible role of purine catabolism and its intermediate allantoin under abiotic stress is described in detail, highlighting the current research gaps of this field.

Chapter 3 (*The allantoin pathway in bread wheat is responsive to abiotic stresses and nitrogen supply*), written in manuscript style, provides the fundaments on the molecular characterisation of purine

catabolic genes in bread wheat genome and their relationship with orthologous genes from other grass genomes. The manuscript also provides preliminary observations regarding the regulation of purine catabolic genes transcription under abiotic stress and allantoin levels under N deficiency.

Chapter 4 (*The dual role of the purine catabolic intermediate allantoin under stress and N homeostasis in bread wheat*), written in manuscript style, provides detailed analysis of allantoin accumulation and the transcriptional response of purine catabolic genes to drought and N deficiency in two Australian bread wheat genotypes. Furthermore, pieces of evidence on the role of allantoin as a N source for seedling growth and storage in wheat grains are provided and discussed in relation to other N sources.

Chapter 5 (*Allantoinase (ALN) genetic diversity in bread wheat*) is an in-depth analysis of the key purine catabolic gene *TaALN* and provides evidence of putative regulatory elements and transcription factors that are possibly involved in regulating *TaALN* expression under stress. The chapter also includes the identification of candidate wheat accessions with *TaALN* allelic variants that could be used for future studies.

Chapter 6 (*General discussion & future directions*) provides a discussion on how the findings reported in this dissertation contributed to the field of purine catabolism research and provides strategies for future research directions.

This thesis also contains three appendices:

Appendix 1: **Casartelli A, Riewe D, Hubberten HM, Altmann T, Hoefgen R and Heuer S**. *Exploring traditional aus-type rice for metabolites conferring drought tolerance*. Rice **11**, 9. The data reported in this manuscript served as fundamentals for designing this PhD research project in wheat. The experimental data was generated before commencing the PhD candidature. The manuscript was prepared for publication during the course of the PhD candidature.

Appendix 2: *Arabidopsis constitutive and inducible ALN transgenic lines*. These materials were produced in collaboration with Dr Rainer Höfgen (MPI-Golm). The observations and hypothesis made during this PhD project led to this experimental approach in order to further investigate the function of ALN in N homeostasis.

Appendix 3: *Rice insertional mutants in purine catabolic genes*. These lines will be used as genetic material to further assess the function of key purine catabolic genes in cereals. Rice was chosen as a model because it is a diploid organism and several mutant collections are currently available.

In the manuscript-style chapters (3 and 4), minor changes have been made to the numbering of tables and figures to provide a consistent format throughout the thesis. The following chapters presented in manuscript style are cross-referenced in other chapters as Chapter 3, Casartelli *et al.* (Chapter 3); Chapter 4 Casartelli *et al.* (Chapter 4); Appendix 1,Casartelli *et al.*,2018 (Appendix 1).

Chapter 2: Literature review

2.1 Increasing cereal production in harsh environments

2.1.1 Global food production

Currently, cereals are the major staple food worldwide, representing more than the 50% of the total human daily calories intake (Hawkesford, 2014) and wheat alone provides more than 20% of the daily recommended intake of calories and proteins (Hawkesford *et al.*, 2013). The 'green revolution' occurring between 1930s and the late 1960s with the introduction of dwarf high-yielding new crop varieties and changes of agricultural practises enabled a tremendous increase of crop productivity, especially for cereals, such as wheat (Farmer, 2008). As a result, during the last fifty years, global wheat production has increased despite having little or no expansion in the total wheat growing area during the same time (Fig. 2-1A). This major achievement is in part attributed to wheat yields that have more than tripled over time (Fig. 2-1B), allowing farmers to harvest more grain per single unit area. For example, in Europe the average wheat yield that in 1961 was 1.3 t ha^{-1} increased to 4.2 t ha^{-1} by 2014 (FAOSTAT 2017; www.fao.org).

Current predictions revealed that food production will need to increase by 2-3% every year to meet the increasing food demand of the projected growth of world population and ensure food security. However, the yields of wheat, rice and maize, the major cereal crops, have increased at less than half this rate (Hawkesford *et al.*, 2013). For example, wheat yield tripled in Asia between 1961 and 1986, from 0.76 t ha^{-1} to 2.29 t ha^{-1}, however, the increase was halved after that recording 3.1 t ha^{-1} in 2014 (Fig. 2-1B and FAOSTAT 2017; www.fao.org). Possible causes for this are the depletion of natural resources, such as water reservoirs and fertile lands due to climate change (Lobell *et al.*, 2011) and the use of crop varieties poorly adapted to unfavourable environments (Asfaw and Lipper, 2012).

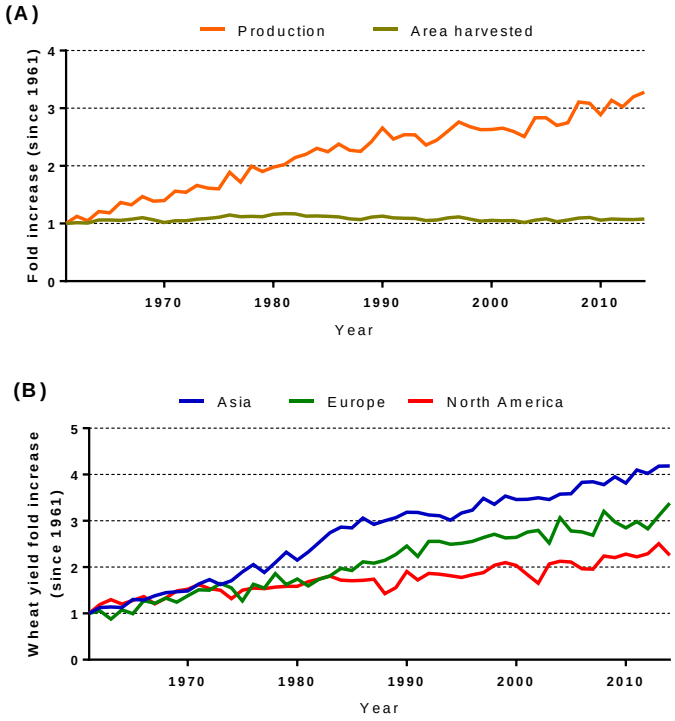


Fig. 2-1. (A) Fold-Increase of wheat production and wheat harvested area and **(B)** wheat yields fold-increase in Asia, Europe and North America area between 1961 and 2014. Data was sourced from FAOSTAT (2017) (www.fao.org).

2.1.2 Nitrogen (N) is required for high yields

One of the key components of the green revolution's success was the application of fertilizers in large amounts, especially fertilizers containing inorganic nitrogen (N). Plants need large amounts of N to develop and reproduce, in fact, N is the most abundant macronutrient obtained from the soil. Therefore, the maintenance of an optimal equilibrium between N demand and supply (N homeostasis) is crucial for plant growth and survival. Overall, N accounts for 1-5% of the total plant dry mass and it is necessary for the biosynthesis of cellular components such as nucleic acids and proteins, chlorophyll, co-enzymes, phytohormones and secondary metabolites (Marschner and Marschner, 2012). N deficiency is one of the most easily recognised nutrient stresses where the plant shows signs of chlorosis in older leaves and, in fact, growers use this as an indicator of when crops require fertilizer. The supply of N fertilizers to agricultural systems is essential to meet the yield potential of a crop. As an example of the importance of

fertilisation, at the end of the 1960s the newly introduced rice variety IR8, which yielded approximately 5 t ha^{-1} with no fertilizers, yielded 10 t ha^{-1} under optimal conditions in India (De Datta *et al.*, 1968). Therefore, it is not surprising that N fertilizer consumption has increased almost 10-fold worldwide since 1961, however, this was not paralleled by crop productivity due to a combination of other limitations, such as abiotic and biotic stresses, improper agricultural management practises and low performing germplasms (Fig. 2-2) (FAOSTAT 2017; www.fao.org).

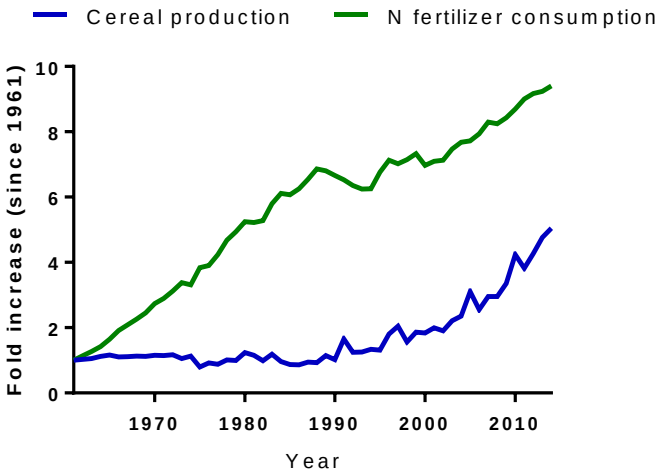


Fig. 2-2. Fold-increase in the global N fertilizer consumption and cereal production from 1961 to 2014. Data was sourced from FAOSTAT (2017) (www.fao.org).

The industrial production of N fertilizers relies on the Haber-Bosch process that requires extremely high temperatures and pressure and, therefore, is an energy intensive process. For example, in 1999, 90% of ammonia produced in the USA was used in the fertilizer industry and its production required 3% of the total natural gas production in the US, (Clark and Kelly, 2004), thus contributing to both atmospheric greenhouse gas emissions and the depletion of non-renewable energy sources. Agricultural systems heavily depend on the use of fertilizers and some fertilizer-producing countries, such as China, apply N fertilizer in excess as an "insurance policy" to ensure high yields (Kong *et al.*, 2013). However, excess nitrogen can be lost to the agricultural system via surface runoff, leaching into groundwater, soil

denitrification and volatilization (Ehdaie *et al.*, 2010; Butterbach-Bahl and Dannenmann, 2011) and can cause pollution and eutrophication of underground and superficial bodies of water, and emission of greenhouse gases, such as N_2O (Matson *et al.*, 1998). Despite being present in significant amounts in agricultural soils, only a small portion of the available N is taken up by the plant. Although, it has been reported that N uptake efficiency can be up to 80% for wheat under optimal conditions (Barraclough *et al.*, 2010). Raun and Johnson (1999) estimated that in 1996 the global cereal N recovery in grain from applied fertilizer (corrected for the amounts acquired from the soil and atmosphere) was as low as 33%, representing a $15.9 billion loss. Currently, plant breeders are investing large efforts and resources in improving N recovery from the soil employing the Nitrogen Use Efficiency (NUE) trait (McAllister *et al.*, 2012; Han *et al.*, 2015). NUE was defined by Moll *et al.* (1982) as grain dry mass at harvest per kilogram available N (from soil and fertilizer) and is divided into two components: (i) N-uptake efficiency (NUpE) that is the above-ground N per kilogram of supplied N, and (ii) N-utilization efficiency (NUtE) that is the grain dry mass per kilogram of above-ground plant N. Improving cereal NUE is a priority for a more sustainable agriculture as it will decrease the energy used for N fertilizer production and minimise pollution caused by N losses to the environment.

2.1.3 Drought largely affects crop yields worldwide

The adoption of irrigation in agricultural systems also represented a major drive in boosting crop productivity in the past. Water is an essential component for plants as it is required for turgor pressure, cell division and elongation, enzyme activity and energy metabolism. However, 44% of all agricultural systems worldwide are currently located within drylands (Millenium Ecosystem Assesment, 2005). These areas are prone to drought, a condition in which the amount of plant available water causes severe yield losses (Manivannan *et al.*, 2008). In fact, in these parts of the world, drought is the single most important factor affecting yields (Lambers *et al.*, 2008). The effects of drought on crop yield is determined by the intensity of the stress in itself and by the developmental stage of the plant experiencing the water

limitation. For cereals, drought events occurring during grain filling period (known as terminal drought) have a larger impact on yields as the stress greatly affects evapotranspiration, awn photosynthesis and translocation of assimilates which all contribute to grain yield (GY). In fact, it has been shown that prolonged mild drought during this crucial developmental stage can decrease GY by up to 92% in Indian wheat varieties (Dhanda and Sethi, 2002). Drought can also occur during vegetative stages, which lead to reduced growth and yield although to a smaller extent than during grain filling stages as shown by Daryanto *et al.* (2016) for wheat and maize. Climate change is expected to exacerbate drought events as vast areas of the world are predicted to become drier by the end of the century (Fig. 2-3A). A comparison of these drying regions with those currently employed for wheat production (Fig. 2-3B) clearly shows that many agriculture systems, especially in Europe and Australia, will face rainfall shortage in future decades.

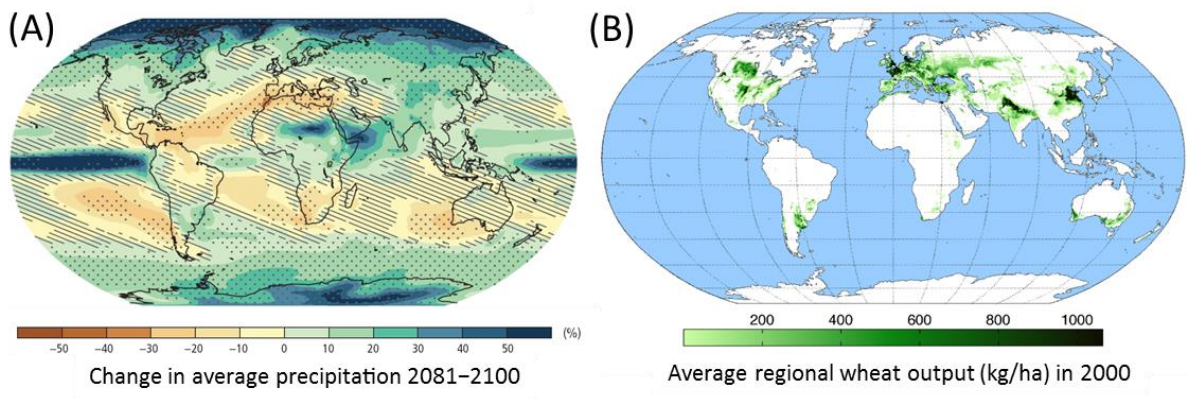


Fig. 2-3. (A) Patterns of global changes in precipitation between 2081 and 2100 estimated by coupled model intercomparison project phase 5 (CMIP5)(source: Kharin *et al.*, 2013). **(B)** Map of global wheat production as kg per hectare (source: Wikipedia based on international wheat production statistics from FAOSTAT; www.fao.org.).

In order to devise strategies for crop improvement that are targeted to the specific geographical location in which particular wheat varieties are grown, breeders at the International Maize and Wheat Improvement Centre (CIMMYT) divided the global wheat growing areas into 12 mega-environments that shares common climate features (Braun *et al.*, 2010). Drought occurs regularly in four of the mega-environments which includes regions of Australia, the Mediterranean basin, North and South America, South Africa,

parts of central Asia and northern China (Hodson and White, 2007). As drought reduces crop yields, a useful definition of 'drought tolerance' is the maintenance of yield under water limitation (Tricker *et al.*, 2016). Wheat GY under water limitation was defined by Passioura (1976) as:

$$GY = WU \times WUE \times HI$$

WU: was water uptake from the soil; WUE: water use efficiency, the biomass produced per unit of water transpired; HI: harvest index, the proportion between grain weight and above-ground biomass.

By dissecting GY into sub-components breeders can have a more targeted and rational approach in which yield can be increased by optimising each sub-component individually.

2.2 Plant nitrogen economy and its implications under drought stress

In this section the main processes involved in plant N homeostasis will be presented and, particularly, in relation to drought stress. Processes involved in N internal utilisation (e.g. recycling and translocation) will be presented in greater details as important topics of this dissertation. Whilst processes involved in N acquisition from inorganic sources will be briefly discussed to provide a context to the general picture of plant N economy.

2.2.1 Inorganic nitrogen uptake

Plants acquire mineral nutrients, including N, mainly from the soil through the root. In agricultural soils inorganic N, as nitrate (NO_3^-) or ammonium (NH_4^+), is the most abundant N source, whilst organic N forms are negligible. Nitrate is generally present in higher concentration (1-5 mM) than ammonium (20-200 µM) (Owen and Jones, 2001). Nitrate is also more mobile than ammonium and therefore more available to roots (Miller and Cramer, 2005).

Nitrate and ammonium (and all mineral nutrients in general) are present in the soil as solutes and their movement toward the soil/root interface for uptake requires adequate amounts of soil moisture. In fact,

the water transpired from above-ground tissue creates a below-ground water mass flow that transports solutes closer to the root. In addition, solutes can move in the soil solution by diffusion following a concentration gradient that is generally lower in the proximity of the root system (Lambers *et al.*, 2008). There are at least two components mediating nitrate uptake from the soil and redistribution among plant organs: the high affinity system (HATS), operating when nitrate concentration are in the µM range and the low affinity system (LATS) operating in the mM range (Tsay *et al.*, 2007; Léran *et al.*, 2014). The NRT2 (NITRATE TRANSPORTER 2) protein family constitutes the HATS. NRT2 proteins contain 12 trans-membrane domains and, to mediate nitrate uptake, they need to interact with a second component called NRT3.1 (NAR2) (Okamoto *et al.*, 2006; Orsel *et al.*, 2006). The NRT1/PTR (NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER) family, mediates the low-affinity nitrate uptake system (LATS). In both systems, the uptake functions as an active symport of nitrate anions with protons (H^+) (Forde, 2000). In contrast to nitrate transporters, ammonium transporters are less well characterised. It has been established that the ammonium HATS are saturable and function when ammonium levels are below 0.5 mM (Kronzucker *et al.*, 1996), whereas the ammonium LATS are non-saturable and operate at ammonium levels above 0.5 mM (Wang *et al.*, 1993; Kronzucker *et al.*, 1996; Rawat *et al.*, 1999). Proteins belonging to the ammonium transporter/methylammonium permease/rheus (AMT/MEP/Rh) family were showed to mediate the high-affinity transport of ammonium (Loqué and von Wirén, 2004). During drought plant transpiration rate is affected due to the abscisic acid (ABA)-mediated stomatal closure and the dryness of the soil also inhibits solute movement by diffusion. This potentially affects the plants' ability to acquire nutrients from the soil. Whilst a large number of studies have focused on the response of the N uptake systems in relation to different N concentration supply and genotypic variations (e.g. Okamoto *et al.*, 2003; Liu *et al.*, 2009; Garnett *et al.*, 2013; Melino *et al.*, 2015), the regulation of N uptake under drought stress is mostly uncharacterised. A study on apple trees (*Malus* x *domestica* Borkh) reported that a high-affinity nitrate transporter gene *MdNRT2.4* was up-regulated in roots under drought relative to re-watered samples (Bassett *et al.*, 2014). In wheat, Duan *et al.* (2016) studied the transcription of three nitrate HATs,

(*TaNRT2.1*, *TaNRT2.2* and *TaNRT2.3*), two nitrate LATs (*TaNRT1.1* and *TaNRT1.2*) and three ammonium transporters (*TaAMT1.1*, *TaAMT1.2* and *TaAMT2.1*) in two winter wheat genotypes contrasting in their N-uptake efficiency under drought and different N application rates. This study revealed a complex transcriptional response of the transporter under drought that varied between genotypes, N applications and plant developmental stages. For example, during vegetative stages the *TaNRT2.3* transcript was induced under drought in both genotypes within plants grown at low N conditions. Under optimal N application *TaNRT2.3* mRNA abundancy appeared to be reduced under drought and, interestingly, with much higher magnitude in the N-inefficient genotype.

2.2.2 N assimilation from inorganic sources

Once the inorganic forms of N are assimilated by the plant, they need to be incorporated into organic molecules in order to become available for the biosynthetic machinery. A key enzyme of this process is nitrate reductase (NR), which mediates the reduction of nitrate to nitrite. NR is localized in the cytoplasm and consists of two identical subunits, each carrying three co-factors that are covalently bound to specific protein domains. Generally, plants possess two NR gene paralogs that can be differentially expressed in roots and shoots (Crawford and Arst, 1993). The nitrite (NO_2^-) that is generated by the aforementioned reaction is transported to the chloroplasts or into plastids, depending on the tissue. Here, the nitrite reductase (NiR) mediates the reduction of NO_2^- into ammonium. Nitrite reductase is a ferredoxin-dependent enzyme. The electrons used by this co-enzyme are sourced from photosystem I in green tissues and from NADPH in root plastids (Bowsher *et al.*, 2007). The majority of higher plants have a single copy of this gene (Rastogi *et al.*, 1997; Kant *et al.*, 2011). The process of nitrate reduction appears to be tightly regulated with carbon (C) reduction mediated by photosynthesis, in fact, NR genes transcription is repressed by amino acids and especially C starvation (Lillo, 2008). Photosynthesis is one of the first processes to be affected under drought and, in fact, the activity of NR appears greatly reduced upon water stress. It has been proposed that NR activity is controlled by the photosynthetic rate via the ATP/AMP ratio that can alter the phosphorylation state of NR proteins (Kaiser and Spill, 1991; Redinbaugh

et al., 1994). Foyer and Valadier (1998) have shown that *NR* transcription is reduced under drought by 80% in maize leaves. Fresneau *et al.* (2007) also showed that NR activity was reduced under drought in durum wheat but in contrast to what was observed for other plants, rewatering the plants did not completely restore NR activity, highlighting that other levels of regulation exist.

The key reactions that ultimately incorporate reduced N from ammonium into organic molecules are carried out by two enzymes named glutamine synthetase (GS) and glutamate synthase (GOGAT; glutamine-2-oxoglutarate aminotransferase) (Fig. 2-4).

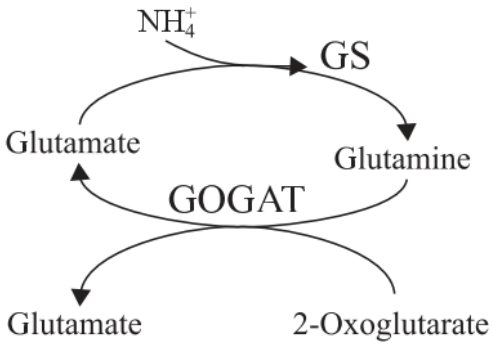


Fig. 2-4. Schematic representation of the GS/GOGAT cycle. Sourced from Bernard and Habash (2009)

The net reaction is:

$$\text{Glutamate} + \text{2-Oxoglutarate} + NH_4^+ \rightarrow 2\ \text{Glutamate} + H^+$$

These two enzymes play a crucial role in regulating the flux of N from ammonium into glutamine and glutamate. The amine ($-NH_2$) group of glutamine and glutamate can be readily used by several aminotransferases as a substrate for the synthesis of low-molecular-weight organic N compound such as amino acids, amides, peptides, amines and ureides. In plants, these organic N compounds serve not only as a precursors for the biosynthesis of high-molecular-weight compounds (i.e. proteins and nucleic acids) but also as N transport forms, transferring N from source to sink tissues, and as soluble N reserves (Forde and Lea, 2007; Marschner and Marschner, 2012). GS catalyses the ATP-dependent condensation of

ammonium (or ammonia) to the δ-carboxyl group of glutamate to form glutamine. In higher plants there are two isoforms of GS which have distinct functions, cytosolic GS1 and plastidic GS2 isoform (Zozaya-Hinchliffe *et al.*, 2005; Bernard *et al.*, 2008; Bernard and Habash, 2009) (Fig. 2-5).

Generally, GS2 is encoded by a single gene, whereas GS1 is encoded by two to five genes, depending on the plant species (Bernard and Habash, 2009). In hexaploid wheat Bernard and Habash (2009) identified and cloned ten GS sequences that share about 80% similarity at the protein level. Specifically, three genes encode the plastidic isoform (*TaGS2a*, *TaGS2b* and *TaGS2c*) since they possess a plastid transit peptide. Whereas the other genes (*TaGS1a*, *TaGS1b*, *TaGS1c*, *TaGSr1*, *TAGSr2*, *TaGSe1* and *TaGSe2*) lack this motif in their sequence and, therefore, were classified as cytosolic isoforms. GS1 is mainly involved in ammonium assimilation from inorganic nitrogen uptake and during N remobilisation and recycling (Bernard and Habash, 2009). In contrast, the major role of GS2 is to reassimilate ammonium generated through the photorespiratory nitrogen cycle that normally occurs during photosynthesis of C_3 plants (Leegood *et al.*, 1995) (Fig. 2-5).

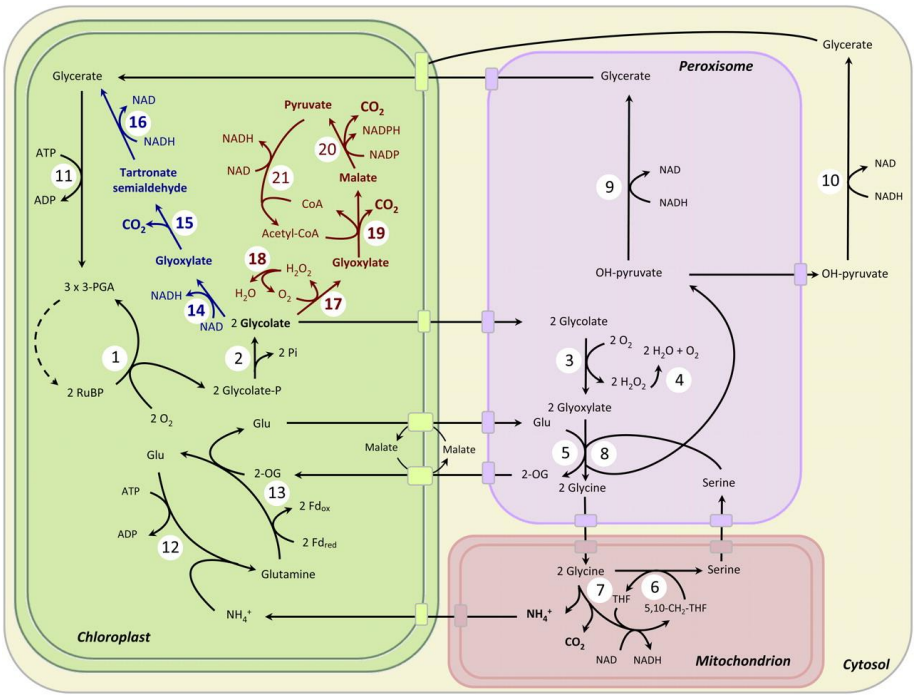


Figure 2-5. Schematic representation of photorespiration in C_3 plants. The photorespiratory pathway is shown in black arrows. 1, Rubisco; 2, 2-PG phosphatase; 3 and 17, glycolate oxidase (GO); 4 and 18, catalase; 5, Glu-glyoxylate aminotransferase; 6, Gly decarboxylase; 7, Ser hydroxymethyl transferase; 8, Ser-glyoxylate aminotransferase; 9 and 10, hydroxypyruvate reductase; 11, glycerate kinase; 12, Gln synthetase (GS); 13, Gln-oxoglutarate aminotransferase (GOGAT). Blue arrows show the bacterial glycolate pathway. 14, GlcDH; 15, glyoxylate carboligase; 16, tartronate semialdehyde reductase. Brown arrows show the intrachloroplastic glycolate oxidation pathway. 19, malate synthase; 20, NADP-malic enzyme; 21, pyruvate dehydrogenase. THF, Tetrahydrofolate; 5,10-CH2-THF, 5,10-methylenetetrahydrofolate. Figure sourced from Peterhanse and Maurino (2011).

The other key enzyme involved in ammonium assimilation is GOGAT. This enzyme catalyses the transfer of the amine ($-NH_2$) group of glutamine to 2-oxoglutarate, thus generating two molecules of glutamate (Forde and Lea, 2007). Plants possess two different types of GOGAT enzymes based on the electron donor and both of them are localised in plastids. The ferredoxin-dependent (Fd-GOGAT) enzyme preferably operates with higher activity in the chloroplast of photosynthetic tissues, especially in phloem companion cells in leaf veins (Masclaux-Daubresse *et al.*, 2006). The NADH-dependent (NADH-GOGAT) enzyme appears to be expressed in non-photosynthetic cells, in fact, it is the predominant form in roots (Tabuchi *et al.*, 2007). Two recent studies identified the bread wheat GOGAT homeologs and showed that NADH-GOGAT and Fd-GOGAT genes have 22 and 33 exons, respectively (Nigro *et al.*, 2013; Nigro *et al.*, 2014). A schematic representation of GS/GOGAT function in the different plant tissues is provided in Fig.2-6.

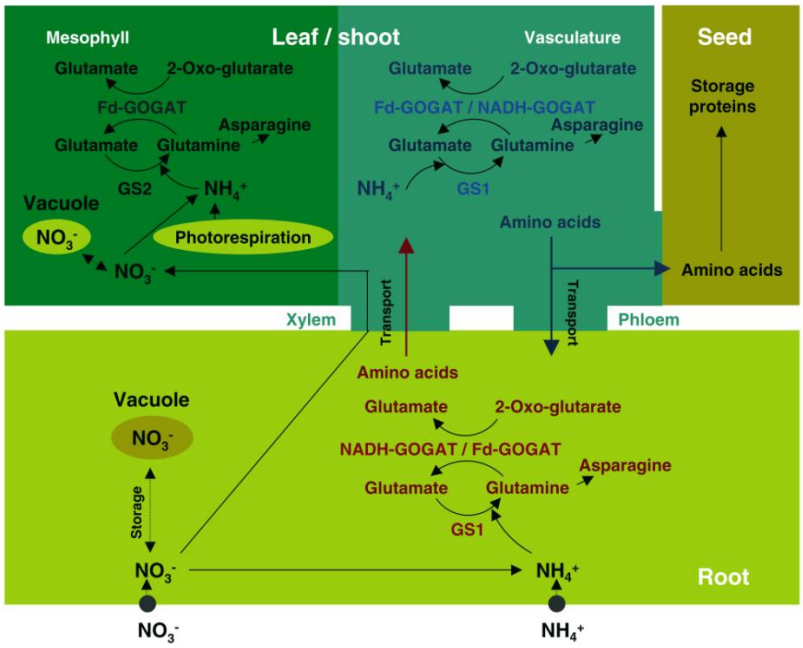


Fig. 2-6. Schematic representation of GS/GOGAT-mediated ammonium assimilation in plants. This figure was sourced from Suzuki and Knaff (2005). Nitrate and ammonium are taken up by root specific transporters. Nitrate can either be stored in vacuoles, transported in the shoot or converted into ammonium in the root. Here, both NADH- and Fd-GOGAT are present, although the former is the predominant, and they operate together with GS1. The vasculature operates similarly to root, although the predominant form of GOGAT is the Fd-dependent. In contrast, in leaf mesophyll, where ammonium is also produced by the photorespiration, the predominant forms are GS2 and Fd-GOGAT.

While GS and GOGAT catalyse irreversible reactions for the synthesis of glutamate, a third enzyme, glutamate dehydrogenase (GDH) mediates the reversible amination/deamination leading to glutamate synthesis or degradation (Forde and Lea, 2007). Masclaux-Daubresse *et al.* (2006) showed in young and old tobacco leaves that [^{15}N] ammonium assimilation to [^{15}N] glutamate occurred via GS/GOGAT cycle. Whilst GDH enzymatic activity appeared involved in deamination of [2-^{15}N] glutamate to provide [^{15}N] ammonium and 2-oxoglutarate.The GS/GOGAT cycle requires C skeletons and energy (1 ATP per glutamate produced) to operate. The C assimilated during photosynthesis can be used for the synthesis of ketoacids, such as 2-oxoglutarate, which is used by GS/GOGAT as backbone for ammonium. During water stress the inhibition of photosynthesis has a direct impact on GS/GOGAT substrates and attempts to study how GS is affected by water stress have been reported in cereals. For example, Singh and Ghosh

(2013) compared the GS activity of drought-tolerant and sensitive cultivars to controls under well-watered conditions and found that the drought-sensitive cultivar IR64 had a 50% reduction in total GS activity but the drought-tolerant rice cultivar Khitish could maintain total leaf GS activity (measures as mol s^{-1}) constant over 12 days of water stress. Dissection of individual GS1 and GS2 contributions to total GS activity revealed GS2 activity to be the most affected by drought in IR64 leaves. In fact, GS2 transcript and protein levels were reduced in IR64 compared to well-watered conditions from day eight of the drought treatment. Analysis of GS activity (measured as mol s^{-1} g^{-1} FW) in flag leaf of different wheat cultivars revealed that the drought-tolerant genotype Plainsman V could maintain GS activity to similar levels of well-watered plants, whilst the sensitive genotype Capelle Desprez had reduced GS activity (Nagy et al., 2013). However, protein levels of Plainsman V (tolerant) flag leaf under drought were maintained to similar levels of well-watered samples, whilst they were reduced in Capelle Desprez (sensitive), prompting to question whether the difference between genotypes under drought lies in their rate of protein degradation and/or synthesis rather than in GS activity in itself.

2.2.3 N remobilisation and recycling

In plants, ammonium is continuously generated not only through the primary uptake of inorganic N but also through several metabolic pathways, such as the turn-over of proteins and nucleic acids (DNA or RNA) and the photorespiration pathway (Leegood *et al.*, 1995). Intracellular concentrations of ammonium generally range from 1 to 30 mM, although the values can vary depending on the methodology used for its quantification (Miller *et al.*, 2001; Marschner and Marschner, 2012). Excessive accumulation of this compound may be toxic to the plant (Britto *et al.*, 2001; Miller *et al.*, 2001; Britto and Kronzucker, 2002). Given that plants do not preferably use ammonium for long-distance transport (Forde and Lea, 2007; Marschner and Marschner, 2012), recycling the released ammonium into low-molecular-weight organic N compounds is crucial for maintaining ammonium homeostasis in the plant. In particular, during leaf senescence several metabolic degradation pathways are up-regulated, leading to the production of large quantities of ammonium. Recycling this portion of N is of extreme importance for plants, especially for

those that use this source for grain filling. In fact, 60 to 92% of the total nitrogen in cereal grains comes from tissues undergoing leaf senescence (Barbottin *et al.*, 2005). Leaf senescence is a highly regulated physiological process that leads to cell death. Changes in cell metabolism and ultrastructure, such as chlorophyll degradation, a decrease of photosynthetic activity and disassembly of cellular integrity, are related to nutrient remobilisation and source-to-sink transport (Munné-Bosch and Alegre, 2004). The N remobilised from protein degradation, especially from chloroplasts, is the main source of nitrogen for grain filling in annual species and, for example, it accounts for about 80% of the total N in rice grains (Tabuchi *et al.*, 2007). A number of studies have shown that enzymes involved in ammonium assimilation play a key role in leaf senescence-induced nutrient remobilisation and have been linked with NUE and improved grain protein content (GPC). Studies on rice and maize, have suggested that GS1 is required for grain filling (for a review see Hirel *et al.*, 2007). Habash *et al.* (2007) highlighted that total GS activity in wheat flag leaves is positively correlated with the N content in the grain and stem. In addition, analyses of rice and maize mutants deficient in leaf cytosolic GS strongly support the notion that the GS1 isoforms are actively involved in the process of grain filling during leaf senescence (Tabuchi *et al.*, 2005; Martin *et al.*, 2006).

Drought induces premature leaf senescence and this plays a major role in enhancing the survival rate of several plant species experiencing water shortage. Under drought plants prioritise the maintenance of sink tissues, such as young and expanding leaves (vegetative stages) or grains (reproductive stages), by remobilising the nutrients accumulated in older tissues, allowing the plant to survive and reproduce (Munné-Bosch and Alegre, 2004). In addition, photorespiration which alone is the largest source of ammonium in C_3 plants, has been shown to increase in barley (*Hordeum vulgare* L.) under moderate drought stress (Wingler *et al.*, 1999). The major pathway involved in reasimmilating the ammonium derived from photorespiration is GS2 located in the plastids. Under drought, GS2 activity was reduced as aforementioned (Singh and Ghosh, 2013). Furthermore, the study of barley heterozygous mutants lacking GS2 showed increased levels of ammonium in leaves that was accompained by a decrease in amino

acids and protein (Häusler *et al.*, 1994). Interestingly Mattsson *et al.* (1997) showed the same barley mutants, with a 66% reduction in total GS activity, had ammonia (NH_3) gas emissions almost six times higher than wild type barley. More recently, Kumagai *et al.* (2011) reported that the rice cultivar Kasalath with reduced GS activity compared to the cultivar Akenohoshi produced more ammonia emissions with increasing light intensity, temperature and O_2 concentration. Although, a correlation of ammonia emission with increased photorespiration was observed in both genotypes. In addition, the results suggested that transpiration and stomatal conductance are not involved in regulating ammonia emissions from rice leaves. Taken together, these findings suggest that the regulation of ammonium assimilation and N remobilisation during drought may be important to maintain plant N homeostasis. Thus, it is not surprising that enzyme involved in N assimilation, such as GS, are differentially regulated under drought (see above). Under drought, degradation processes induced by premature senescence and increased rates of photorespiration can generate an excess of ammonium. Increased ammonium production under drought may lead to increased emissions of volatile ammonia leading to N losses from the shoot. However, it is still not known what extent this process affects the N economy of C_3 plants under drought.

2.3 Australian wheat production is co-limited by water and nitrogen

Spring wheat in Australia is grown in the so-called Australian Wheatbelt that includes the southern states of Western Australia, South Australia, Victoria, New South Wales and a small part of south Queensland. This area is large and includes different climate zones. Growing areas in Western Australia and South Australia are generally characterised by a Mediterranean-type climate where the majority of rainfalls occur during winter months (Australia: 1st June – 31st August). Winter also coincides with the beginning of the wheat growing season that generally starts with sowing in May/early June. As an example, a schematic representation of a mid-maturity Australian cv. Mace grown at Minnipa (South Australia) during the 2015 season is provided in Fig. 2-7.

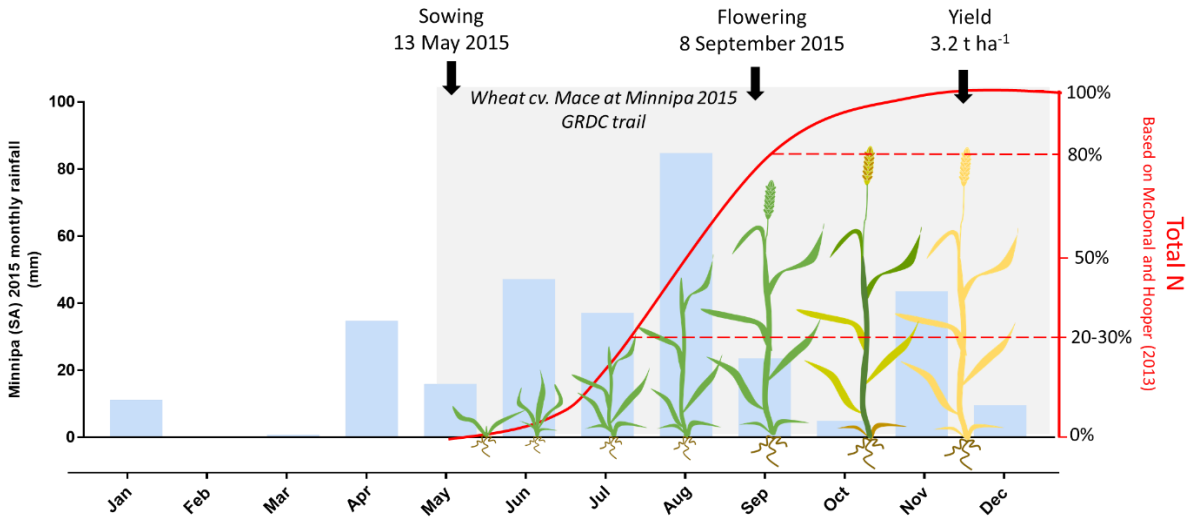


Fig. 2-7. Schematic representation of a mid-maturity spring wheat cv. Mace grown in Minnipa (South Australia, 32.84 °S, 135.15°E) during 2015. Rainfall data was sourced from Minnipa weather station (18195) available on Australian Bureau of Meteorology website (www.bom.gov.au). Wheat growth stages are depicted from germination to maturity. Sowing, flowering dates and yield data were sourced from James *et al.* (2016) (grdc.com.au). N pattern uptake is depicted with a red line and is based on general trends for wheat grown in southern Australia environments reported by McDonald and Hooper (2013) (grdc.com.au).

The vegetative phase of growth usually extends until September, when plants reach flowering stage, this period correspond with high rainfall events and, thus, water availability does not limit growth. From flowering stages until harvest, which usually occurs in late November – early December, plants undergo senescence and grain filling. During this time rain events start to become more sporadic and with the progression of the season, water stress and heat waves can occur at any time varying in duration and intensity (Izanloo *et al.*, 2008; Sadras *et al.*, 2016). Wheat production in southern Australia is under the threat of climate change (Hope *et al.*, 2015) and, in fact, during the past twenty-five years it has seen a significant reduction in the wheat yield potential, which is estimated from the difference of the actual yield and the predicted water-limited yield, also called the yield gap (www.yieldgapaustralia.com.au) (Fig. 2-8A). Surprisingly, Australian growers, despite the decrease of theoretical maximum attainable yield, have managed to maintain wheat yield mostly stable over the years (Fig. 2-8B) thanks to the adoption of improved practices and wheat germplasm. However, severe losses can occur and, for example, when in

2002 Australia suffered one of the most widespread drought events of the last decades (Horridge *et al.*, 2005), wheat yields were reduced by 57% compared to just the year before (see red asterisk in Fig. 2-7B).

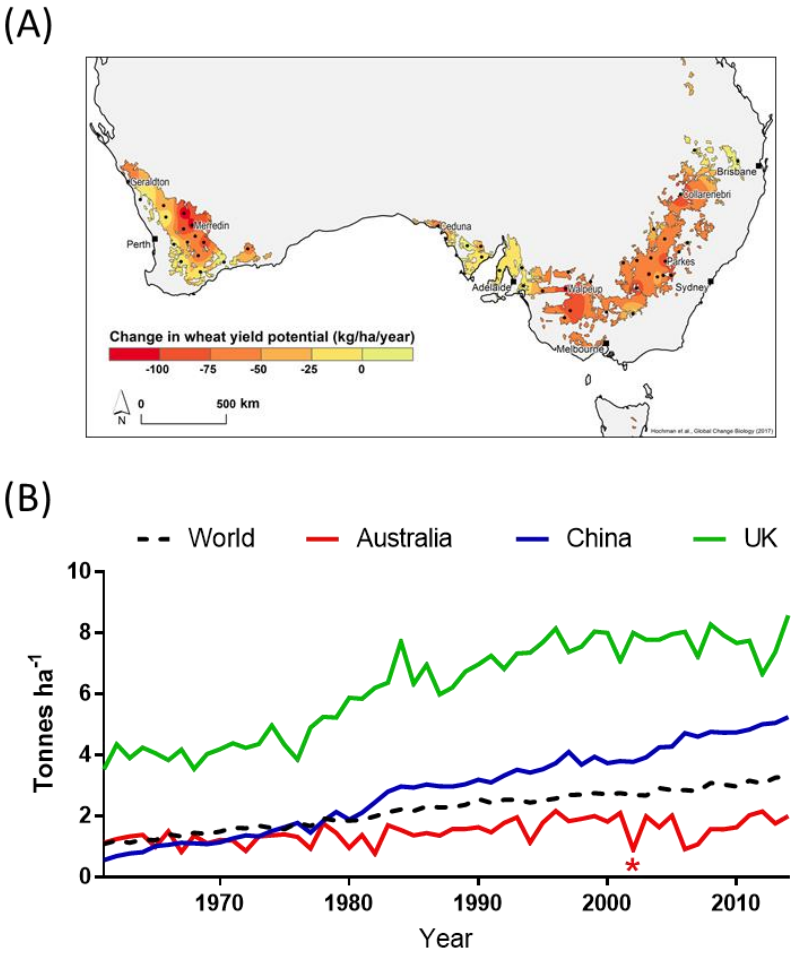


Fig. 2-8. (A) Distribution of annual change in wheat yield potential of Australian Wheatbelt estimated between 1990 and 2015 (Hochman *et al.*, 2017). **(B)** Wheat yields in the world, Australia, China and UK between 1961 and 2014 expressed as tonnes per hectare. Data was sourced from FAOSTAT (2017) (www.fao.org).

In agricultural systems the concomitant interaction of various rather than singular factors determines yields. In fact, N homeostasis and water use are closely associated as discussed in the previous section. Ample water resources improve plant transpiration causing an increase of N availability at the root surface for uptake (Tricker *et al.*, 2016). Higher transpiration also has positive effects on plant biomass formation

which in turn promotes root growth and access to more nutrients in the soil. Similarly, higher N availability produces more plant biomass and, thus, larger root system which allows better capture of water resources as well (Tricker *et al.*, 2016). While it is logical to assume that water is the main determinant affecting wheat yield in dry regions, multi environmental studies of drought-prone areas have shown that wheat yields in southern Australia (Sadras, 2005) and north-eastern Spain (Cossani *et al.*, 2010) increase as the degree of water and N co-limitation increases. In other words, yield is limited by the interaction of water and N rather than being severely limited by water or N alone. Interestingly, Sadras and Lawson (2013) analysed 13 Australian spring wheat cultivars released between 1958 to 2007 and demonstrated that breeding for yield over five decades has concomitantly improved yield and N uptake. The authors showed that yield and N uptake increased with same rates over time, while yield per unit nitrogen uptake has remained stable (Sadras and Lawson, 2013). Generally, wheat plants uptake the largest portion of N during vegetative stages (Lemaire *et al.*, 2007). In particular, wheat crops grown in southern Australia by the start of stem elongation phase have already taken up 20-30% of the total plant N and by flowering, which marks the end of vegetative growth, about 80% of the total N is taken up (Fig. 2-7) (McDonald and Hooper, 2013). As aforementioned, in southern Australia regions drought is more likely to occur in the field during terminal growth of spring wheat (Izanloo *et al.*, 2008). In this scenario, if post-flowering N uptake (Kichey *et al.*, 2007) is inhibited during terminal drought, this is not likely to have a large impact on plant total N, as the largest portion of it is usually acquired during pre-flowering stage (Fig.2-6). Therefore, terminal drought more likely affects processes of N recycling and loading to the grain rather than N uptake from the soil.

2.4 Metabolomics as a tool for exploring N metabolism under stress

The previous sections discussed drought impacts on N remobilisation and recycling in crop plants and it was suggested that this processes may be relevant for maintaining plant N homeostasis during terminal drought. In recent times, metabolomics has been explored as a valuable and versatile *-omics* technique

for characterising plant responses to different environmental changes. In particular, it represents a powerful tool to study how pathways involved in N remobilisation and recycling are affected under stress.

Metabolomics can be defined as the description of the metabolic state of a given biological system in response to environmental and genetic perturbations (Beckles and Roessner, 2012). The metabolome consists of a huge number of compounds presenting different chemicals and physical properties, such as molecular weight and size, polarity, volatility, solubility, stability and pk_a (Villas-Bôas *et al.*, 2007). A single plant can produce between 5,000 and 25,000 compounds at any given time (Roychoudhury *et al.*, 2011) and it has been estimated that approximately 200,000-1,000,000 compounds can be synthesised in the plant kingdom (Saito and Matsuda, 2010).

Metabolomics analyses are usually accomplished combining a separation technique with a detection device, such as a mass spectrometry (MS) analyser. The most popular techniques employed for separating complex metabolic matrices are gas chromatography (GC) and liquid chromatography (LC). GC-MS is usually employed for separating volatile compounds or compounds that can be volatilised after chemical derivatisation. This technique allows for a great power of metabolite separation and reproducibility although it is generally used to detect low-molecular-weight compounds (Beckles and Roessner, 2012). On the other hand, LC-MS presents two main advantages. Firstly, the compounds can be analysed without derivatisation and therefore the chemical structure is not altered. Secondly, high-molecular-weight compounds can be analysed with this technique, thus expanding the range of metabolites that can be analysed (Beckles and Roessner, 2012).

Generally, there are two approaches in metabolomics, untargeted and targeted. The former aims to detect and analyse as many metabolites as possible, whilst with the latter is limited to monitoring and detecting only a fixed pool of metabolites. The main advantages of the untargeted approach is that it is comprehensive and high-throughput, whereas the disadvantages are the high number of unknown metabolites and that most commonly is used as semi-quantitative method. On the other hand, with the

targeted approach, it is easier to perform absolute quantification of the metabolite concentration in the sample and this can be done with high accuracy and selectivity. In contrast, it is not comprehensive, expensive due to the use of standards and presents constraints for high-throughput analysis (Beckles and Roessner, 2012). Therefore, the best approach is dependent on the specific purpose of the analysis.

2.4.1 Versatility of metabolomics

Metabolomics is regarded as the most transversal among the *–omics* technologies, mostly because it does not require any previous genetic information, allowing the analysis of different organisms with minor modifications (Arbona *et al.*, 2013). In addition, the metabolic phenotype reflects the integration of gene transcription (mRNA), protein interaction and other regulatory pathways (Beckles and Roessner, 2012).

Plant species and varieties that are tolerant of abiotic stresses have been identified, however, genome sequences and resources (e.g. genetic stocks or transcriptomics datasets) of these organisms are still scarcely available. In fact, most of the research focussing on stress tolerance is carried out on model organisms, such as *Arabidopsis thaliana*, or old experimental cultivars, such as bread wheat Chinese Spring ,for which a first version of the reference genome became recently available (IWGSC RefSeq v1.0; www.wheatgenome.org). In particular, as reported by Sears and Miller (1985), this cultivar became the standard for research by chance, as it was initially used for producing wheat-rye hybrids more than a century ago. Sears and Miller (1985) described Chinese Spring as susceptible to most wheat diseases and insects and poorly adapted to the major wheat-growing regions around the world. A number of studies have recently reported Chinese Spring to be susceptible to abiotic stresses such as heat (Qin *et al.*, 2008), drought (Li *et al.*, 2012; Cheng *et al.*, 2015) and salt stress (Mott and Wang, 2007). Therefore, the disadvantage of studying such cultivars is that they may present a stress-intolerant response which might be unrelated to pathways responding to stress in tolerant genotypes (Pariasca-Tanaka *et al.*, 2009). Therefore, metabolomics represents an effective alternative for studying stress responses in organisms for which no or little genetic information is available.

There is a plethora of different types of molecular interactions and cross-talks that control plant adaptive responses (Schlüter *et al.*, 2013; Venu *et al.*, 2013). In this scenario, it is often difficult to identify candidate genes and determine whether their function is protective or merely part of a cascade of genes responding to the stress. Metabolomics to some extent can overcome this complexity, because it allows to analyse the "outcomes" of all the upstream interactions and regulatory processes and, actually, it can serve as a tool to identify pathways and candidate genes that may play a relevant role in mediating the tolerance. In fact, metabolomics is regarded closer to the plant phenotype than transcriptomics and proteomics alone (Beckles and Roessner, 2012; Arbona *et al.*, 2013).

The advantages of this technology can potentially be exploited in crop breeding programs, since it does not rely on genomic sequence availability and overcome the complexity of environmentally-controlled responses. Of particular interest is the statistical association of metabolic markers with genomic markers that leads to the development of metabolic quantitative trait loci (mQTLs). mQTLs have proven to be effective in predicting phenotypical properties before these features become apparent, which is a relevant aspect for developing biomarkers (Beckles and Roessner, 2012; Arbona *et al.*, 2013). GC-MS metabolic profiling of tomato introgression lines (ILs) obtained by crossing a wild species (*Solanum pennellii*) with a cultivated variety (*S. lycopersium* cv "M82") generated 889 mQTL of which almost half of them were linked with a yield-related traits (Schauer *et al.*, 2006). Another example is a study carried out by Meyer *et al.* (2007) in which the authors analysed the relationship between the metabolic profile and biomass of Arabidopsis recombinant inbred lines (RIL). They highlighted a positive correlation between enhanced biomass and N-containing compounds such as ornithine and polyamines. If confirmed in crops, this could potentially allow the selection of highly productive genotypes already at early stages of growth, with obvious advantages in terms of time and costs. Riedelsheimer *et al.* (2012) recently showed that by integrating metabolomics and genomics data of a large number of maize inbred lines it was possible to identify 26 metabolites in leaves of young plants that correlated to agronomic traits, such as plant height and dry matter yield. In wheat, Hill *et al.* (2015) used LC-MS to assess 179 double haploid lines of a cross

between an Australian bread wheat cultivars Excalibur (drought tolerant) and Kukri (drought sensitive). The authors annotated 197 known metabolites and found several phenolic acids to be positively correlated with grain yield; whilst a number of organic acids to be negatively correlated with grain yield. These studies demonstrate that the changes in metabolites accumulation are heritable and that the use of these powerful metabolomics techniques in combinations with genetic information may prove useful for plant breeding.

2.4.2. Exploring traditional aus-type rice for metabolites conferring drought tolerance

In a recent study the metabolic changes under water stress were assessed in four rice cultivars presenting contrasting tolerance to drought (Casartelli *et al.*, 2018; Appendix 1). Aus-type rice cultivars Dular and N22 were selected as tolerant varieties, whilst IR64 and IR74, belonging to the indica subgroup, were selected as sensitive varieties. Aus-type rice was originally bred and is still used in certain areas of India and Bangladesh characterised by dry environments (rain-fed) and soils poor in nutrients. As a results, aus-type rice cultivars are generally highly tolerant to environmental stresses, such as drought and heat, and have been recently used as interesting model for studying stress tolerance (Londo *et al.*, 2006). In fact, high valuable stress tolerant genes, such as submergence tolerance (OsSUB1A) and phosphorous (P)-starvation tolerance (OsPSTOL1) genes, have been identified in genotypes within the aus-type rice group (Xu *et al.*, 2006; Gamuyao *et al.*, 2012). Several studies also showed the superior tolerance of aus-type rice compared to modern irrigated varieties under heat stress (Li *et al.*, 2015; González-Schain *et al.*, 2016), drought stress (Henry *et al.*, 2011) and higher P uptake under P deficiency (Wissuwa and Ae, 2001).

In Casartelli *et al.* (2018) (Appendix 1) rice plants were grown in pots under well-watered conditions until 18 days after sowing (DAS) and then drought was applied by dry-down to half of the pots. At 32 DAS signs of leaf rolling appeared on leaves, at 33 DAS shoots samples were harvested and then analysed with an untargeted GC-MS approach and ion chromatography. A total of 328 metabolic features were identified, of which 102 were annotated. Principal component analysis (PCA) showed that the two first

principal components separate well the data according to treatment (PC1) and genotype (PC2) (Fig. 2-9). Specifically, aus-type rice Dular and N22 data plotted separately with respect to the indica verities IR64 and IR74, and that was particularly evident under drought, although already under control conditions the PCA analysis showed a certain degree of separation between tolerant and sensitive genotypes (Fig. 2-9).

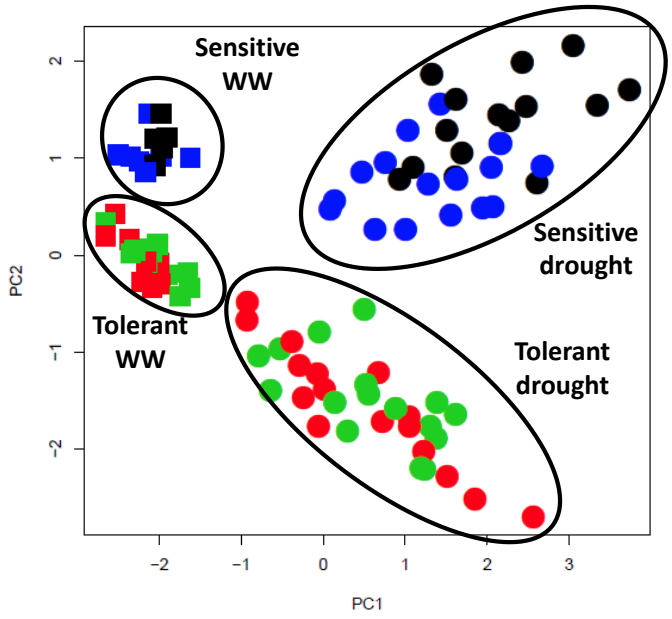


Fig. 2-9. PCA analysis metabolite profiling of shoot tissue. The first two principal components shoot samples were plotted against each other. Squares= well-watered; circles= drought; black=Dular; blue= N22; red= IR64; green= IR74; WW, well-watered.

Given the distinct metabolic composition of Dular and N22 with respect to IR64 and IR74, individual genotypes were combined into tolerant (aus-type) and sensitive (IR64 and IR74) groups. In order to identify tolerant-specific metabolites, a two-way ANOVA analysis was performed using treatment (well-watered, drought) and group of tolerance (tolerant, sensitive) as factors. Among the annotated metabolites, seventy-nine were identified as significantly changed because of the treatment, twenty-six because of the genotype and fifteen because of the interaction of the treatment and tolerance group. Thus, the latter included metabolites that accumulated under drought with significant differences between

genotypes. These metabolites were further classified depending on their magnitude of change (reported as log_2 of the fold-change between well-watered and drought) and associated with either the tolerant of sensitive group (Table 2-1).

Table 2-1. Metabolites with significant interaction (p<0,05) in the shoot for the two-way ANOVA analysis.

	Two-way ANOVA analysis			Log_2(Fold-change; D/WW)				Response
	Treatment	Genotype	Interaction	Dular	N22	IR64	IR74	type
Shoot								
Allantoin	1.29E-39	3.21E-10	3.77E-10	5.19	4.90	3.02	2.77	T
Proline	1.11E-24	1.95E-02	1.28E-02	4.91	4.13	2.22	3.45	T
Arginine	7.57E-38	1.94E-02	2.07E-02	4.11	3.57	3.07	2.51	T
Asparagine	6.44E-26	6.86E+00	1.95E-03	3.45	3.77	1.94	2.25	T
Ornithine	8.29E-23	1.32E-01	1.53E-03	2.97	2.37	1.73	1.76	T
Threonine	1.09E-33	6.93E-01	2.97E-02	2.71	2.56	1.76	1.76	T
Methionine	5.10E-24	8.60E-02	8.32E-04	2.71	2.03	1.26	1.33	T
Serine	6.48E-32	2.64E+01	1.11E-07	2.43	2.30	1.05	1.30	T
2-amino-butanoic acid (AABA)	4.79E-26	9.28E-02	4.36E-02	1.82	1.70	0.94	1.22	T
Uridine	1.77E-18	2.10E+02	3.08E-02	1.46	1.61	0.70	0.81	T
Chloride	1.29E-08	9.28E-01	5.39E-04	0.37	0.76	0.11	0.09	T
Raffinose	4.54E-20	6.85E-01	8.23E-04	0.61	0.98	1.69	2.50	S
Secologanin	1.57E-09	6.39E-11	1.35E-07	-0.41	0.29	1.26	1.20	S
Quinic acid	2.53E-03	7.05E-02	3.68E-03	0.25	-0.24	0.79	1.13	S
O-acetyl-serine (OAS)	3.83E-09	1.30E+01	2.92E-02	-0.26	-0.24	-1.03	-1.14	S

For each metabolite the log_2 value of the fold-change (drought/well-watered) is shown for individual genotypes. Metabolites were further classified based on their response type (T= tolerant or S=sensitive) and ordered according to their magnitude of change.

2.4.2.1 Drought tolerant rice cultivars accumulate N-rich metabolites in their shoots

Metabolites with significant interaction treatment × tolerant group are of particular interest as they can be indicative of pathways that are specifically involved in tolerance. Overall, all the metabolites that were associated with tolerant genotypes showed higher accumulation in Dular and N22 than in IR64 and IR74 and the majority were N-containing metabolites, with the exception of chloride. Amino acid was the most represented class of metabolites, including six proteinogenic amino acids (proline, arginine, asparagine, threonine, methionine and serine) and three intermediates (ornithine, AABA and OAS). Amino acids accumulation under drought stress is well documented in the literature (for example see Rai, 2002;

Planchet and Limami, 2015) and is still debated whether amino acid accumulation under stress is the result of increased protein degradation, decreased protein synthesis or enhanced synthesis and/or interconversion. However, it has been long established that several amino acids have a functional role in tolerance. Among all, proline is probably the most studied amino acid and it was demonstrated that it plays a role as osmolyte (Yoshiba *et al.*, 1995), regulator of redox potential (Hare and Cress, 1997), molecular chaperone (Verbruggen and Hermans, 2008; Szabados and Savoure, 2010), ROS scavenger (Mohanty and Matysik, 2001) and signalling molecule (Khedr *et al.*, 2003). It is therefore not surprising that proline accumulates with higher magnitude in Dular and N22 tolerant genotypes (Table 2-1). Vanrensburg *et al*. proposed back in 1993 the use of proline accumulation under drought as a criteria to screen tobacco lines for drought tolerance. Proline belongs to the amino acid family of glutamate. As discussed in the previous section, glutamate represents an important node of N homeostasis as the GS/GOGAT cycle assimilates the cytosolic ammonium, for which the net product is one molecule of glutamate per molecule of assimilated ammonium. In particular, it was highlighted how N recycling pathways are possibly inhibited under drought causing ammonium to accumulate leading to toxicity and N losses. Arginine and ornithine, also belonging to the glutamate family, were found associated with drought tolerant genotypes (Table 1). Arginine is the amino acid with the highest nitrogen to carbon ratio (N:C, 2:3) and, for example, it accounts for 50% of the N in the free amino acid pool of soybean and pea developing embryos (de Ruiter and Kollöffel, 1983; Micallef and Shelp, 1989). Degradation of arginine in the mitochondria, mediated by the arginase enzyme, liberates urea that is further degraded in the cytosol to produce CO_2 and two molecules of ammonium (Witte, 2011). Zonia *et al.* (1995) showed that arginase activity increases in Arabidopsis seeds during germination, suggesting that the stored N is remobilised via the arginine catabolic pathway during early stages of development. Therefore, accumulation of arginine under drought may indicate the prevention of further degradation thereby maintaining low cellular ammonium concentrations. Accumulating arginine may also serve as a temporary N reserve that can be stored or translocated and used for growth after recovery. Nonetheless, arginine may play a regulatory

role as it is involved in nitric oxide (NO)-signalling that regulates a plethora of cellular processes including responses to abiotic stresses, although the molecular mechanism is yet to be unveiled (Winter *et al.*, 2015). Ornithine, is a non-proteinogenic amino acid and a product of arginine catabolism mediated by arginase. However, ornithine is also synthesised from glutamate through cyclic and linear pathways involving several acetylated intermediates. Ornithine can be used as a substrate for arginine synthesis via the intermediates citrulline and argininosuccinate (for a review see Winter *et al.*, 2015). It is also worth mentioning that arginine and ornithine are key intermediates for polyamines biosynthesis, for which their involvement in plant stress-related processes has been extensively investigated although the results are sometimes controversial (polyamines are involved in complex interactions with metabolic pathways and hormones) (for a review see Gupta *et al.*, 2013). Putrescine, is formed from ornithine in a reaction mediated by ornithine decarboxylase (ODC). This polyamine is generally the most abundant in plants and it can constitute up to 1.2% of plant dry weight (Marschner and Marschner, 2012) indicating that putrescine may represent an important plant N pool. The synthesis of polyamines spermidine and spermine requires the use of S-adenosyl methionine (SAM) that is produced from the amino acid methionine. Methionine is derived from aspartate and from serine and its derivate O-acetyl serine (OAS) (Galili *et al.*, 2005). Interestingly, methionine, serine and OAS were also found to be more accumulated in rice tolerant genotypes under drought in Casartelli *et al.* (2018) (Appendix 1; Table 2-1). Taken together, the close connection of the amino acids associated with drought tolerant rice genotypes aforementioned and their tight link with ammonium and the GS/GOGAT cycle prompted us to speculate that changes of these amino acids under drought may be coordinated.

2.4.2.2 Allantoin accumulated with highest magnitude in aus-type rice under drought

Interestingly, the metabolite that presented the most significant interaction (treatment × tolerant group) and also higher accumulation in drought-tolerant genotypes was allantoin (Table 2-1). Allantoin is an intermediate of purine catabolism, a pathway involved in the degradation of purines that yields glyoxylate and four molecules of ammonium that can be recycled by GS/GOGAT (Zrenner *et al.*, 2006; Werner and

Witte, 2011). Purines are the most abundant N heterocycles in plants; besides being building blocks of nucleic acid (DNA and RNA), purines also participate in energy metabolism (ATP and GTP) and are precursors of certain alkaloids (theobromine, caffeine) coenzymes (NAD/NADP, FAD and Coenzyme A) and phytohormones (cytokinins) (reviewed in Smith and Atkins, 2002). Interestingly, the pyrimidine-analog uridine was found to be significant for the interaction and accumulated with higher magnitude in tolerant rice genotypes. However, pyrimidine catabolism yields only one molecule of ammonium, thus, representing a smaller pool of potential N for recycling as compared to purines.

2.5 Purine catabolism involvement in plant responses to stress

In the previous section allantoin was described as a drought responsive metabolite, accumulating in drought tolerant rice cultivars (Casartelli *et al*., 2018; Appendix 1). Allantoin belongs to the purine catabolic pathway, for which the enzymatic components have been recently identified and characterised in plants as summarised by Werner and Witte (2011) (Fig. 2-10).

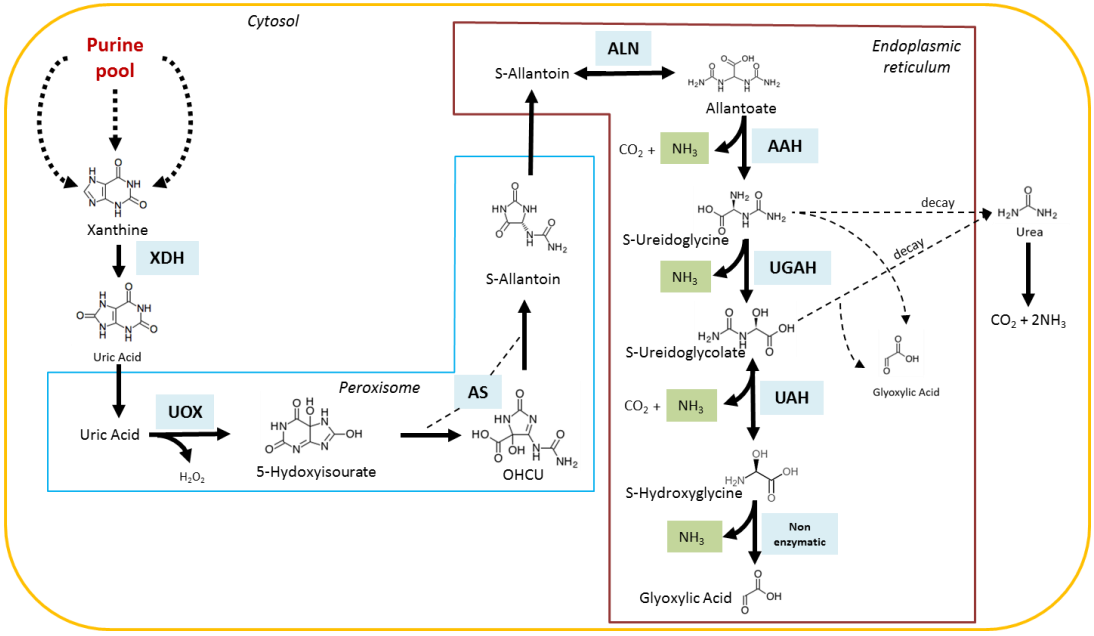


Fig. 2-10. Overview of the purine catabolic pathway. Xanthine is derived from three different substrates, namely hypoxanthine, xanthosine and guanosine, and is represented here by three dotted arrows. The degradation reactions are shown beginning with xanthine, the first intermediate of the pathway. Abbreviations of enzymes are enclosed within a blue box. The putative subcellular localisation of the enzymes and metabolites is indicated. Abbreviations include XDH, xanthine dehydrogenase; UOX, urate oxidase; AS, allantoin synthase; ALN, allantoinase; AAH, allantoate amidohydrolase; UGAH, ureidoglycine aminohydrolase; UAH, ureidoglycolate amidohydrolase; OHCU, 2-Oxo-4-hydroxy-4carbocy-5-ureido-imidazoline.

This pathway allows the complete degradation of metabolites with a purine ring. The end products, glyoxylate and ammonia, are recycled to synthesize new molecules, although the physiological role of this pathway in relation to N-remobilisation and nutritional aspect is still not fully characterised. The first enzyme of the pathway, xanthine dehydrogenase (XDH), catalyses the oxidation of xanthine to uric acid (urate) in the cytosol (Triplett *et al*., 1982; Werner and Witte, 2011). Uric acid then moves into the

peroxisomes, although the process is still unknown, where it is further degraded to allantoin by urate oxidase (UOX) and allantoin synthase (AS) (Hanks *et al.*, 1981b; Ramazzina *et al.*, 2006; Kim *et al.*, 2007; Lamberto *et al.*, 2010; Pessoa *et al.*, 2010). Allantoin is then transported into the endoplasmatic reticulum; the transport process is yet to be characterised (Werner and Witte, 2011). Here, allantoin is oxidised to allantoate by allantoinase (ALN) and then to ureidoglycine by allantoate amidohydrolase (AAH) (Yang and Han, 2004; Todd and Polacco, 2006; Werner *et al.*, 2008). The last two enzymatic reactions are mediated by ureidoglycine aminohydrolase (UGAH), that converts ureidoglycine to ureidoglycolate (Serventi *et al.*, 2010), and ureidoglycolate amidohydrolase (UAH), which metabolises ureidoglycolate to hydroxyglycine. Lastly, hydroxyglycine decays to glyoxylate by a non-enzymatic reaction (Werner *et al.*, 2010).

2.5.1 Ureides are differentially accumulated under stress

The relevance of allantoin as an N-storage and transport molecule was recognized in the early 1960s for maple trees (Reinbothe and Mothes, 1962). In addition, allantoin and allantoate, collectively referred to as ureides, are particularly important for ureidic legumes, since they represent the main N-transport form from the dinitrogen (N_2) fixed in nodules (Schubert, 1981). In these species, synthesis of allantoin and allantoate begins in the symbiosome, the physical barrier between the bacteroid and the infected plant cell. In this structure, N_2 is reduced to ammonia, ammonium or amino acids that are then converted to glutamine in the cytosol (Smith and Emerich, 1993; Day *et al.*, 2001; Lodwig *et al.*, 2003). The N present in glutamine is used for *de novo* purine synthesis in plastids or mitochondria of infected cells (Shelp *et al.*, 1983; Schubert, 1986; Atkins and Storer, 1997). The xanthine produced is then oxidised to allantoin and allantoate in uninfected cells (Hanks *et al.*, 1981a).

Recently, accumulation of allantoin was reported in several studies carried out on a range of plant species subjected to different stresses (Table 2-2). For example, higher levels of allantoin were observed in the desiccation-tolerant plant *Sporobolus stapfianus* compared to the desiccation-sensitive *Sporobolus*

pyramidalis (Oliver *et al.*, 2011). Allantoin proved to accumulate in the dry stage of the extremely drought-tolerant plant *Salaginella lepidophylla*, a lycophyte belonging to the class of resurrection plants (Yobi *et al.*, 2013). Metabolite profiling of 21 different rice cultivars under water stress highlighted a positive correlation between allantoin and drought tolerance in rice (Degenkolbe *et al.*, 2013), thus, supporting the findings reported in Casartelli *et al*., (2018) (Appendix 1). In addition, Wang *et al.* (2012) showed positive correlation of allantoin concentrations in rice grains and survival under drought and low temperature stress during vegetative growth. Bowne *et al.* (2011) studied drought adaptation of three South Australian bread wheat cultivars (RAC875, Kukri and Excalibur) using a GC-MS targeted approach. The authors subjected the plants to a cyclic drought treatment, which resembled field conditions of South Australian climate (Izanloo *et al.*, 2008). They observed that allantoin specifically accumulated in the drought-tolerant wheat cultivars RAC875 and Excalibur under mild to moderate drought conditions, whereas no significant accumulation was observed in the less tolerant variety Kukri.

Interestingly, allantoin was reported to be responsive also to nutrient stress. Allantoin accumulated under combined bicarbonate stress and zinc deficiency (Rose *et al.*, 2012) and under sulfur (S) deprivation (Nikiforova *et al.*, 2005). On the other hand, two studies reporting plant metabolic responses to nitrogen deficiency showed allantoin to be reduced with respect to high N treatments. (Amiour *et al.*, 2012; Coneva *et al.*, 2014). Specifically, Amiour *et al.* (2012) studied the metabolic responses of long-term N starvation of maize plants. The authors highlighted that allantoin was significantly reduced under N starvation in maize vegetative leaves at 55 days after sowing. Coneva *et al.* (2014) performed a hydroponic experiment in rice and showed that a number of nucleotide-related metabolites were differentially accumulated under N deficiency and when nutrient solution was changed from high N (HN) to low N (LN) and vice versa. Interestingly, allantoin concentration was more than seven-fold higher in rice roots grown under HN compared to LN. In addition, allantoin accumulated more than 8-fold in leaves when plant grown under HN were exposed to LN conditions for two hours. These studies therefore suggest that purine catabolism is responsive to N supply and that prolonged sub-optimal N applications generally lead to a reduction of

allantoin. This reduction may indicate enhanced remobilisation of purine nucleotides under N deficiency resulting in higher amounts of liberated ammonium that can be recycled into amino acids by GS/GOGAT.

Table 2-2. List of literature were allantoin levels were changed under stress.

Organism	Stress	Observation	Author	Year
Common bean	Drought	Ureide (allantoate) higher accumulation in leaf of sensitive genotypes	Coleto *et al.*	2014
Rice	Drought	Positive correlation between allantoin levels and drought tolerance	Degenkolbe *et al.*	2013
Selaginella lepidophylla	Drought	Accumulation during the dry stage	Yobi *et al.*	2013
Wheat	Drought	Accumulation under mild stress in drought tolerant genotype RAC875	Bowne e*t al.*	2011
Soybean	Drought	Accumulation in leaf of sensitive genotype and reduction in nodules of tolerant genotypes	Silvente *et al.*	2012
Sporobolus stapfianu	Drought	Accumulation in desiccation-tolerant species	Oliver et al.	2011
French bean	Drought	Allantoate accumulates in roots, shoots and leaves under drought	Alamillo *et al.*	2010
Arabidopsis	Stress signalling	Allantoin-mediated ABA activation triggers JA-stress related response via MYC2	Takagi *et al.*	2016
Arabidopsis	Stress signalling	Constitutive and exogenous accumulation of allantoin enhance the ABA-related stress signalling pathway	Watanabe *et al.*	2014
Arabidopsis	Abiotic stress	Ureide metabolism and accumulation contribute to stress response	Irani and Todd	2016
Arabidopsis	Cadmium toxicity	Allantoin and purine catabolism involved in response to Cd treatment	Nourimand and Todd	2016
Rice	Salt stress	Accumulation in shoot and root under salt and salt + ABA conditions	Wang et al.	2016
Rice	Salt stress	Higher accumulation in salt-tolerant genotypes than in salt-sensitive genotypes	Nam *et al.*	2015
Arabidopsis	Salt stress	Allantoin accumulation and ureide transporter UPS5 involved in tolerance	Lescano *et al.*	2016
Suaeda salsa	Salt stress	Accumulation in the aboveground part seedlings	Wu *et al.*	2012
Arabidopsis	Salt stress	Accumulation in 12 days seedlings exposed to high salt concentration (up to 30 hr treatment)	Kanani *et al.*	2010
Arabidopsis	Cold	Accumulation in aerial tissue	Kaplan *et al.*	2004
Rice	Nutrient constraint	Higher levels correlates with root solute leakage and increased conc. Of hydrogen peroxide due to Zn deficiency and high bicarbonate	Rose *et al.*	2015
Rice	Nutrient constraint	Differential accumulation in response to alteration of N concentrations	Coneva *et al.*	2014
Maize	Nutrient constraint	Lower amounts in leaf of long-term N starved plants during vegetative stage	Amiour *et al.*	2012
Arabidopsis	Nutrient constraint	Accumulation in seedlings in response to S deprivation	Nikiforova *et al.*	2005
Arabidopsis	Extended darkness	Enhanced allantoin and allantoate amounts in WT leaves upon exposing plants to dark treatment	Brychkova *et al.*	2008

2.5.2 Purine catabolic genes are important for plant growth and stress adaptation

The use of Arabidopsis lines carrying detrimental mutations in purine catabolic genes highlighted the importance of allantoin formation and degradation for plant growth and stress responses. Nakagawa *et al.* (2007) first attempted to silence the two Arabidopsis *XDH* genes, namely *AtXDH1* and *AtXDH2*, by RNA interference (RNAi). The authors reported that *xdh* mutants had reduced growth and fertility and an early onset of senescence. Interestingly, six week-old *xdh* mutant lines displayed significant reduction of chlorophyll and higher GS1 to GS2 ratio as compared to wild-type (WT) plants. GS1 is the cytosolic form of GS enzyme and it is generally induced during senescence while chloroplastic GS2 is usually reduced (Kawakami and Watanabe, 1988; Kamachi *et al.*, 1991). This suggests that purine catabolism and N assimilation pathways are indeed tightly linked. Watanabe *et al.* (2010) showed the importance of allantoin synthesis under drought, in fact, the authors reported that *xdh* RNAi mutants were highly sensitive to water stress and displayed lower biomass and chlorophyll with higher rates of cell death and hydrogen peroxide accumulation under drought as compared to WT plants. Brychkova *et al.* (2008a) employed XDH RNAi mutant lines and one XDH T-DNA insertion mutant line (KO) to show that *xdh* mutants failed to recover from extended dark stress, displaying higher levels of reactive oxygen species (ROS) and mortality rates as compared to WT plants. The authors of the aforementioned studies could revert or at least attenuate the *xdh* mutant phenotypes by supplementing uric acid (Nakagawa *et al.*, 2007; Watanabe *et al.*, 2010) or allantoin and allantoate (Brychkova *et al.*, 2008a) to the growth media.

The Arabidopsis *ALN* gene T-DNA insertion lines displayed a contrasting phenotype as compared to the *xdh* mutants. *aln* mutants constitutively accumulated allantoin, more than four times that of WT plants (Watanabe *et al.*, 2014) and performed better under dehydration, drought, salt, osmotic and heavy metal stress as compared to WT plants (Watanabe *et al.*, 2014; Irani and Todd, 2016; Lescano *et al.*, 2016; Nourimand and Todd, 2016). In particular, Irani and Todd (2016) reported that *aln* mutants subjected to drought stress had higher biomass and chlorophyll as compared to WT plants. These lines also maintained higher F_v/F_m ratio, indicating better energy allocation to photosystem II. The authors further

showed that accumulation of ROS under drought was reduced in *aln* mutants. However, a previous study reported that allantoin does not possess any *in-vitro* antioxidant activity (Wang *et al.*, 2012) prompting them to speculate that allantoin may play a role in stress sensing and regulation. Supporting this hypothesis, Watanabe *et al.* (2014) and Takagi *et al.* (2016) recently showed that increased levels of allantoin by loss of function mutants of *ALN* gene or by exogenous supply allantoin to the media enhanced abscisic acid (ABA) and jasmonic acid (JA) metabolism. ABA and JA are well known to play a key role in stress signalling in plants (Wasternack and Hause, 2013; Yoshida *et al.*, 2014).

Details of the possible molecular mechanisms regulating the levels of ureides under stress conditions have been characterised in the model organism *Arabidopsis thaliana*. Irani and Todd (2016) reported that the transcription of several Arabidopsis purine catabolic genes was strongly induced under drought stress. The authors showed that genes leading to allantoin synthesis, *AtXDH1*, *AtXDH2*, *AtUOX* and *AtAS*, were up-regulated after withholding water for five-ten days, whilst *AtALN* coding for the allantoin-degrading enzyme was only marginally increased after ten days of drought stress. Interestingly, *AtAAH*, operating just downstream *AtALN*, was up-regulated at ten days. Concurrently, allantoin accumulation under drought occurred from day five and peaked at day ten of the treatment. Irani and Todd (2016) also reported allantoin and allantoate accumulation in plants subjected to salt (NaCl) and osmotic stress (mannitol). *AtALN* transcripts were reduced after treatment with NaCl or mannitol. Reduced levels of *AtAAH* transcripts were also observed with NaCl treatment and *AtAS* transcripts were reduced in plants treated with NaCl. Similar results were observed by Lescano *et al.* (2016) who reported that allantoin concentration in Arabidopsis seedlings progressively increased with the concentration of NaCl supplied. In contrast, the concentration of allantoate progressively decreased with increasing NaCl supply. Transcriptomic analysis of Arabidopsis purine catabolic genes revealed that *AtUOX* and *AtAS* were significantly up-regulated, whilst *AtALN* was down-regulated in plants grown on 150 mM NaCl as compared to control conditions. Taken together, these results highlighted that accumulation of ureides (allantoin and allantoate) in Arabidopsis seedlings under stress is paralleled by changes in purine

catabolic genes transcripts. Increased transcription of genes coding for enzymes involved in the steps leading to allantoin formation (XDH, UOX, AS) and/or repression of genes coding for ureide degrading enzymes (ALN, AAH) may regulate allantoin and allantoate accumulation under abiotic stress.

Less information is available on the regulation of purine catabolic genes under N deficiency. Yang and Han (2004), who first isolated the *ALN* gene in Arabidopsis and in the leguminous tree black locust (*Robina pseudoacacia*), showed that both *AtALN* and *RpALN* transcripts levels were increased when plant were grown under N deficiency. Moreover, *AtALN* transcript levels were also increased when Arabidopsis seedlings were supplied with 10 mM allantoin as a sole source of N. Similar results for *AtALN* gene were reported by Werner *et al.* (2008), and they additionally showed that *AAH* transcript abundance was unaltered under both low N and with allantoin as a sole N source. Nonetheless, these studies proved that Arabidopsis can take up ureides supplied externally and that they can be used as a nitrogen source to grow. However, Desimone *et al.* (2002) demonstrated that despite providing equal amounts of N in the media as allantoin (5 mM) or ammonium nitrate (10 mM), Arabidopsis seedlings displayed better growth with the latter. This may indicate that inorganic N is more efficiently taken up and/or more efficiently utilised by Arabidopsis seedlings. Additional studies have demonstrated that the utilisation of xanthine and allantoin as a N source is dependent on the presence of functional purine catabolic enzymes. Severe delays in Arabidopsis growth were observed when *xdh* mutants were grown on xanthine as sole N source (Brychkova *et al.*, 2008b) and when either *aln* or *aah* mutants were supplied with allantoin as a sole N source (Desimone *et al.*, 2002; Todd and Polacco, 2006; Werner *et al.*, 2013). Interestingly, Werner *et al.* (2013) reported that the dry weight (DW) of Arabidopsis *uah* mutants grown on 10 mM allantoin was not significantly reduced compared to WT plants. The *UAH* gene encodes ureidoglycolate amidohydrolase, which is the last enzyme of the pathway that oxidises ureidoglycolate to hydroxylglycine (Fig. 2-9). In the absence of a functional UAH enzyme, ureidoglycolate decays *in-vitro* to glyoxylate and urea with a half-life of a few hours (Gravenmade *et al.*, 1970). This indicates that *uah* mutants can remobilise the N from allantoin via urea catabolism. To reinforce this assumption, Arabidopsis lines carrying T-DNA insertions

in both *UAH* and *URE* (urease gene mediating the oxidation of urea to ammonium and CO_2) showed significant DW reduction with respect of WT plants when grown on 10 mM allantoin (Werner *et al.*, 2013). However, this non-enzymatic decay is not likely to play a major role *in-vivo* as ureidoglycolate is expected to be metabolised by UAH at a faster rate (e.g. Km OsUAH = 188 µM and Km GmUAH = 40 µM; Werner *et al.*, 2013).

2.5.3 Peculiarities of ureidic legumes

As mentioned above, ureidic legumes, such as common bean and soybean, synthesise the ureides allantoin and allantoate as the product of N_2 fixation in root nodules; ureides are then translocated through the vasculature system (Schubert, 1981). As a result, ureide concentration in these species is particularly high when compared to plants that employ amino acids as main form to transport organic N, such as Arabidopsis or cereals. For example, a concentration of 5 µmol g^{-1}DW of allantoin and up to 30 µmol g^{-1}DW of allantoate was reported in common bean leaves (Coleto *et al.*, 2014) and 25 µmol g^{-1}DW of allantoin was found in nodulated common bean seed coats (Pèlissier *et al.,* 2007); whistl a concentration of more than 24 µmol g^{-1}DW of leaf ureides was recently reported in soybean (Gil-Quintana *et al.*, 2013). In contrast, the concentrations found in non-ureidic plants are considerably less: e.g. in Arabidopsis Watanabe *et al.* (2014) found a concentration of just 0.005 µmol g^{-1}DW and 0.002 µmol g^{-1}DW for allantoin and allantoate, respectively. In cereals, Montalbini *et al.* (1992) reported wheat grains and leaves containing approximately 1 µmol g^{-1}FW and 0.250 µmol g^{-1}FW allantoin, respectively. Interestingly, corn silk, that finds use in dermatology and cosmetology, contains up to 1.8 µmol g^{-1}DW of allantoin (Maksimović *et al.*, 2004).

Early studies had already shown that allantoin and allantoate accumulate under drought in ureide transporting legumes (Sinclair and Serraj, 1995; deSilva *et al.*, 1996) and this was more recently demonstrated in non-nodulated legumes too (King and Purcell, 2005; Alamillo *et al.*, 2010). In particular, allantoate, rather than allantoin, significantly accumulates in the leaves and stems of common bean

(*Phaseolus vulgaris*) subjected to drought stress (Alamillo *et al.*, 2010; Coleto *et al.*, 2014). Interestingly, the accumulation of allantoate under water stress appears to be correlated with an up-regulation of *PvALN* gene expression and a coordinated inhibition of the allantoate degradation activity of PvAAH, which has been suggested to be mediated at a post-translational level in common bean (Alamillo *et al.*, 2010) and soybean (Gil-Quintana *et al.*, 2013). In addition Alamillo *et al.* (2010) reported that exogenous application of ABA induces the expression of the *PvALN* gene in both stems and leaves of common bean. It is interesting to compare this finding with Watanabe *et al.* (2014) who reported that the accumulation of allantoin, as a result of *AtALN* loss of function, induced ABA metabolism leading to ABA accumulation. Another relevant difference of the drought response of ureidic legumes from non-ureidic plants was highlighted by Coleto *et al.* (2014). The authors studied ureide accumulation in four genotypes of common bean with different levels of tolerance to drought. Whilst they observed a general accumulation of ureides, especially allantoate, in all the genotypes, they reported that these metabolites were more concentrated in the tissues of drought-sensitive genotypes. This finding is in contrast to those showing that allantoin specifically accumulates in tolerant plants or cultivars of non-ureidic plants including wheat, rice and resurrection plants (Bowne *et al.*, 2011; Oliver *et al.*, 2011; Degenkolbe *et al.*, 2013; Yobi *et al.*, 2013; Casartelli *et al.*, 2018). This may suggest that purine catabolism is regulated differently under drought stress by ureidic legumes. Investigating the relation between the expression of the purine catabolism genes, their enzymatic activity and ureides accumulation under drought stress may shed light on the role that allantoin has in stress adaptation in these plants.

Another aspect that highlights the difference between ureidic legumes and other plants is the analysis of the ureide permease (UPS) transporter family. This gene family was firstly characterised by Desimone *et al.* (2002) who identified five *UPS* genes (AtUPS1-5) in *Arabidopsis thaliana* genome and characterised the *AtUPS1* orthologous. *AtUPS1* and *AtUPS3* may derive from a recent event of tandem duplication and *AtUPS2* and *AtUPS4* possibly followed the same fate. Pélissier *et al.* (2004) identified a putative allantoin transporter in common bean, the transporter PvUPS1 has been described to transport allantoin and the

precursors xanthine and uric acid. In addition, *PvUPS1* was found expressed in several plant organs including the phloem and xylem and that *PvUPS1* expression levels were in accordance with allantoin concentration among different tissues. These findings support the hypothesis that PvUPS1 is involved in source-sink transport of allantoin (Pélissier *et al.*, 2004; Pèlissier and Tegeder, 2007). Interestingly, PvUPS1 was localised in the phloem and xylem parenchyma which is in agreement with the high levels of allantoin in the vasculature tissue leading to speculation that PvUPS1 is also involved in the storage of allantoin in this tissue, as suggested by Pèlissier and Tegeder (2007). Nevertheless, in Arabidopsis the proteins AtUPS1 and AtUPS2 have been characterised and, interestingly, Schmidt *et al.* (2004) reported that AtUPS1 and AtUPS2 affinity for allantoin (~75 μM and ~26 μM; tested in Xenopus oocytes system) were lower than their affinity for the pyrimidine base uracil (~5.9 μM and ~6.2 μM). The fact that Arabidopsis mainly uses amino acids for long-distance transport may explain this evolutionary divergence in transporter function. However, Desimone *et al.* (2002) showed that Arabidopsis seedling can grow on a media containing allantoin as a sole source of N and that *AtUPS1* is up-regulated under N-starvation. This suggests that *AtUPS1* might play a role when primary sources of N are limited. In addition, *AtUPS2* was shown to be expressed in cells of the root stele (Schmidt *et al.*, 2004) and, thus, might be involved in long distance translocation of allantoin. *AtUPS5,* another member of the Arabidopsis UPS family, has two different splice variants: AtUPS5l (long variant) and AtUPS5s (short variant) (Schmidt *et al.*, 2006). AtUPS5l has similarities with AtUPS1 and AtUPS2 in terms of protein structure and recognition of purine degradation products and pyrimidines. However, AtUPS5l had highest affinity for xanthine (~6.8 μM), whilst similar affinity for allantoin and uracil (~35.6 μM and ~38.5 μM, respectively) (Schmidt *et al.*, 2006).On the other hand, AtUPS5s was not able to transport these substrates into yeast cells and, therefore, it may have a different function (Schmidt *et al.*, 2006). Long-distance transport of allantoin in Arabidopsis has not been investigated. Determining whether allantoin under abiotic and N stress is stored or translocated in non-ureidic plants might shed light on its physiological role. Lescano *et al.* (2016) recently showed that a transgenic Arabidopsis line expressing the GUS reporter gene under the control

of the *AtUPS5* promoter had increased GUS staining intensity after seven days of 150 mM NaCl stress. Moreover, the authors showed *ups5* T-DNA insertion lines to be more susceptible to salt stress. This suggests that *UPS* genes are also involved in regulating ureide levels under salt stress, thus, represent interesting targets for further investigations with other abiotic stresses.

2.6 Physiological role of allantoin in wheat under stress and knowledge gaps

The accumulation of the ureides allantoin and allantoate in plants experiencing water stress has been documented in several studies, although the specifics of its function still remains to be tested in plants of agricultural importance. Accumulating high levels of allantoin under stress may prevent the formation of ammonium that would not be efficiently re-assimilated by GS/GOGAT (Nagy *et al.*, 2013; Singh and Ghosh, 2013). The high N:C ratio (1:1) of allantoin suggests that it is an ideal and energy effective form of N to be translocated or stored during drought stress. Therefore, from a nutrition perspective, controlling the metabolic flux through the purine catabolic pathway could represent a strategy that plants adopt to better tolerate drought. On the other hand, evidence showing reduced amounts of allantoin in plants under N deficiency (Amiour *et al.*, 2012; Coneva *et al.*, 2014) indicates that plants may compensate for low external N supply by enhancing allantoin catabolism providing an internal source of recycled N to maintain growth. The fact that products of nucleotide catabolism may have a role during plant nutrient stress is not surprising. Purines (and pyrimidines) are abundant in DNA and different RNA species (mRNA, t-RNA, rRNA, and sRNA), as well as in many other cellular components, such as nucleotide sugars, ATP, or GTP. The complete oxidation of these compounds liberates C, N and P that can be recycled into central metabolism. For example, Hewitt *et al.* (2005) reported that Arabidopsis plants reduced their RNA content by 90% after 14 days of P starvation. Therefore, fine tuning of purine catabolism under abiotic stress and nitrogen deficiency may also improve the carbon and nitrogen homeostasis of the plant.

Allantoin has been quantified by metabolomics techniques in cereals such as wheat, rice and maize, however, no information is currently available in relation to the function of purine catabolic genes in such organisms. Especially for wheat, gene sequence information is still highly fragmented. In common bean and Arabidopsis accumulation of allantoin and allantoate under drought was described to be coordinated with transcriptional changes of specific purine catabolic genes. However, this still needs to be described in cereals in which allantoin was reported to accumulate in response to drought. In addition, certain studies have shown allantoin levels to be reduced in plants under N deficiency, whilst other authors reported that transcription of *ALN* was increased under low N conditions. To date, there are no comprehensive reports describing the coordination of purine catabolic genes transcription and allantoin levels under N deficiency. While the nutritional value of allantoin has been long established in ureidic legumes, its involvement in N economy of non-ureidic plants still remains to be elucidated in details. This is especially important in crop plants such as bread wheat, in which N remobilisation to the grain is an important determinant of grain quality. Furthermore, allantoin was shown to differentially accumulate in genotypes with contrasting tolerance to abiotic stress (Bowne *et al.,* 2011; Oliver *et al.,* 2011; Degenkolbe *et al.,* 2013; Casartelli *et al.*, 2018) , however, no study has yet assessed whether allelic variants of purine catabolic genes are associated with this tolerance.

2.7 Thesis scope and outline

The objective of this study was to investigate the role of purine catabolism and allantoin degradation in Australian bread wheat genotypes and their contribution to the N economy under N deficiency or water deficit. The specific aims of this project were:

a) To identify, annotate and compare purine catabolic genes in the bread wheat hexaploid genome to those in other grasses;

b) To quantify and compare accumulation of allantoin in two distinct Australian bread wheat genotypes subjected to drought and N deficiency and simultaneously assess whether the response of purine catabolic genes to these stresses is coordinated with allantoin accumulation.

c) To estimate the contribution of allantoin to N recycling and remobilisation in wheat;

d) To identify bread wheat accessions with allelic variation in purine catabolic genes that can be used for future genetic and functional studies.

Chapter 3: The allantoin pathway in bread wheat is responsive to abiotic stresses and nitrogen supply

Statement of Authorship

Title of Paper	The allantoin pathway in bread wheat is responsive to abiotic stresses and nitrogen supply
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Manuscript prepared in accordance with the guidelines for the Journal Plant, Cell and Environment

Principal Author

Name of Principal Author (Candidate)	Alberto Casartelli		
Contribution to the Paper	Executed the study, performed bioinformatics and molecular analysis, interpreted data, wrote manuscript		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	

Co-Author Contributions

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

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr. Vanesa J. Melino		
Contribution to the Paper	Supervised development of work, help in data interpretation and manuscript evaluation and editing		
Signature		Date	26/10/17

Name of Co-Author	Dr. Ute Baumann		
Contribution to the Paper	Provided help and critical comments of bioinformatics analyses		
Signature		Date	10/10/17

Name of Co-Author	Prof. Ute Roessner		
Contribution to the Paper	Provided help on allantoin quantification and evaluation of the manuscript		
Signature		Date	10/10/17

Name of Co-Author	Dr. Melissa R. Pitman		
Contribution to the Paper	Performed *in silico* protein modelling, provided help with its interpretation and comments on the results section		
Signature		Date	11/10/17

Name of Co-Author	Prof. Stuart M. Pitson		
Contribution to the Paper	Performed *in silico* protein modelling and provided help with its interpretation		
Signature		Date	10/10/17

Name of Co-Author	Dr. Matteo Riboni		
Contribution to the Paper	Performed qRT-PCR analysis, provided support with data interpretation and evaluation of the results section		
Signature		Date	26/10/2017

Name of Co-Author	Nirupama S. Jayasinghe		
Contribution to the Paper	Provided technical support on allantoin quantification and evaluation of the methods section		
Signature		Date	16/10/2017

Name of Co-Author	Dr. Radoslaw Suchecki		
Contribution to the Paper	Performed RNA-seq data mapping and provided help in its interpretation		
Signature		Date	26/10/2017

Name of Co-Author	Dr. Mamoru Okamoto		
Contribution to the Paper	Supervised development of work, help in data interpretation and manuscript evaluation and editing		
Signature		Date	26/10/2017

Name of Co-Author	Dr. Sigrid Heuer		
Contribution to the Paper	Supervised development of work, help in data interpretation and manuscript evaluation and editing		
Signature		Date	10/10/17

Title

The allantoin pathway in bread wheat is responsive to abiotic stresses and nitrogen supply.

Alberto Casartelli[1], Vanessa Melino[1*], Ute Baumann[1], Ute Roessner[2], Melissa R. Pitman[3], Stuart M. Pitson[3,4], Matteo Riboni[1], Nirupama S. Jayasinghe[2], Radoslaw Suchecki[1], Mamoru Okamoto[1], Sigrid Heuer[1, 5]

Affiliations:

[1] School of Agriculture Food and Wine, The University of Adelaide, Urrbrae, SA 5064, Australia

[2]Metabolomics Australia, The University of Melbourne, Parkville, VIC 310, Australia

[3] Centre for Cancer Biology, University of South Australia and SA Pathology, Frome Road, Adelaide, SA 5000, Australia

[4] School of Biological Sciences, University of Adelaide, SA 5000, Australia

[5] Plant Biology and Crop Science department, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UK

*Corresponding author: PMB1 Glen Osmond 5064 SA, Australia; vanessa.melino@adelaide.edu.au

Abstract

Purines are nitrogen (N) rich heterocyclic compounds present in DNA, RNA and other cellular components. In plants, the purine catabolic pathway liberates four molecules of ammonium that may be recycled by N assimilating enzymes (GS/GOGAT cycle). Intermediates of the pathway (the ureides allantoin and allantoate) are the major compounds used for long-distance transport of N in nodulated tropical legumes, and have been associated with abiotic stress responses in a range of plant species.

Here we describe identification of purine catabolic genes in hexaploid bread wheat (*Triticum aestivum* L.) and their syntenic position in other grass genomes, highlighting a duplication event of *TaXDH* genes (TaXHD1, TaXHD2) unique to bread wheat and its ancestors. However, detailed protein analyses suggest that only one TaXDH2 homeolog is functional. Wheat purine catabolic genes are differentially expressed in flag leaves and grain suggesting that ureides contribute to grain N content. In agreement with this, allantoin levels of field-grown wheat grains increased in response to N-fertilizer supply. Furthermore, the purine pathway was transcriptionally responsive to dehydration and cold stress. Data presented here provides the groundwork to explore the possible dual role of ureides in wheat, both in N recycling and protection against abiotic stresses.

Keyword index: nitrogen homeostasis; purine catabolism; nucleotide metabolism; allantoinase; xanthine dehydrogenase; allantoin

Introduction

Nitrogen (N) is a macro nutrient that supports plant growth and reproduction. To ensure high yields in agricultural production systems, N can be supplied in the form of either inorganic or organic fertiliser. Once N is taken up and assimilated into organic compounds, significant amounts of ammonium (NH_4^+) are produced as by-product of plant metabolism, mainly during carbon oxidation. NH_4^+ is readily recycled by the glutamine synthetase-glutamate synthase (GS/GOGAT) cycle into low-molecular-weight organic N compounds, such as amino acids and amides (Lea & Miflin, 2010). In cereals undergoing natural senescence, recycling and translocation of these compounds from vegetative tissues to the grain accounts for 60% to 92% of grain N (Barbottin *et al.*, 2005). In bread wheat, grain N content is an important quality trait since grain protein content (GPC) is a major determinant of wheat flour baking quality which is valued economically (Bogard *et al.*, 2010; Hawkesford, 2017).

An important portion of recycled N comes from the breakdown of macromolecules such as proteins and nucleic acids (DNA and RNA). Purines are the most widely distributed N heterocyclic compound, as present in DNA and RNA, and are also a component of important cellular functions, such as energy metabolism and secondary metabolism (Werner & Witte, 2011). Purines may be derived from either the breakdown of nucleic acids or *de novo* synthesis. Plants undergo the complete breakdown of the purine ring via a catabolic pathway enabling the recycling of both carbon and N (Fig. 3-1). Overall, the oxidation of 1 molecule of xanthine to 1 molecule of glyoxylate liberates three molecules of CO_2 and four molecules of NH_4^+, which are likely to be reassimilated by the GS/GOGAT cycle into amino acids, although reassimilation of purine derived NH_4^+ is yet to be experimentally verified. The pathway starts with the conversion of xanthine to urate catalysed by xanthine dehydrogenase (XDH) that is localised in the cytosol (Triplett *et al.*, 1982; Werner & Witte, 2011). Urate is further processed in the peroxisome by urate oxidase (UOX) producing 5-hydroxyisourate (5-HIU), 5-HIU is then converted to allantoin via the 2-oxo-4-hydroxy-4-carboxy-5-ureido-imidazoline (OHCU) intermediate by allantoin synthase (AS) (Hanks *et al.*, 1981; Ramazzina *et al.*, 2006; Kim *et al.*, 2007; Lamberto *et al.*, 2010; Pessoa *et al.*, 2010). Allantoin is metabolised in the endoplasmatic reticulum (ER) to allantoate by allantoinase (ALN) and then to ureidoglycine by allantoate amidohydrolase (AAH) (Yang & Han, 2004; Todd & Polacco, 2006; Werner *et al.*, 2008). The last two enzymatic steps are catalysed by ureidoglycine aminohydrolase (UGAH), which converts ureidoglycine to ureidoglycolate (Serventi *et al.*, 2010). Ureidoglycolate amidohydrolase (UAH) converts ureidoglycolate to hydroxyglycine and, lastly, hydroxyglycine decays to glyoxylate by a non-enzymatic reaction (Werner *et al.*, 2010).

In addition to its housekeeping role in N recycling, the purine catabolic pathway has an important function in warm-season N_2-fixing legumes (Schubert, 1986). Allantoin and allantoate (also known as ureides) are the main products of atmospheric N_2 fixation in root nodules which are then translocated to the shoot, mainly via the xylem, accounting for up to the 90% of xylem N (Herridge *et al.*, 1978; Pate *et al.*, 1980). In common bean (*Phaseolus vulgaris* L.), allantoin is actively translocated to sink leaves and pods during

reproductive stages, with the greatest accumulation in stem and seed coats (Pèlissier & Tegeder, 2007). There is some evidence to support the role of ureides in N remobilization in other non-leguminous plants, for instance, in ageing leaves of *Arabidopsis thaliana*, ureide accumulation was observed concurrently with changes in the expression levels of purine catabolic genes (Brychkova *et al.*, 2008).

In recent years, purine catabolism has gained further attention because of its involvement in plant stress responses. Early studies had already shown that allantoin and allantoate accumulate under drought in ureide transporting legumes (Sinclair & Serraj, 1995; de Silva *et al.*, 1996) and this was more recently demonstrated in non-nodulated legumes too (King & Purcell, 2005; Alamillo *et al.*, 2010) and in a wide range of plant species under different abiotic stress (Bowne *et al.*, 2011; Oliver *et al.*, 2011; Wu *et al.*, 2012; Degenkolbe *et al.*, 2013; Yobi *et al.*, 2013; Nam *et al.*, 2015; Wang *et al.*, 2016, Casartelli *et al*., 2018). Studies in Arabidopsis further showed that the genes involved in purine catabolism are differentially regulated under drought, salt, osmotic, heavy metal stress and prolonged dark exposure (Brychkova *et al.*, 2008; Irani & Todd, 2016; Lescano *et al.*, 2016; Nourimand & Todd, 2016). In addition, *AtALN* loss-of-function mutants (*Ataln*) revealed that constitutive accumulation of allantoin enhanced the abscisic acid (ABA) and jasmonic acid (JA) pathways and increased tolerance of dehydration, drought, salt and osmotic stress in Arabidopsis (Watanabe *et al.*, 2014b; Takagi *et al.*, 2016). In support of this, RNAi-mediated inhibition of *AtXDH1* led to disrupted growth with early onset of senescence and hypersensitivity to drought (Nakagawa *et al.*, 2007; Watanabe *et al.*, 2010; Watanabe *et al.*, 2014a). Despite the large number of studies available, purine catabolism remained largely uncharacterised in cereals. However, early reports indicate an involvement of XDH in the response of wheat to wheat leaf rust (Montalbini, 1992), and an accumulation of allantoin has been observed in both wheat and barley subjected to drought stress (Bowne *et al.*, 2011; Chmielewska *et al.*, 2016). The aim of the present study was to identify wheat orthologs of the purine catabolic genes and to investigate their role during plant development and abiotic stresses. The gene loci were identified based on the reference genome of the cultivar Chinese Spring and their chromosomal location was experimentally verified. Gene expression was examined in above-ground

tissues during various stages of development, as well as under dehydration and after a 24-hour cold treatment using the Australian bread wheat genotype RAC875. In addition, a comparative study using three different wheat genotypes showed altered levels of grain allantoin under N-limited growth conditions.

Materials and Methods

Plant material and experimental conditions

Triticum aestivum L. genotype RAC875 (RAC655/3/Sr21/4*LANCE//4*BAYONET), a drought tolerant breeding line developed by Roseworthy Agricultural Campus (Izanloo *et al.*, 2008), was used for gene expression analysis. Seeds were germinated in black plastic pots (11 x 11 x 14 cm) containing 1.5 kg of potting mixture (river sand and coco-peat; for detail see Melino *et al.*, 2015) with complete nutrients, except N which was supplemented as basal application of urea at a rate of 150 mg N kg^{-1} soil. A second application of urea (50 mg N kg^{-1} soil) was performed at stem elongation by soil drenching. Plants were grown in a controlled environment chamber until maturity with day/night cycle of 12h/12h at a flux density at canopy level of 300 $mmolm^{-2}s^{-1}$, 20°C/15°C and 82% average humidity. Tissue samples were collected from six individual plant replicates and immediately snap-frozen in liquid nitrogen and stored at -80 °C until further use. Growth conditions for RAC875 under dehydration and cold treatments are described in Kovalchuk *et al.* (2013). Briefly, two-week old hydroponically grown seedlings were subjected to dehydration stress by withholding growth solutions and leaf samples were harvested one and two hours after induction of the stress. For the cold treatment, three-week old soil grown seedlings were transferred to a cold cabinet kept at a constant temperature of 4°C and leaf samples were harvested after four and 24 hours. Mature grains harvested from plants of three Australian wheat genotypes, RAC875, Spitfire (Drysdale/Kukri) and Kukri (76ECN44/76ECN36//MADDEN/6*RAC177) grown under high N (98 kg N ha^{-1}; HN) and low N (8 kg N ha^{-1}; LN) conditions in a field trial conducted in Tarlee (South Australia; 34°16'46.1"S 138°46'17.3"E) in 2013 were used for allantoin analysis.

Identification of the purine catabolic genes in wheat and other grass genomes

The purine catabolic genes of *Oryza sativa*, *Zea mays*, *Sorghum bicolor* and *Brachypodium distachyon* were identified with a BLASTP search using *Arabidopsis thaliana* protein sequences as a query in Phytozome (https://phytozome.jgi.doe.gov). *B. distachyon* protein sequences were then used for a TBLASTN search on the barley WGS Morex Assembly version 3 (International Barley Genome Sequencing, 2012) and the barley predicted protein sequences were used for a TBLASTN search against the Chinese Spring TGACv1 genome assembly (http://pre.plants.ensembl.org/Triticum_aestivum/Info/Index) (Clavijo *et al.*, 2017). In cases where full-length sequences were not found in Chinese Spring TGACv1, we searched the Chinese Spring IWGSC chromosome survey sequence version 3 (IWGSC, 2014 – http://www.wheatgenome.org) and the sequences were merged with TGACv1 sequences using Geneious version 10.0.2 (www.geneious.com) (See Table S3-1 for details). The same approach was used for the wheat ancestor genomes (Table S3-4). Genomic regions were viewed, exon/intron structure annotated and CDS and protein sequences retrieved using Artemis (Rutherford *et al.*, 2000) (http://www.sanger.ac.uk/science/tools/artemis). Gene structures (exon/introns) and coding sequences (CDS) were illustrated using the software Geneious (10.0.2.). Each purine catabolic gene was assigned with the name of the Arabidopsis orthologous according to Watanabe *et al.* (2014b). For wheat, the sub-genome localisation (e.g., Chr 1AL) was additionally included in the gene name to allow distinction of the three homeologous sequences (Table S3-1). Alignments of the predicted protein sequences were performed using ClustalW in Geneious 10.0.2 with the default settings (IUB cost matrix, 15 gap open cost and 6.66 gap extend cost).

Synteny analysis

Synteny analysis were conducted based on available genome sequences of *Sorghum bicolor*, *Oryza sativa* and *Brachypodium distachyon* using the Phytozome 11 'Ancestry' tool (phytozome.jgi.doe.gov). Generally, five genes upstream and downstream of the respective purine catabolic pathway gene were included in the orthology analysis. These genes in bread wheat were identified by TBLASTN searches of Brachypodium genes against the TGACv1 wheat genome assembly (http://pre.plants.ensembl.org/Triticum_aestivum/Info/Index). Genes were considered syntenic if they were present on the same TGACv1 scaffold where wheat purine catabolic genes were annotated.

Wheat nulli-tetrasomic lines

To verify the chromosomal location of the wheat purine catabolic gene Chinese Spring nulli-tetrasomic (NT) lines were used in which individual wheat chromosomes were replaced (nullisomic) by an extra pair (tetrasomic) of their homeologs (Sears *et al.*, 1954). For this analysis, homeolog-specific primers were designed using Geneious 10.0.2 and tested for secondary structures using Primer3Plus (http://primer3plus.com/) (Table S3-1). PCR was performed using 150 ng of genomic DNA of the NT lines as template in a final volume of 12 µl consisting of 1X ThermoPol® Buffer (BioLabs, USA), 5% DMSO, 0.2 mM dNTPs, 0.25 µM forward and reverse primers, 1X Taq DNA polymerase (in-house), and milliQ water. Amplification cycles consisted of 1 cycle = 5 min 95 °C, 35 cycles = 20 s 95 °C, 30 s 60-63 °C, 2 min 72 °C, 1 cycle = 1 min 72 °C. PCR products were analysed in 1% (w/v) agarose gels stained with SYBR safe DNA Gel Stain (Thermo Fisher Scientific, USA) and visualized using a ChemiScope 2850 fluorescence and chemiluminescence imaging system (Clinx Science Instruments, China).

RNA-Seq reads mapping

Publically available bread wheat (cv. Chinese Spring) RNA-Seq data was used for generating the per-base expression values. This data set covers 15 duplicated samples corresponding to five organs (root,

leaf, stem, spike, grain), each at three developmental stages (http://urgi.versailles.inra.fr/files/RNASeqWheat/) (Choulet *et al*., 2014). The RNA-Seq reads were quality-, adapter- and length-trimmed using Trimmomatic (Bolger et al., 2014), version 0.30 with a custom list of adapter sequences and the following settings: 'ILLUMINACLIP:adapters.fa:1:6:6 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:6 MINLEN:60'. The reads were aligned to the scaffolds (version 3) from the Chinese Spring whole genome assembly (version 0.4) using STAR (version 2.5.1b) (Dobin et al., 2013), with the following settings: --outFilterMultimapScoreRange 0; --outFilterMultimapNmax 5; --outFilterMismatchNoverLmax 0 --outFilterMatchNminOverLread 1; --outSJfilterOverhangMin 35,20,20,20; --outSJfilterCountTotalMin 10,3,3,3; --outSJfilterCountUniqueMin 5,1,1,1 --alignEndsType EndToEnd; --alignSoftClipAtReferenceEnds No; --alignIntronMax 10000; --alignMatesGapMax 10000. The remaining settings were left at their defaults. The resulting BAM files were merged using samtools merge (version 1.2) (Li *et al.*, 2009). The manually annotated coordinates of the genes define the regions for which the depth of the aligned reads was computed from the aligned BAM file using the depth module of samtools.

RNA preparation

Preparation of cDNA libraries from the wheat genotype RAC875 were prepared from snap-frozen samples of youngest fully emerged leaves harvested at 24 days after sowing (Z22) and from flag leaves, stems, spikelets and developing grain harvested from the main stem at 17 days after anthesis (Z79). Samples were ground to a fine powder using a 2010 Geno/Grinder® (SPEX SamplePrep, USA) and total RNA was extracted from frozen samples with a phenol and guanidine thiocyanate buffer according to Chomczynski (1993). To extract RNA from developing grains which have a high polysaccharide content, an extraction buffer (1% (w/v) sarcosyl, 150 mM NaCl, pH 9) and a guanidine hydrochloride-based buffer for purification according to Singh *et al.* (2003) was employed. Genomic DNA was removed using the TURBO DNA-

free™ Kit (Ambion®, Thermo Fisher Scientific, USA) following the manufacturers' instructions. High RNA quality was confirmed in a 2% (w/v) agarose gel visualized under UV light and RNA concentrations were quantified using a ND-1000 spectrophotometer (NanoDrop Technologies, USA). 1 µg of RNA was used for cDNA synthesis using the SuperScript® III kit (Thermo Fisher Scientific, USA) as per manufacturers' instructions. cDNA quality was verified by PCR amplification of the actin gene (Table S3-5). cDNA libraries from dehydration and cold stressed plants used for this study were kindly provided by Kovalchuk *et al.* (2013).

Primer design and quantitative reverse transcription PCR (qRT-PCR) analysis

Quantitative real–time PCR was performed with KAPA SYBR® Fast qPCR kit Master Mix, and amplification was real–time monitored on a QuantStudio™ 6 Flex Real-Time PCR System (Applied Biosystems, USA). Gene-specific primers targeted to amplify all three homeologs simultaneously were designed with AlleleID® software (Premiere Biosoft) (Suppl. Table S3-5) and the specificity of each pair was verified by melting curve analysis and sequencing of the products. Change in gene expression were calculated using equations from Hellemans *et al*. 2007 for Fig. 3-6 and qBASE+ software (Hellemans *et al*. 2007) for Fig. 3-7. qBASE+ software reports gene expression as normalised relative quantity (NRQ):

$$NRQ = \frac{E_{goi}^{\Delta Ct,goi}}{\sqrt[f]{\prod_{o}^{f} E_{ref_o}^{\Delta Ct,ref_o}}}$$

E: efficiency
ΔCt: delta-Ct
Ct: cycle threshold
goi: gene of interest
ref: reference

Metabolite extraction and derivatisation

Metabolites were extracted based on a modified version of the method used by Bowne *et al*. (2011). Briefly, 35 mg of mature grains ground to a fine powder under liquid N_2 were extracted with 500 µl of

100% (v/v) methanol containing $^{13}C_5$-^{15}N-Valine (0.01 mg ml^{-1}) and incubated at 70°C for 15 min. Samples were centrifuged for 15 min at 15000 rpm and the supernatant transferred to a new tube. 500 µl of milli-Q-water was added to the remaining pellet, vortexed for 30 sec and centrifuged again as described above. The supernatant was combined with the previous one and 300 µl 100% (v/v) chloroform was added to the combined supernatant, vortexed thoroughly and centrifuged for 15 min at 15000 rpm. 50 µl of the upper polar phase were aliquoted in a glass insert and dried under vacuum (RVC 2-33 CD plus, John Morris Scientific, Pty Ltd, Melbourne, Australia). Derivatisation was carried out with 10 µl of Methoxyamine hydrochloride (30 mg ml^{-1} Pyridine) at 45°C for 2 h before adding 20 µl of N-methyl-N-(tert-butyldimethylsilyl)trifluoroacetamide (MTBSTFA) with 1% (w/v) trimethyl chlorosilane (TMCS) and incubation at 45°C for 45 min.

GC-MS detection and data analysis

The GC-MS system used for allantoin detection comprised of a Gerstel 2.5.2 autosampler, a 7890A Agilent gas chromatograph and a 5975C Agilent quadrupole mass spectrometer (Agilent, USA). The mass spectrometer was tuned according to the manufacturers' recommendations using tris-(perfluorobutyl)-amine (CF43). 1 µl of derivatised sample was injected onto the GC column (30 m Agilent J & W VF-5MS column with 0.25 µm film thickness and 0.25 mm internal diameter and a 10 m Integra guard column). The injection temperature was set to 250°C, the MS transfer line to 280°C and the ion source to 230°C. Helium was used as the carrier gas at flow rate of 1 ml min^{-1}. Initial oven temperature was set to 40°C with a ramp of 25°C per min to 325°C. The detector was set to single ion monitoring at 398 m/z for allantoin as previously described by Pavitt *et al.* (2002) and 191 m/z for valine as the internal standard. Normalisation of the data was performed by dividing the area response value of allantoin by the area response value of valine and by the dry weight of the extracted sample.

Protein homology modelling

Modelling of the three-dimensional structure of the predicted wheat proteins TaXDH1 and TaXDH2 was performed using the Bovine Xanthine Dehydrogenase crystal structure as a template (3UNI) (Ishikita *et al.*, 2012) on the ICM-Pro platform (version 3.8-6 Molsoft, San Diego, CA, USA) (Abagyan & Totrov, 1994). Homology models were developed with template ligands and all loops and side-chains refined by Cartesian energy minimisation. Receptor maps were calculated for xanthine and NADH binding pockets and substrates were docked using full-atom-biased Monte Carlo sampling of the ligand and receptor binding pocket side chains in internal coordinates (Abagyan & Totrov, 1994). Models were compared to both bovine XDH (3UNI) and rat XDH (4YSW) (Nishino *et al.*, 2015) enzymes by superpositioning using the ICM-Pro suite.

Statistical analysis

To assess differences in gene expression among wheat tissues and between control and stress (dehydration and cold) conditions a one-way ANOVA analysis followed by Tukey's multiple comparison test was performed. Differences in allantoin grain content between high N and low N were assessed by Student's *t*-test. Statistical analyses were performed using GraphPad Prism version 7.00 for Windows (GraphPad Software, La Jolla California USA, www.graphpad.com).

Results

Identification of the wheat purine catabolic genes and their chromosomal localization

Based on comparative sequence analyses, a total of 24 wheat genes that are orthologous to the purine catabolic genes of Arabidopsis and rice were identified (Fig. 3-2A, Table S3-1). For the *TaXDH* genes, two allelic variants are present in wheat located on the long arm of chromosome group 1 (*TaXDH1*-

1AL/TaXDH1-1BL/TaXDH1-1DL) and short arm of chromosome group 6 (*TaXDH2-6AS/TaXDH2-6BS/TaXDH2-6DS*), respectively. The *TaXDH1* and *TaXDH2* paralogs share 89.6% to 90.5% sequence identity within the coding region (CDS) and the *TaXDH1* homeologs showed a higher CDS identity (98.6 to 98.8 %) than the *TaXDH2* homeologs (93.5-95.3 %), Table S3-2. the two paralogs of *TaAS* genes were identified located on the short arm (*TaAS-4AS*) and long arm (*TaAS-4DL*) of chromosome 4. A third copy was identified but the genomic location was not determined in the TGACv1 scaffold (U, unknown). However, based on the IWGSC css assembly this gene is part of a contig (4BL_6969139) associated with the long arm of chromosome 4 (*TaAS-4BL*). The well-known homology of chromosome 4AS with 4BL and 4DL and 4AL with 4BS and 4DS (Mickelson-Young *et al.*, 1995) may explain the distribution of *TaAS* genes observed here. The *TaUOX* genes (*TaUOX-3AL/TaUOX-3BL/TaUOX-3DL*) were localised to the long arm of chromosome 3 and the *TaALN* genes (*TaALN-2AL/TaALN-2BL/TaALN2DL)* were localised to the long arm of chromosome 2. The *TaAAH* genes (*TaAAH-7AL/TaAAH-7BL/TaAAH-7DL*) are located on the long arm of chromosome 7 and were the least conserved among the analysed genes with only 95-92.9% sequence identity within the CDS (Table S3-2). The *TaUGAH* genes (*TaUGAH-2AS/TaUGAH-2BS/TaUGAH-2DS*) are located on the short arm of chromosome group 2 and the last genes of the pathway, *TaUAH* (*TaUAH-5AS/TaUAH-5BS/TaUAH-5DS*), were localised to the short arm of chromosome 5 (Fig. 3-2A).

In order to confirm the chromosomal localization of the wheat genes, nulli-tetrasomic (NT) lines of the wheat cultivar Chinese Spring were used (Fig. 3-2B; Table S3-1) (Sears *et al.*, 1954; see M&M for details). For this analysis, homeolog-specific primers were designed and used for PCR analysis of genomic DNA from the NT lines, the absence of an amplicon confirms that the target gene is physically located on the nullisomic chromosome pair. This analysis was conducted for all 24 genes (Fig. 3-2B) and the results were in agreement with the TAGCv1 and IWGSC css assembly. To predict the subcellular localisation of the putative proteins, WoLF PSORT and TargetP1.1 were used (Table S3-3). XDH, which in the literature is reported localised in the cytosol (Werner & Witte, 2011), was suggested to be located in either

chloroplasts or plastids by WoLF PSORT or in unspecified organelles (other) by TargetP1.1. For the enzymes reported located in the peroxisomes (UOX and AS) WoLF PSORT suggested TaUOX proteins to be localised in the peroxisomes but predicted chloroplastic localisation for TaAS protein; whilst, TargetP1.1 predicted both enzymes located in mitochondria. The remaining enzymes (ALN, AAH, UGAH and UAH) are reported localised in the endoplasmic reticulum (E. R.), however, WoLF PSORT suggested localisation in the E.R. for just TaALN, TaUAH and TaUGAH-2DS homeolog; whilst, TargetP1.1 predicted all the enzymes to be located within the secretory pathway (SP).

To further corroborate the orthology between the identified loci we analysed the synteny of the genomic regions between bread wheat, Brachypodium, rice and sorghum. The fragmented nature of the bread wheat TGACv1 assembly (Clavijo *et al.*, 2017) may have reduced the resolution of the analysis, however, wheat showed a high degree of synteny with the three diploid genomes included in the analysis and the position of several purine catabolic genes and their neighbouring genes was highly conserved (Fig. 3-3; Table S3-4). The only exception was *XDH* for which we could identify only one syntenic gene (Bradi1g15910, depicted in blue) among all analysed genomes. This gene is located upstream of *XDH* in sorghum, downstream of *XDH* in rice and Brachypodium, and on chromosome group 1 in wheat. Interestingly, we identified two Brachypodium genes (Bradi1g15820 and Bradi1g15826, depicted in brown and grey, respectively) with orthologous sequences on wheat chromosome group 6, supporting the evidence that a duplication of the genomic region harbouring *XDH* may have occurred during wheat genome evolution. We also identified two copies of *XDH* in the genomes of wheat ancestors, which will be described in more details below. This analysis further revealed the presence of two copies of *UAH* in the Brachypodium genome of which only one gene (*BdUAH1*) showed a syntenic relationship with the other genomes.

***In-silico* analysis of the two *TaXDH* genes and encoded proteins highlights unique features in *TaXDH2-6DS*.**

Analysis of the annotated open reading frames (ORFs) of the *TaXDH* genes revealed five premature stop codons in *TaXDH2-6AS* and several mutations in *TaXDH2-6BS*, both likely to result in non-functional proteins. In contrast, the *TaXDH2-6DS* ORF appeared to translate into a functional protein (Fig. S3-1). To investigate this further, other grass genomes closely related to bread wheat were used in a search for *TaXDH2* orthologous sequences. The analysis revealed only one *XDH* ortholog in the barley genome (*HvXDH*) located on chromosome H1, whereas orthologous sequences of Chinese Spring chromosome group 6 *TaXDH2* homeologs were identified in diploid wheat ancestor genomes of *Aegilops tauschii* (*tauXDH2*) and *Triticum urartu* (*uraXDH2*), tetraploid durum wheat cv. Kronos (*TdXDH2-A/TdXDH2-B*) and in hexaploid bread wheat cv. Cadenza (*TaXDH2-6AS-Cad/ TaXDH2-6BS-Cad/ TaXDH2-6DS-Cad*) (Table S3-4). In all genomes containing *TaXDH2* orthologous genes, analysis of the CDSs revealed that *XDH2-A* genes have premature stop codons at positions also conserved in *TaXDH2-6AS* (Fig. S3-2). The *XDH-B* genes have severely mutated ORFs as *TaXDH2-6BS* (data not shown), while *XDH-D* genes have putative functional CDSs. To further characterise the evolutionary relationship of *TaXDH1* and *TaXDH2*, a phylogenetic analysis was conducted using *XDH* CDSs from grass diploid genomes, wheat ancestors and Arabidopsis. *TaXDH2-6BS* was not included in the analysis as the annotated CDS presented several mutations (Fig. S3-1). The resulting phylogenetic tree (Fig. S3-3) revealed that *TaXDH1* and *TaXDH2* separated into two distinct clades and that the barley ortholog, *HvXDH*, is more closely related to the *TaXDH1* clade than the *TaXDH2* clade. The *XDH* duplication in Arabidopsis consists of a tandem duplication as *AtXDH1* and *AtXDH2* genes are adjacent on chromosome 4, therefore, it is not likely related to the duplication observed in wheat and its progenitors. A protein alignment of XDH from grasses and Arabidopsis genomes highlighted that the TaXDH2 clade orthologous has an 11 amino acid insertion in the molybdenum cofactor (moco)-binding domain and an arginine to histidine substitution at the xanthine binding site (corresponding to position 909 in AtXDH1) (Fig. 3-4). Interestingly, site-direct mutagenesis of

AtXDH1^{Arg909} and insertion of alanine (AtXDH1^{Ala909}) was reported to disrupt XDH enzymatic activity when xanthine was used as substrate (Zarepour *et al.*, 2010).

To investigate this further, TaXDH1 and TaXDH2 protein structures were predicted using homology modelling. TaXDH1 and TaXDH2 share 47% and 46% amino acid identity with the bovine milk XDH, respectively, thus their crystal structures can be used to make high-quality models. Rat and bovine XDH structures dock xanthine in a prime position for nucleophilic attack and the key xanthine hydrogen bonding interactions include XDH^{Glu1261}, XDH^{Glu802} and XDH^{Arg880} (Cao *et al.*, 2010) (Fig. 3-5A). Analysis of the TaXDH1-1DL homology model docked with xanthine revealed a highly conserved orientation, with all hydrogen bonding residues maintained in the pocket (Fig. 3-5A). Conversely, in the TaXDH2-6DS model the substitution of arginine to histidine (TaXDH2-6DS^{His920}) significantly alters the properties of the pocket resulting in a rotated conformation of xanthine and an alteration of hydrogen bonding (Fig. 3-5A). Arginine is this position is highly conserved in Arabidopsis, diploid grasses and in the TaXDH1 clade orthologs (Fig. 3-4) and it has been shown to be required for uric acid formation in plants (Zarepour *et al.*, 2010) and bovine (Cao *et al.*, 2010) as the arginine side-chain positive charge is proposed to stabilise the negative charge during the reaction transition state (Cao *et al.*, 2010). To analyse the potential role of the extra TaXDH2-6DS loop (670-680), the model was optimised by database loop searching and energy minimisation which predicted a short helical structure between XDH2^{Arg675} and XDH2^{Lys679}. This loop appears to largely cover the xanthine entry at the TaXDH2-6DS protein surface and to reduce the positive charge that characterises the TaXDH1-1DL xanthine pocket (Fig. 3-5B). Although docking cannot determine whether the loop prevents xanthine entry, the change in electrostatic properties around the xanthine pocket and the significant reduction in the size of the TaXDH2-6DS substrate entrance suggests that the binding efficiency of xanthine is likely to be reduced. On the other hand, the residues lining the iron-sulfur clusters (2Fe-2S) and FAD-binding domain are highly conserved and both wheat TaXDH1-1DL and TaXDH2-6DS displayed high structure identity with bovine XDH (Fig. S3-4).

Expression analysis of purine catabolic genes in wheat

To assess the temporal and spatial expression of the purine catabolic genes, the drought tolerant Australian wheat genotype RAC875 was selected because it is widely used in Australian breeding programs and well characterised at the physiological level (Izanloo *et al.*, 2008; Bennett *et al.*, 2012; Bonneau *et al.*, 2013; Melino *et al.*, 2015). For this analysis qRT-PCR primers were designed to amplify all three homeologs of each purine catabolic gene. The only exception was the design of primers to amplify only one homeolog of *TaXDH2*, TaXDH2-*6DS,* which *in-silico* analysis of publicly available Chinese Spring RNA-seq transcriptomics data revealed was the only homeolog expressed (Fig. 3-6B). Analysis was carried out using RNA extracted from the youngest fully emerged leaves (YFEL) collected 24 days after sowing, and from plant tissues (leaves, stems, floral organs) collected at 17 days after anthesis. As a general trend, *TaXDH1*, *TaAS* and *TaUAH* and *TaUOX* were expressed at a relatively high level, whilst *TaALN, TaXDH2, TaUGAH* and *TaAAH* expression was lower (Fig. 3-6A). Specifically, expression of genes involved in ureide synthesis (*TaXDH1* and *TaUOX*) were 2-fold higher in flag leaves compared with YFEL, whereas expression of the ureide utilising genes (*TaALN* and *TaUGAH*) was significantly higher in YFEL than flag leaves. Compared to its high expression in flag leaves, *TaXDH1* showed significantly lower expression in stems and spikelets and a 4.5-fold reduced expression in developing grain. Reduced expression of *TaXDH2* in grains was also observed although at a much lower level. In contrast, *TaALN* showed a 10-fold increased expression in developing grains compared to flag leaves and, similarly, *TaAAH* expression was 3-fold higher in developing grain compared to other tissues, though expression was generally low (Fig. 3-6).

Response of wheat purine catabolic genes to dehydration and cold stress

Purine catabolic genes have been previously reported to be differentially expressed under water stress and low temperatures in Arabidopsis, common bean and rice (Hesberg *et al.*, 2004; Alamillo *et al.*, 2010;

Coleto *et al.*, 2014; Li *et al.*, 2015; Irani & Todd, 2016). To assess stress responsiveness of the genes in wheat, gene expression was analysed in RAC875 after exposure to dehydration and cold stress (Fig. 3-7A, B). Within two hours of dehydration, genes involved in ureide utilization were generally down-regulated and, specifically, expression of *TaALN* was reduced by 3.3-fold and *TaUGAH* was reduced by 2-fold, whilst expression of *TaAAH* was reduced by almost 18-fold. Expression of *TaAS* was marginally reduced (20%) after one and two hours as compared to expression under control conditions. In contrast, expression of the purine catabolic gene *TaUAH* was nearly doubled after two hours of dehydration. These findings are in agreement with a study on two wild emmer accessions subjected to a similar dehydration treatment (Fig. S3-5A) (Ergen *et al.*, 2009). After 24 h of cold stress, expression of *TaXDH1* was down-regulated (1.5-fold), whilst *TaUOX* was strongly up-regulated (9.2-fold) (Fig. 3-7B). Likewise, the ureide utilising genes showed a contrasting expression pattern with up-regulation of *TaAAH* (4.5-fold) and downregulation of *TaUGAH*. Similar trends for the genes described here were reported in a study on Canadian spring and winter wheat, however, in contrast to our data, they found that *TaAAH* expression was down-regulated and *TaUGAH* expression levels were only marginally altered (Laudencia-Chingcuanco *et al.*, 2011) (Fig. S3-5B).

Allantoin differentially accumulates in mature wheat grains depending on N supply

Since purine catabolism is considered to contribute to N remobilisation during plant senescence (Pèlissier & Tegeder, 2007; Brychkova *et al.*, 2008), we investigated whether plants grown under N limiting conditions have a reduced level of allantoin in mature grain. For this, a relative quantitative method for allantoin was developed and allantoin levels were analysed in mature grain of RAC875 and two other Australian genotypes (Spitfire, Kukri) grown under high N (HN) and low N (LN) field conditions. Allantoin was detectable in mature grains of all three genotypes and was 2.5 to 5 times higher in grains from plants

grown under HN compared with plants grown under LN conditions (Fig. 3-8). This suggests a possible N-dependent regulation of this pathway.

Discussion

Wheat and its progenitors have a duplication of *XDH* gene and the homeolog *TaXDH2-6DS* has unique features.

Purine catabolic genes were identified in several grass genomes including hexaploid bread wheat (Fig. 3-2) and showed a high degree of synteny among wheat, Brachypodium, rice and sorghum (Fig. 3-3), suggesting that the loci identified are the true gene orthologs. However, the poor synteny displayed by *XDH* and its adjacent genes even in diploid genomes (Brachypodium, rice and sorghum) suggest that *XDH* is located in an unstable genomic region prone to rearrangements. There is evidence to suggest that *XDH* duplication occurred after speciation of the wheat progenitors from barley (*HvXDH* is closer to the TaXDH1 clade than the TaXDH2 clade) but before the major events of polyploidy took place (*TaXDH2* orthologs were identified in the diploid ancestors of wheat, *A. tauschii* and *T. urartu*). Among the three *TaXDH2* homeologs, only *TaXDH2-6DS* appeared to be functional and expressed in bread wheat cv. Chinese Spring (Fig. 3-6B). However, protein analysis of TaXDH2-6DS revealed an 11 amino acid insertion and an arginine to histidine substitution within the xanthine pocket, located in the moco-binding domain (Fig. 3-4). Protein modelling indicates that these alterations disrupt the polarity of the xanthine binding pocket and may obstruct entrance to the active site of the enzyme (Fig. 3-5). In support of this, site-direct mutagenesis of the corresponding arginine residue in Arabidopsis XDH1 (AtXDH1^{Arg909Ala}) abolished any *in-vitro* enzymatic activity when xanthine was provided as substrate (Zarepour *et al.*, 2010). However, the authors showed that AtXDH1^{Arg909Ala} can still function as an oxidase, producing the superoxide anion ($O_2^{\bullet-}$) in a NADH/O_2-dependent reaction *in-vitro*, suggesting that the function of the FAD-binding site is uncoupled from the moco-binding domain (Zarepour *et al.*, 2010). Our analysis

revealed that both 2Fe-2S clusters and the FAD-binding site of TaXDH1-1DL and TaXDH2-6DS are conserved with bovine XDH (Fig. S3-5). Although TaXDH2-6DS has a defective xanthine pocket, it may still have a functional electron transfer between the NAD, FAD and iron-sulfur cluster binding pockets given the close proximity of these ligands in the protein tertiary structure, suggesting that TaXDH2-6DS may still function as a NADH oxidase. In addition, the expression patterns of *TaXDH2-6DS* as compared to the *TaXDH1* homeologs differed between YFEL and flag leaves (Fig. 3-6A). However, further experimental studies are required to confirm the biological function of the *TaXDH2-6DS* homeolog.

Differential expression of purine catabolic genes in source and sink tissue suggests an involvement of the pathway in nutrient remobilisation in wheat

Grasses use amino acids for long-distance transport of N; in particular glutamine, aspartate, serine and alanine represent 85% of the free amino acids pool in senescing wheat tissues (Lopes *et al.*, 2006). Despite that ureides are not the major N form for long distance transport in wheat, its large free purine and pyrimidine pool in flag leaves (Sawert *et al.*, 1988) suggests that purine precursors are available for catabolism and remobilisation. In legumes it has been shown that allantoin is actively redistributed from source to sink tissues in French bean (Pèlissier & Tegeder, 2007), however, allantoin translocation is yet to be studied in cereals. Nevertheless, results shown here (Fig. 3-8) and those previously demonstrated that wheat and rice accumulate allantoin in their mature grain (Montalbini, 1992; Wang *et al.*, 2012) as a result of either remobilisation from source tissues or *de novo* synthesis in the grain via purine catabolism. The data reported here are in support of ureide remobilisation as expression levels of ureide synthesising genes (*TaXDH1* and *TaUOX*) were higher in flag leaves than YFEL (Fig. 3-6A); the flag leaves have an essential role in nutrient remobilisation and production of photoassimilates during grain filling. In contrast, the purine catabolic pathway genes involved in N liberation from allantoin were either down-regulated (*TaALN* and *TaUGAH*) or not significantly altered (*TaAAH* and *TaUAH*) in source tissues (flag leaf, stem

and spikelet) when compared to YFEL (Fig. 3-6A). These findings are in accordance with reports from Arabidopsis showing that *AtXDH1* and *AtUOX* were up-regulated whilst *AtALN* and *AtAAH* genes were down-regulated in senescing leaves (Brychkova *et al.*, 2008). In support of this, XDH enzymatic activity was reported to be greatly increased in Arabidopsis senescing leaves compared with younger tissues (Hesberg *et al.*, 2004). Similar results were also reported for field pea (*Pisum sativum L.*) where enzymatic activity of PsXDH and PsUOX increased in senescing leaves (Pastori & Del Rio, 1997).

Conversely, in developing grain, the ultimate sink for remobilised nutrients, the combination of relatively lower expression of *TaXDH* (ureide synthesis) and relatively higher levels of *TaALN* and *TaAAH* transcripts (ureide utilisation) compared to flag leaf suggests that ureides are predominantly translocated from other plant tissues and that a portion is catabolised to recycle N. Ureide permease (UPS) and possibly other transporters may take part in the distribution of ureides across plant tissues, including grains, as suggested by Arabidopsis *AtUPS4* specific expression in seeds undergoing later stages of development (Schmidt *et al.*, 2006).

Wheat purine catabolic genes are responsive to dehydration and cold stress

In recent years a growing number of studies have shown that purine catabolic genes are differentially regulated under abiotic stresses, such as drought, salt, cold, dark and heavy metals; these studies were carried out in plants such as Arabidopsis and rice that do not use ureides as a main source for long-distance transport of N (Hesberg *et al.*, 2004; Brychkova *et al.*, 2008; Li *et al.*, 2015; Irani & Todd, 2016; Lescano *et al.*, 2016; Nourimand & Todd, 2016Our data confirms the responsiveness of this pathway in wheat and, despite employing non-homeologs specific primers, the apparent transcriptional changes observed suggest the existence of a specific mechanism regulating the gene expression under stress. Dehydration stress significantly down-regulated ureide-utilising genes (*TaALN, TaAAH* and *TaUGAH*) (Fig. 3-7A) which would result in allantoin accumulation under dehydration stress. This is supported by Bowne *et al.* (2011) who reported an increase of allantoin under drought in flag leaves of the same

genotype (RAC875) used here. However, unexpectedly, expression of the last gene in the pathway, *TaUAH*, increased under dehydration stress in bread wheat and 8 hour dehydration stress in wild emmer (Fig. 3-7A and Fig. S3-5A). This is in agreement with a recent report from rice, which showed a 3-fold upregulation of *OsUAH* after 12 h of air-dry treatment (Li *et al.*, 2015). UAH may have an alternative substrate or function under stress.

In contrast to the response of ureide synthesis and degradation genes to dehydration stress, cold stress (4°C) resulted in down-regulation of *TaXDH1* (Fig. 3-7B). In Arabidopsis, Hesberg *et al.* (2004) observed down-regulation of *AtXDH1* gene expression and reduction of XDH enzymatic activity after 6 to 20 h at both +4 and -4°C, supporting the results shown here. In contrast, global expression analysis showed down-regulation of *TaAAH* after 2 days at 6 °C in spring wheat cv. Manitou whilst no significant changes were detected in winter wheat cv. Norstar (Fig. S3-5B). Additionally, in a study on rice, Li *et al.* (2015) showed that *OsUAH* was up-regulated almost 10-fold after 24 h at 4°C which is in contrast to the down-regulation observed in RAC875 (Fig. 3-7B).

Further analyses are required to address these discrepancies in detail, however, since the primary function of purine catabolism is the recycling of the N present in the purine ring, differences in the developmental stage and/or nutrient status of the plants might at least partly explain the observed differences in gene expression patterns between the fore-mentioned studies. In addition, post-translational regulation may be involved in controlling the flux of metabolites within this pathway under stress, indicating the need for further comprehensive studies of enzyme activities and accumulation of metabolic intermediates.

Low N basal application reduces the amount of allantoin in mature wheat grains in the field

Given the well-established role of purine catabolism in N remobilisation (Pèlissier & Tegeder, 2007; Brychkova *et al.*, 2008) and the importance of ureides for N transport in certain legumes species (Herridge

et al., 1978; Pate *et al.*, 1980), it is surprising that very little information on the regulation of purine catabolism under N restrictions is available in wheat so far. Our findings provide the first evidence that N limitation during wheat growth indeed reduces the accumulation of allantoin in mature grains suggesting that the recycling and remobilisation of purines in the form of allantoin under nitrogen limitation may contribute to maintaining N homeostasis in wheat (Fig. 3-8). Similarly, recent metabolomics studies showed reduced allantoin levels under N restriction in maize vegetative leaves and in hydroponically grown rice roots (Amiour *et al.*, 2012; Coneva *et al.*, 2014). Taken together, there is converging evidence that allantoin derived from purine catabolism has an important role in maintaining N homeostasis in plants warranting a more in-depth analysis under different N treatments.

The data presented in this study show that purine catabolism in wheat is responsive to both developmental and environmental cues. During natural senescence, purine catabolism is involved in in N recycling (Zrenner *et al.*, 2006; Werner & Witte, 2011), whereas under abiotic stress conditions, the pathway is tuned towards accumulations of intermediates, such as allantoin, with possible protective and regulatory functions (Watanabe *et al.*, 2014a). Such a dual role is directly relevant for crops and especially for wheat, for which high grain N content is an important quality trait. During grain filling stages, a prompt response to abiotic stresses and efficient N remobilisation from vegetative tissues to the grain can positively impact grain yield and grain protein content. These are major determinants of the economic value of wheat and are important targets of breeding programs (Branlard *et al.*, 2001; Bogard *et al.*, 2010; Hawkesford, 2017). Therefore, the wheat purine catabolic genes described here represent an interesting and relevant target for crop improvement by either conventional and /or transgenic approaches.

Acknowledgments

We thank Adam Lukaszewski (University of Riverside, California, US) for providing the Chinese Spring NT seeds stock, Margaret Pallotta (University of Adelaide) for providing DNA of the NT lines, Juan Carlos

Sanchez Ferrero (University of Adelaide) for support in annotating the gene sequences, and Himasha Mendis (University of Melbourne) for technical support with the GC-MS analysis. This work was supported by an Australian Research Council linkage grant LP1400100239 in partnership with DuPont-Pioneer, the Fay Fuller Foundation, Senior Research Fellowship (1042589) from the National Health and Medical Research Council of Australia (to S.M. Pitson) and a RAH Early Career Fellowship (to M.R. Pitman). The authors declare that there is no conflict of interest related to this work.

References

Abagyan R. & Totrov M. (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *Journal of molecular biology* **235**, 983-1002.

Alamillo J.M., Diaz-Leal J.L., Sanchez-Moran M.V. & Pineda M. (2010) Molecular analysis of ureide accumulation under drought stress in *Phaseolus vulgaris* L. *Plant Cell & Environment* **33**, 1828-1837.

Amiour N., Imbaud S., Clément G., Agier N., Zivy M., Valot B., ... Hirel B. (2012) The use of metabolomics integrated with transcriptomic and proteomic studies for identifying key steps involved in the control of nitrogen metabolism in crops such as maize. *Journal Of Experimental Botany* **63**, 5017-5033.

Barbottin A., Lecomte C., Bouchard C. & Jeuffroy M.-H. (2005) Nitrogen remobilization during grain filling in wheat: genotypic and environmental effects. *Crop Science* **45**, 1141.

Bennett D., Izanloo A., Reynolds M., Kuchel H., Langridge P. & Schnurbusch T. (2012) Genetic dissection of grain yield and physical grain quality in bread wheat (*Triticum aestivum* L.) under water-limited environments. *Theoretical and Applied Genetics* **125**, 255-271.

Bogard M., Allard V., Brancourt-Hulmel M., Heumez E., Machet J.-M., Jeuffroy M.-H., ... Le Gouis J. (2010) Deviation from the grain protein concentration–grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. *Journal of Experimental Botany* **61**, 4303-4312.

Bolger A.M., Lohse M. & Usadel B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120

Bonneau J., Taylor J., Parent B., Bennett D., Reynolds M., Feuillet C., ... Mather D. (2013) Multi-environment analysis and improved mapping of a yield-related QTL on chromosome 3B of wheat. *Theoretical and Applied Genetics* **126**, 747-761.

Bowne J.B., Erwin T.A., Juttner J., Schnurbusch T., Langridge P., Bacic A. & Roessner U. (2011) Drought responses of leaf tissues from wheat cultivars of differing drought tolerance at the metabolite level. *Molecular Plant* **5**, 418-429.

Branlard G., Dardevet M., Saccomano R., Lagoutte F. & Gourdon J. (2001) Genetic diversity of wheat storage proteins and bread wheat quality. *Euphytica* **119**, 59-67.

Brychkova G., Alikulov Z., Fluhr R. & Sagi M. (2008) A critical role for ureides in dark and senescence-induced purine remobilization is unmasked in the *Atxdh1* Arabidopsis mutant. *Plant Journal* **54**, 496-509.

Casartelli A, Riewe D, Hubberten HM, Altmann T, Hoefgen R, Heuer S. 2018. Exploring traditional aus-type rice for metabolites conferring drought tolerance. *Rice* **11**, 9.

Cao H., Pauff J.M. & Hille R. (2010) Substrate Orientation and Catalytic Specificity in the Action of Xanthine Oxidase: The Sequential Hydroxylation of Hypoxanthine to Uric Acid. *Journal of Biological Chemistry* **285**, 28044-28053.

Chmielewska K., Rodziewicz P., Swarcewicz B., Sawikowska A., Krajewski P., Marczak Ł., ... Stobiecki M. (2016) Analysis of Drought-Induced Proteomic and Metabolomic Changes in Barley (*Hordeum vulgare* L.) Leaves and Roots Unravels Some Aspects of Biochemical Mechanisms Involved in Drought Tolerance. *Frontiers in Plant Science* **7**, 1108.

Chomczynski P. (1993) A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques* **15**, 532-534, 536-537.

Choulet F., Alberti A., Theil S., Glover N., Barbe V., Daron J., ... Paux E. (2014) Structural and functional partitioning of bread wheat chromosome 3B. *Science* **345**, 1249721

Clavijo B.J., Venturini L., Schudoma C., Accinelli G.G., Kaithakottil G., Wright J., ... Clark M.D. (2017) An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Research* **27**, 885-896.

Coleto I., Pineda M., Rodino A.P., De Ron A.M. & Alamillo J.M. (2014) Comparison of inhibition of N2 fixation and ureide accumulation under water deficit in four common bean genotypes of contrasting drought tolerance. *Annals of Botany* **113**, 1071-1082.

Coneva V., Simopoulos C., Casaretto J.A., El-kereamy A., Guevara D.R., Cohn J., ... Rothstein S.J. (2014) Metabolic and co-expression network-based analyses associated with nitrate response in rice. *BMC Genomics* **15**, 1056.

de Silva M., Purcell L.C. & King C.A. (1996) Soybean petiole ureide response to water deficits and decreased transpiration. *Crop Science* **36**, 611.

Degenkolbe T., Do P.T., Kopka J., Zuther E., Hincha D.K. & Köhl K.I. (2013) Identification of drought tolerance markers in a diverse population of rice cultivars by expression and metabolite profiling. *PloS one* **8**, e63637.

Dobin A., Davis C.A., Schlesinger F., Drenkow J., Zaleski C., Jha S., ... Gingeras T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21.

Ergen N.Z., Thimmapuram J., Bohnert H.J. & Budak H. (2009) Transcriptome pathways unique to dehydration tolerant relatives of modern wheat. *Functional & Integrative Genomics* **9**, 377-396.

Gill B.S., Appels R., Botha-Oberholster A.-M., Buell C.R., Bennetzen J.L., Chalhoub B., ... Sasaki T. (2004) A Workshop Report on Wheat Genome Sequencing: International Genome Research on Wheat Consortium. *Genetics* **168**, 1087-1096.

Hanks J.F., Tolbert N.E. & Schubert K.R. (1981) Localization of Enzymes of Ureide Biosynthesis in Peroxisomes and Microsomes of Nodules. *Plant Physiology* **68**, 65-69.

Hawkesford M.J. (2017) Genetic variation in traits for nitrogen use efficiency in wheat. *Journal of Experimental Botany* **68**, 2627-2632.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome biology* **8**, R19.

Herridge D.F., Atkins C.A., Pate J.S. & Rainbird R.M. (1978) Allantoin and Allantoic Acid in the Nitrogen Economy of the Cowpea (*Vigna unguiculata* [L.] Walp.). *Plant Physiology* **62**, 495-498.

Hesberg C., Hänsch R., Mendel R.R. & Bittner F. (2004) Tandem orientation of duplicated xanthine dehydrogenase genes from *Arabidopsis thaliana*: Differential gene expression and enzyme activities. *Journal of Biological Chemistry* **279**, 13547-13554.

Irani S. & Todd C.D. (2016) Ureide metabolism under abiotic stress in *Arabidopsis thaliana*. *Journal of Plant Physiology* **199**, 87-95.

Ishikita H., Eger B.T., Okamoto K., Nishino T. & Pai E.F. (2012) Protein Conformational Gating of Enzymatic Activity in Xanthine Oxidoreductase. *Journal of the American Chemical Society* **134**, 999-1009.

Izanloo A., Condon A.G., Langridge P., Tester M. & Schnurbusch T. (2008) Different mechanisms of adaptation to cyclic water stress in two South Australian bread wheat cultivars. *Journal of Experimental Botany* **59**, 3327-3346.

Kim K., Park J. & Rhee S. (2007) Structural and Functional Basis for (S)-Allantoin Formation in the Ureide Pathway. *Journal of Biological Chemistry* **282**, 23457-23464.

King C.A. & Purcell L.C. (2005) Inhibition of N_2 fixation in soybean is associated with elevated ureides and amino acids. *Plant physiology* **137**, 1389.

Kovalchuk N., Jia W., Eini O., Morran S., Pyvovarenko T., Fletcher S., ... Lopato S. (2013) Optimization of TaDREB3 gene expression in transgenic barley using cold-inducible promoters. *Plant Biotechnology Journal* **11**, 659-670.

Lamberto I., Percudani R., Gatti R., Folli C. & Petrucco S. (2010) Conserved Alternative Splicing of Arabidopsis Transthyretin-Like Determines Protein Localization and S-Allantoin Synthesis in Peroxisomes. *The Plant Cell* **22**, 1564-1574.

Laudencia-Chingcuanco D., Ganeshan S., You F., Fowler B., Chibbar R. & Anderson O. (2011) Genome-wide gene expression analysis supports a developmental model of low temperature tolerance gene regulation in wheat (*Triticum aestivum* L.). *BMC Genomics* **12**, 299.

Lea P.J. & Miflin B.J. (2010) Nitrogen Assimilation and its Relevance to Crop Improvement. *Annual Plant Reviews* **42**, 1-40.

Lescano C., Martini C., González C. & Desimone M. (2016) Allantoin accumulation mediated by allantoinase downregulation and transport by Ureide Permease 5 confers salt stress tolerance to Arabidopsis plants. *Plant Molecular Biology* **91**, 581-595.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., ... Durbin R. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079.

Li J., Qin R.-Y., Li H., Xu R.-F., Yang Y.-C., Ni D.-H., ... Yang J.-B. (2015) Low-Temperature-Induced Expression of Rice Ureidoglycolate Amidohydrolase is Mediated by a C-Repeat/Dehydration-Responsive Element that Specifically Interacts with Rice C-Repeat-Binding Factor 3. *Frontiers in Plant Science* **6**, 1011.

Lopes M., Cortadellas N., Kichey T., Dubois F., Habash D. & Araus J. (2006) Wheat nitrogen metabolism during grain filling: comparative role of glumes and the flag leaf. *Planta* **225**, 165-181.

Melino V.J., Fiene G., Enju A., Cai J., Buchner P. & Heuer S. (2015) Genetic diversity for root plasticity and nitrogen uptake in wheat seedlings. *Functional Plant Biology* **42**, 942.

Mickelson-Young L., Endo T.R. & Gill B.S. (1995) A cytogenetic ladder-map of the wheat homoeologous group-4 chromosomes. *Theoretical and Applied Genetics* **90**, 1007-1011.

Montalbini P. (1992) Ureides and enzymes of ureide synthesis in wheat seeds and leaves and effect of allopurinol on *Puccinia recondita* f. sp. *tritici* infection. *Plant Science* **87**, 225-231.

Nakagawa A., Sakamoto S., Takahashi M., Morikawa H. & Sakamoto A. (2007) The RNAi-mediated silencing of xanthine dehydrogenase impairs growth and fertility and accelerates leaf senescence in transgenic Arabidopsis plants. *Plant and Cell Physiology* **48**, 1484-1495.

Nam M., Bang E., Kwon T., Kim Y., Kim E., Cho K., ... Yoon I. (2015) Metabolite Profiling of Diverse Rice Germplasm and Identification of Conserved Metabolic Markers of Rice Roots in Response to Long-Term Mild Salinity Stress. *International Journal of Molecular Sciences* **16**, 21959-21974.

Nishino T., Okamoto K., Kawaguchi Y., Matsumura T., Eger B.T., Pai E.F. & Nishino T. (2015) The C-terminal peptide plays a role in the formation of an intermediate form during the transition between xanthine dehydrogenase and xanthine oxidase. The *FEBS Journal* **282**, 3075-3090.

Nourimand M. & Todd C.D. (2016) Allantoin increases cadmium tolerance in Arabidopsis via activation of antioxidant mechanisms. *Plant and Cell Physiology* **57**, 2485-2496.

Oliver M.J., Guo L., Alexander D.C., Ryals J.A., Wone B.W. & Cushman J.C. (2011) A sister group contrast using untargeted global metabolomic analysis delineates the biochemical regulation underlying desiccation tolerance in *Sporobolus stapfianus*. *Plant Cell* **23**, 1231-1248.

Pastori G.M. & Del Rio L.A. (1997) Natural Senescence of Pea Leaves (An Activated Oxygen-Mediated Function for Peroxisomes). *Plant Physiology* **113**, 411-418.

Pate J.S., Atkins C.A., White S.T., Rainbird R.M. & Woo K.C. (1980) Nitrogen Nutrition and Xylem Transport of Nitrogen in Ureide-producing Grain Legumes. *Plant Physiology* **65**, 961-965.

Pavitt D.V., De Fonseka S., Al-Khalaf N., Cam J.M. & Reaveley D.A. (2002) Assay of serum allantoin in humans by gas chromatography-mass spectrometry. *Clinica Chimica Acta* **318**, 63-70.

Pèlissier H.C. & Tegeder M. (2007) PvUPS1 plays a role in source–sink transport of allantoin in French bean (*Phaseolus vulgaris*). *Functional Plant Biology* **34**, 282.

Pessoa J., Sárkány Z., Ferreira-da-Silva F., Martins S., Almeida M.R., Li J. & Damas A.M. (2010) Functional characterization of *Arabidopsis thaliana* transthyretin-like protein. *BMC Plant Biology* **10**, 30.

Ramazzina I., Folli C., Secchi A., Berni R. & Percudani R. (2006) Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nature Chemical Biology* **2**, 144-148.

Rosemeyer H. (2004) The Chemodiversity of Purine as a Constituent of Natural Products. *Chemistry & Biodiversity* **1**, 361-401.

Rutherford K., Parkhill J., Crook J., Horsnell T., Rice P., Rajandream M.A. & Barrell B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics (Oxford, England)* **16**, 944-945.

Sawert A., Backer A.I. & Wagner K.G. (1988) Age-Dependent Decrease of Nucleoside Pools in Cereal Leaves. *Plant and Cell Physiology* **29**, 61-65.

Schmidt A., Baumann N., Schwarzkopf A., Frommer W. & Desimone M. (2006) Comparative studies on Ureide Permeases in *Arabidopsis thaliana* and analysis of two alternative splice variants of AtUPS5. *Planta* **224**, 1329-1340.

Schubert K.R. (1986) Products of biological nitrogen fixation in higher plants: synthesis, transport, and metabolism. *Annual Review of Plant Physiology* **37**, 539-574.

Sears E.R., University of M. & Agricultural Experiment S. (1954) The aneuploids of common wheat. *Missouri Agricultural Experiment Station Annual Bulletin* **572**, 1-59.

Serventi F., Ramazzina I., Lamberto I., Puggioni V., Gatti R. & Percudani R. (2010) Chemical basis of nitrogen recovery through the ureide pathway: formation and hydrolysis of S-ureidoglycine in plants and bacteria. *ACS chemical biology* **5**, 203.

Sinclair T.R. & Serraj R. (1995) Legume nitrogen fixation and drought. *Nature* **378**, 344-344.

Singh G., Kumar S. & Singh P. (2003) A quick method to isolate RNA from wheat and other carbohydrate-rich seeds. *Plant Molecular Biology Reporter* **21**, 93-93.

Takagi H., Ishiga Y., Watanabe S., Konishi T., Egusa M., Akiyoshi N., ... Sakamoto A. (2016) Allantoin, a stress-related purine metabolite, can activate jasmonate signaling in a MYC2-regulated and abscisic acid-dependent manner. *Journal of Experimental Botany* **67**, 2519.

Todd C.D. & Polacco J.C. (2006) AtAAH encodes a protein with allantoate amidohydrolase activity from *Arabidopsis thaliana*. *Planta* **223**, 1108-1113.

Triplett E.W., Blevins D.G. & Randall D.D. (1982) Purification and properties of soybean nodule xanthine dehydrogenase. *Archives of Biochemistry and Biophysics* **219**, 39-46.

Wang P., Kong C.H., Sun B. & Xu X.H. (2012) Distribution and function of allantoin (5-ureidohydantoin) in rice grains. *Journal of Agricoltural and Food Chemestry* **60**, 2793-2798.

Wang W.-S., Zhao X.-Q., Li M., Huang L.-Y., Xu J.-L., Zhang F., ... Li Z.-K. (2016) Complex molecular mechanisms underlying seedling salt tolerance in rice revealed by comparative transcriptome and metabolomic profiling. *Journal of Experimental Botany* **67**, 405-419.

Watanabe S., Kounosu Y., Shimada H. & Sakamoto A. (2014a) Arabidopsis xanthine dehydrogenase mutants defective in purine degradation show a compromised protective response to drought and oxidative stress. *Plant Biotechnology* **31**, 173-178.

Watanabe S., Matsumoto M., Hakomori Y., Takagi H., Shimada H. & Sakamoto A. (2014b) The purine metabolite allantoin enhances abiotic stress tolerance through synergistic activation of abscisic acid metabolism. *Plant Cell & Environment* **37**, 1022-1036.
Watanabe S., Nakagawa A., Izumi S., Shimada H. & Sakamoto A. (2010) RNA interference-mediated suppression of xanthine dehydrogenase reveals the role of purine metabolism in drought tolerance in Arabidopsis. *FEBS Letter* **584**, 1181-1186.
Werner A.K., Romeis T. & Witte C.P. (2010) Ureide catabolism in *Arabidopsis thaliana* and *Escherichia coli*. *Nature Chemical Biology* **6**, 19-21.
Werner A.K., Sparkes I.A., Romeis T. & Witte C.P. (2008) Identification, biochemical characterization, and subcellular localization of allantoate amidohydrolases from Arabidopsis and soybean. *Plant Physiology* **146**, 418-430.
Werner A.K. & Witte C.P. (2011) The biochemistry of nitrogen mobilization: purine ring catabolism. *Trends in Plant Science* **16**, 381-387.
Wu H., Liu X., You L., Zhang L., Zhou D., Feng J., ... Yu J. (2012) Effects of Salinity on Metabolic Profiles, Gene Expressions, and Antioxidant Enzymes in Halophyte *Suaeda salsa*. *Journal of Plant Growth Regulation* **31**, 332-341.
Yang J. & Han K.H. (2004) Functional characterization of allantoinase genes from Arabidopsis and a nonureide-type legume black locust. *Plant Physiology* **134**, 1039-1049.
Yobi A., Wone B.W., Xu W., Alexander D.C., Guo L., Ryals J.A., ... Cushman J.C. (2013) Metabolomic profiling in *Selaginella lepidophylla* at various hydration states provides new insights into the mechanistic basis of desiccation tolerance. *Molecular Plant* **6**, 369-385.
Zarepour M., Kaspari K., Stagge S., Rethmeier R., Mendel R.R. & Bittner F. (2010) Xanthine dehydrogenase AtXDH1 from *Arabidopsis thaliana* is a potent producer of superoxide anions via its NADH oxidase activity. *Plant Molecular Biology* **72**, 301-310.
Zrenner R., Stitt M., Sonnewald U. & Boldt R. (2006) Pyrimidine and purine biosynthesis and degradation in plants. *Annual review of plant biology* **57**, 805.

Figure legends

Fig. 3-1. Outline of the purine catabolic pathway. The degradation reactions are shown beginning with xanthine, the first intermediate of the pathway. Abbreviations of enzymes are enclosed within a box. The putative subcellular localisation of the enzymes and metabolites is indicated. Abbreviations include XDH, xanthine dehydrogenase; UOX, urate oxidase; AS, allantoin synthase; ALN, allantoinase; AAH, allantoate amidohydrolase; UGAH, ureidoglycine aminohydrolase; UAH, ureidoglycolate amidohydrolase; ER, endoplasmatic reticulum. The chemical structure of certain metabolites are shown to demonstrate the loss of the amine groups and carbon skeleton.

Fig. 3-2. Distribution of the purine catabolic genes across the bread wheat allohexaploid genome (*Triticum aestivum* cv. Chinese Spring). (A) The 24 purine catabolic genes are represented by grey bars and their putative location on chromosomes was estimated based on the related Munich Information Centre for Protein Sequences (MIPS) gene annotation on EnsemblPlants (plants.ensembl.org). Fig. adapted from Gill *et al.* (2004). **(B)** Homoelog-specific primers were used in a PCR amplification employing cv. Chinese Spring nulli-tetrasomic (NT) lines as template. For clarity, well titles are limited to the name of the nullisomic chromosome (e.g. N1A) (see Table S3-1 for details). The caption indicates the primer set employed to amplify the NT DNA, for which full details are also reported in Table S3-1. For example, (i) XDH1-1AL,-1BL, -1DL primer sets used to amplify nullisomic lines for chromosome subgroup

1. Absence of a band for a specific NT line whilst displaying amplification of the other two products indicates localisation of the gene on that specific chromosome. (ii to viii) as described for (i).

Fig. 3-3. Syntenic relations of the genomic regions harbouring purine catabolic genes in sorghum, rice, Brachypodium and wheat. Organisms are separated in four distinct rows, genes are represented as coloured shapes and abbreviations of purine catabolic genes are shown. For each purine catabolic gene, the presence of the same coloured shape in different genomes denotes the presence of gene orthologs. Solid lines represent chromosomes and the chromosome number is indicated. This Fig. is not drawn to scale and therefore genome distances cannot be determined from the length of lines here. Dashed lines indicate that genes are located on the same chromosome but not adjacent to the genes of interest (GOI). In the bread wheat genome, genes that were located on the same TGACv1 scaffold as the purine catabolic genes of interest are represented on the same solid line. Genes that were identified on a different scaffold but assembled on the same chromosome are separated from the GOI by a dashed line.

Fig. 3-4. Amino acids alignment of the XDH proteins from grasses and Arabidopsis. Black bar indicates disagreement of a specific protein sequence with the consensus sequence, grey boxes indicates the three main protein domains. Residue Arg909 of AtXDH1 crucial for substrate recognition (Zarepour *et al.*, 2010) is highlighted with a black dotted line. Details of the alignment of the 11 amino acid insertion and AtXDH1Arg909 residue are represented in the bottom panel. Fig.s were prepared with the support of Geneious 10.0.2 software.

Fig. 3-5. TaXDH2-6DS displays disrupted xanthine binding orientation and obstructed entry to the xanthine pocket. (A) Structures are represented in stick format. Hydrogen bonding is depicted as dotted lines. The molybdenum (mo) co-factors are represented as spheres. Non-hydrogen bonding residues were omitted from pictures to aid visibility. Xanthine is coloured by atom type blue= oxygen, red= nitrogen yellow = carbon and white = hydrogen. Bovine XDH (3UNI) and rat XDH (4YSW) crystal structure was docked with xanthine. Both crystal structures show the xanthine molecule is coordinated through hydrogen bonding between amine groups of Arg880(side-chain/backbone) and Thr1010(backbone) and the hydroxyl group of Glu802(side-chain). Glu1216 is positioned to arm diooxothiomolybdenum hydroxyl for nucleophilic attack at the C-8 of the xanthine ring. TaXDH1-1DL binds xanthine in a similar conformation to rat and bovine structures with similar hydrogen bonding interactions. In the TaXDH2-6DS homology model xanthine is docked in a rotated conformation due to a lack of positive charge from His920 substitution in the pocket. Xanthine instead forms hydrogen bonds with catalytic residue Glu1308(sidechain) and Ser1048(backbone) and ring amines. This conformation is predicted to be inactive. **(B)** Top: TaXDH1-1DL and TaXDH2-6DS electrostatic surface. Red indicates positively charged

and blue negatively charged surfaces, white indicates neutral charge. Xanthine substrate is shown in sticks. Xanthine entry is partially covered by helix 670-680 in TaXDH2-6DS. Bottom: TaXDH1-1DL (cyan) and TaXDH2-6DS (orange) ribbon structure. TaXDH2-6DS loop 670-680 is coloured blue. Xanthine is shown in sticks and molybdenum is shown in red spheres.

Fig. 3-6. Gene expression of purine catabolic genes in bread wheat. (A) Relative expression of purine catabolic genes of bread wheat genotype RAC875 were determined by quantitative RT-PCR using gene-specific primers and shown as normalized relative quantities (NRQ) with respect to the most stable reference genes (*TaActin* and *TaEFA2*) . RNA was extracted from the youngest fully emerged leaf (YFEL) of the main stem at 24 days after sowing (tillering, Zadoks growth stage, Z22) and from flag leaf, stem, spikelets and developing grain of the main stem of plants 17 days after anthesis (grain filling, Zadoks growth stage,Z77). Data are weighted mean ±SEM of six biological replicates and letters indicate significant differences for all tissue types within a single gene (one-way ANOVA, Tukey's test, $p<0.05$). **(B)** In-silico analysis of TaXDH2 homeologous gene expression. RNA-seq data from the IWGSC Chinese Spring tissue series dataset (https://wheat-urgi.versailles.inra.fr/Seq-Repository/Expression) was mapped against the IWGSC Chinese Spring WGA V0.3 scaffold containing the *TaXDH2* genes. RNA-seq reads per base pair are displayed in dark grey and exon position annotated on the scaffolds are coloured in light grey. At the top is represented *TaXDH2-6AS*, *TaXDH2-2BS* in the middle and *TaXDH2-6DS* at the bottom.

Fig. 3-7. Relative expression of purine catabolic genes of bread wheat genotype RAC875 under dehydration and cold stress. Levels of gene expression were determined by quantitative RT-PCR using gene-specific primers and shown as normalized relative quantities (NRQ) with respect to the most stable reference genes: *TaGAPdH* and *TaEFA2* for **(A)**, *TaActin* and *TaGAPdH* for **(B)**. **(A)** Purine catabolic gene expression in dehydrated leaves after 0, 1 and 2 hours of stress; **(B)** purine catabolic gene expression in leaves after 0, 4 and 24 hours of cold stress (4 °C). Data are mean ±SEM of 3 biological replicates and letters indicate significant differences ($p<0.05$ by one way ANOVA, Tukey's test).

Fig. 3-8. Allantoin levels in mature grains of field-grown Australian wheat cultivars grown under different N supplies. Mature grains from three different Australian wheat cultivars (RAC875, Spitfire and Kukri) were sourced from a field trial conducted in 2013 (Tarlee, Australia) where plants were grown under low N (LN – 8 kg N ha^{-1}) and high N (LN – 98 kg N ha^{-1}). Allantoin levels were determined by single quad mass spectrometry (QMS) and are represented as normalised relative abundances. Data are mean ±SEM of 3 biological replicates and the asterisk indicates significant differences between the LN and HN treatment within a single cultivar (*$p<0.05$; by Student's *t*-test).

Supporting information:

Table S3-1. Details of the identification and localisation of the purine catabolism genes in bread wheat genome.

Table S3-2. Percentage of nucleic acid identity between the coding sequences of wheat purine catabolism gene.

Table S3-3. Prediction of subcellular localisation of wheat purine catabolic protein.

Table S3-4. Details of the genes used for phylogenetic analysis.

Table S3-5. Details of primers employed for quantitative reverse transcription PCR (qRT-PCR) analysis.

Fig. S3-1. TaXDH2 homeologs coding sequence (CDS) structure.

Fig. S3-2. Genome A-XDH2 proteins alignment from hexaploid bread wheat cv. Chinese Spring and Cadenza, tetraploid durum wheat cv. Kronos and diploid *Triticum urartu* genomes.

Fig. S3-3. Phylogenetic analysis of XDH among grass genomes.

Fig. S3-4. TaXDH1-1DL and XDH2-6DS have highly conserved FAD-binding site and 2Fe-2S clusters.

Fig. S3-5. Expression data of purine catabolism genes from data mining of publicly available data bases.

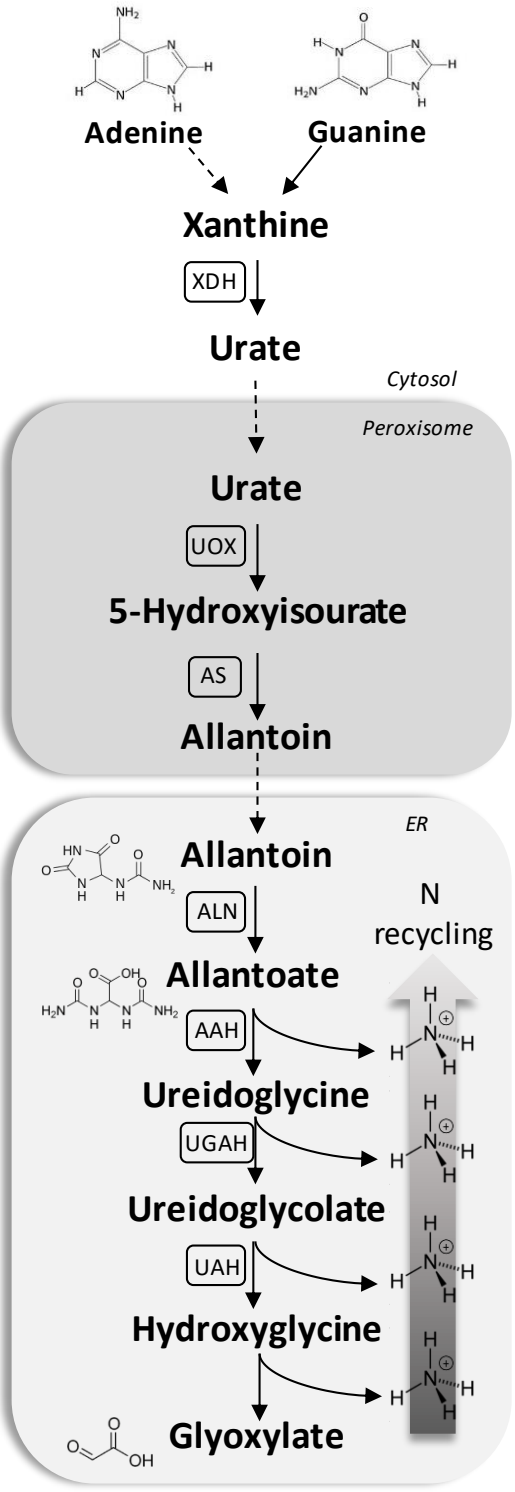


Fig. 3-1. Outline of the purine catabolic pathway.

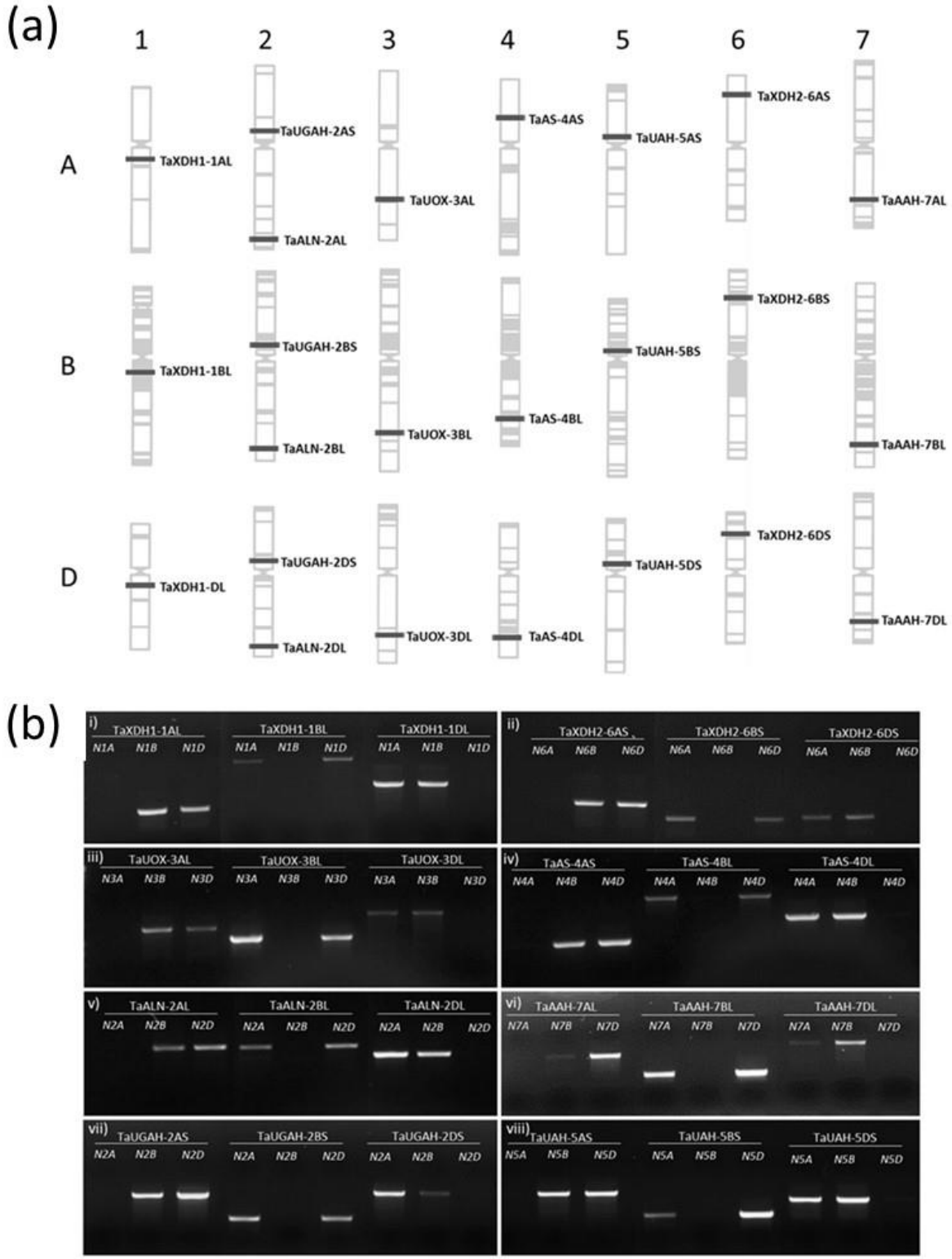


Fig. 3-2. Distribution of the purine catabolic genes across the bread wheat allohexaploid genome (*Triticum aestivum* cv. Chinese Spring).

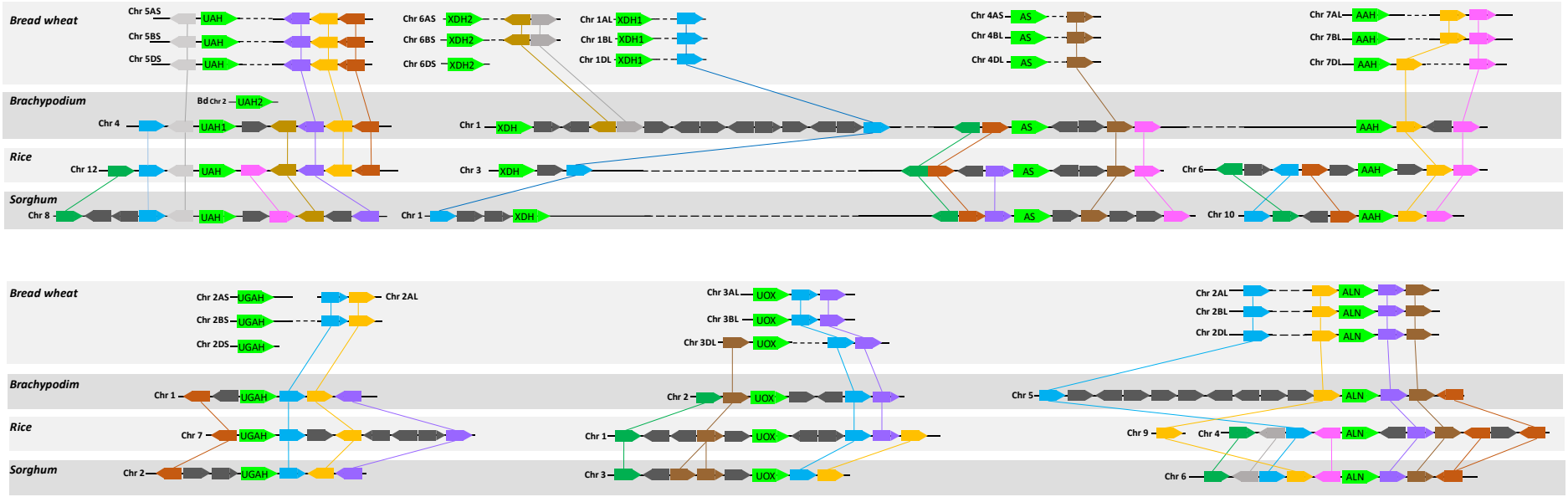


Fig. 3-3. Syntenic relations of the genomic regions harbouring purine catabolic genes in sorghum, rice, Brachypodium and wheat.

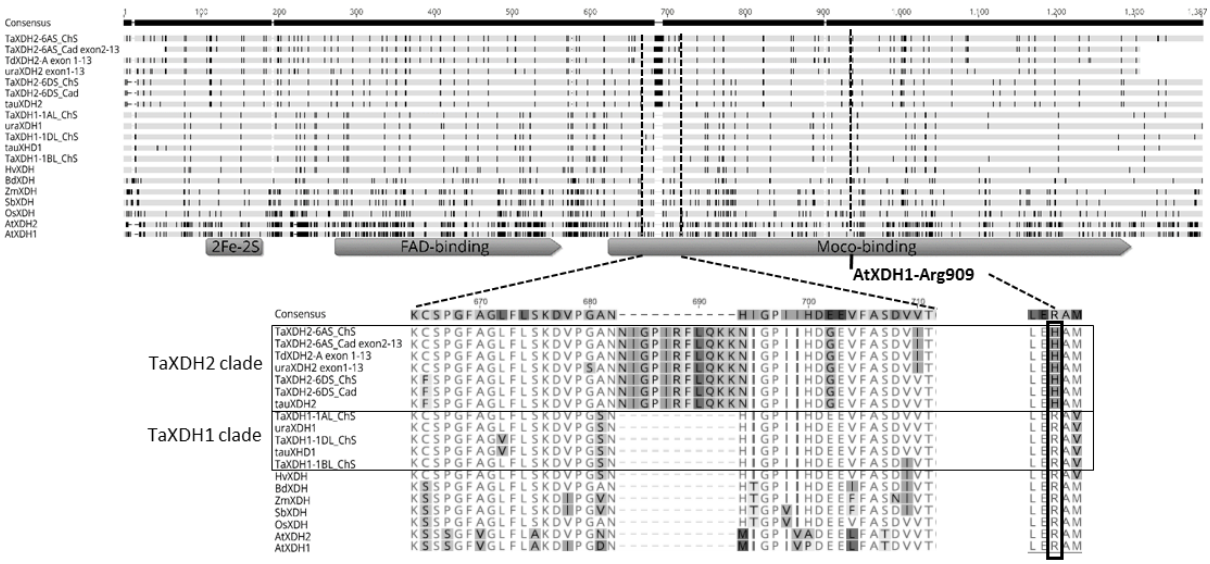


Fig. 3-4. Amino acids alignment of the XDH proteins from grasses and Arabidopsis.

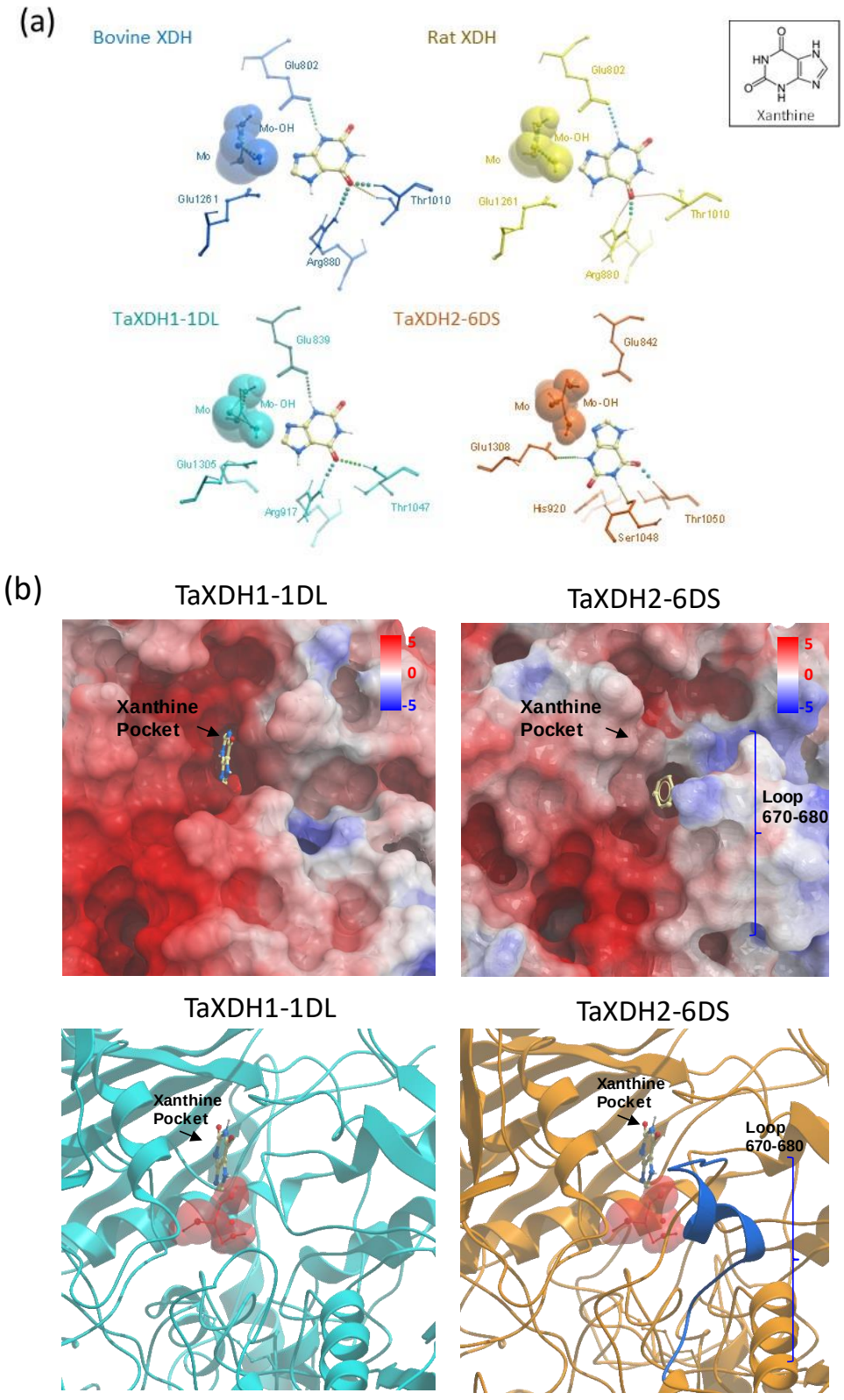


Fig. 3-5. TaXDH2-6DS displays disrupted xanthine binding orientation and obstructed entry to the xanthine pocket.

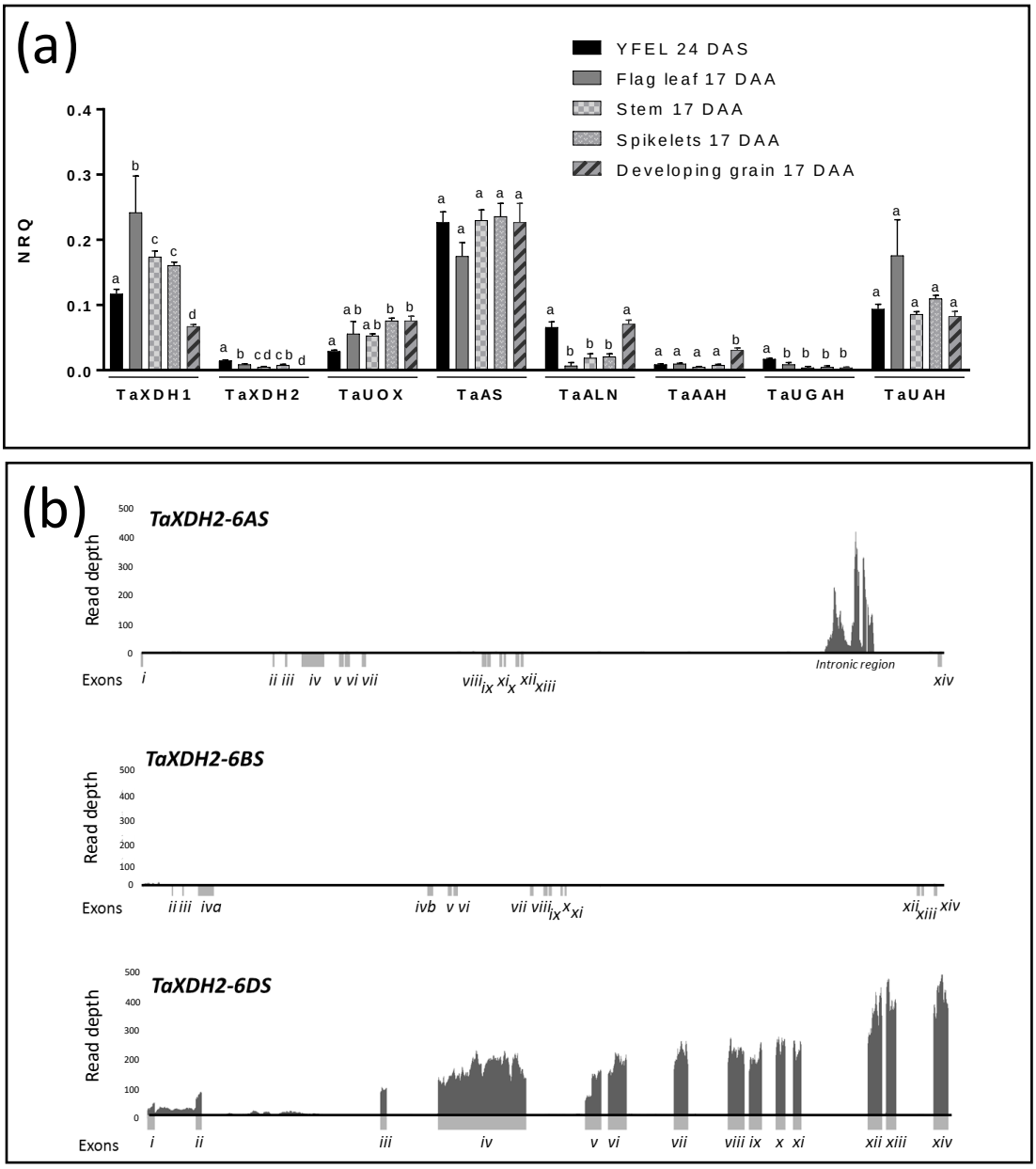


Fig. 3-6. Gene expression of purine catabolic genes in bread wheat.

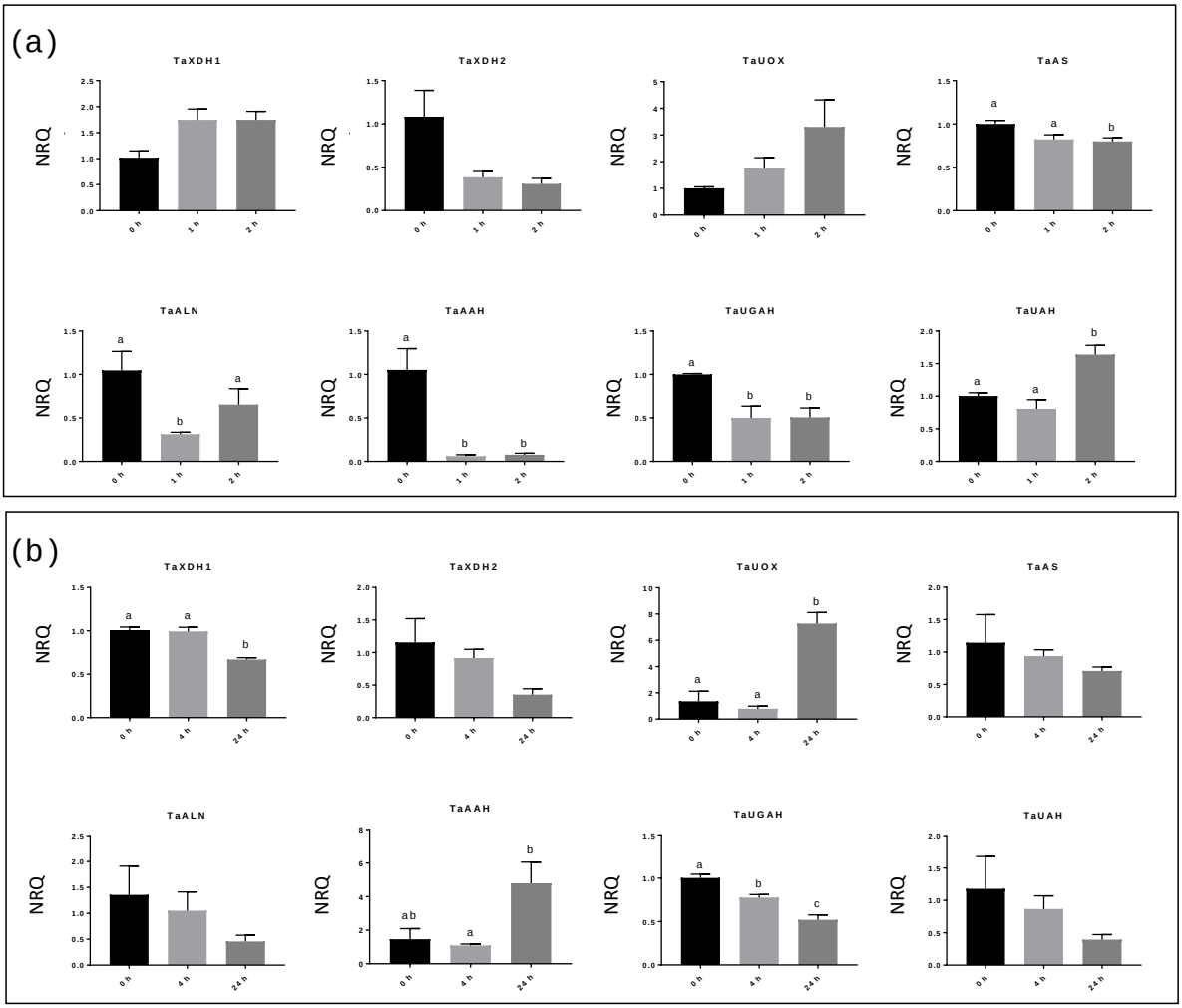


Fig. 3-7. Relative expression of purine catabolic genes of bread wheat genotype RAC875 under dehydration and cold stress.

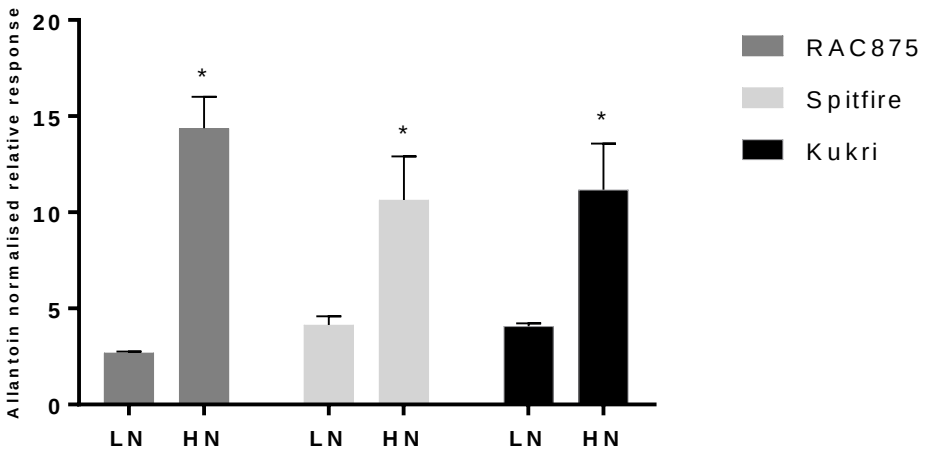


Fig. 3-8. Allantoin levels in mature grains of field-grown Australian wheat cultivars grown under different N supplies.

Chapter 3

Supplementary material

Table S3-1. Details of the identification and localisation of the purine catabolism genes in bread wheat genome.

											Nulli-tetrasomic line			
Gene name	Chromosome location	TGACv1 scaffold ID	Position	Orientation	IWGSC survey sequence v3	TGAC gene prediction	Predicted CDS length (bp)	Forward primer (5'-3)	Reverse primer (5'-3)	Length (bp)	A genome	B genome	D genome	Tm (°C)
XDH1-1AL	1AL	TGACv1_scaffold_001495_1AL	42879-56371	forward	1AL_3917256 (exon iv, v, vi)		4107	CACACGTGAGAGGGGGTGTTACAG	GATTCCCCACTTCCCAAGTCATATC	919	N1A-T1B	N1B-T1A	N1D-T1D	60
XDH1-1BL	1BL	TGACv1_scaffold_030308_1BL	84913-55999	reverse	1BL_3852509 (exon iv)	TRIAE_CS42_1BL_TGACv1_030308_AA0086270	4107	TCCTCTCAGCATTCTTACTTACAGGCTA	ATTTGCATGTTCCCCGTGAATCAC	2056	N1A-T1B	N1B-T1A	N1D-T1D	60
XDH1-1DL	1DL	TGACv1_scaffold_062000_1DL	22603-3578	reverse	1DL_2290396 (exon iv, v, vi)	TRIAE_CS42_1DL_TGACv1_062000_AA0207160	4107	TCTTGGAGCGTGCTGTTTTCCA	GCTTGGAGGCTAACTTGGAAGTGCT	1224	N1A-T1B	N1B-T1A	N1D-T1D	60
XDH2-6AS	6AS	TGACv1_scaffold_485762_6AS, TGACv1_scaffold_488141_6AS	36105-11077 7493-7735	reverse forward			n.a.	TGGCAGTATAGCCCGTTAGTCTTAGG	GATCCCGTTGAACCCATTCAGTC	1262	N6A-T6B	N6B-T6A	N6D_T6B	62
XDH2-6BS	6BS	TGACv1_scaffold_515662_6BS TGACv1_scaffold_516491_6BS TGACv1_scaffold_514923_6BS	30440-33390, 2532-12270 939,410,835	forward forward reverse	exon 1 not found		n.a.	TTCCCAAACACGAAAGAAAGCTAGAG	AGTCAAGCCAAGGTCCAGAGTGAAC	1052	N6A-T6B	N6B-T6A	N6D_T6B	62
XDH2-6DS	6DS	TGACv1_scaffold_544391_6DS	5257-18628	forward		TRIAE_CS42_6DS_TGACv1_544391_AA1748110	4116	TTCCAGAACTCCACACCCTTAAAGTG	GTTGTAGAGGTGACGAAGATGGGAGA	1083	N6A-T6B	N6B-T6A	N6D_T6B	62
UOX-3AL	3AL	TGACv1_scaffold_195992_3AL	17573-20980	forward		TRIAE_CS42_3AL_TGACv1_195992_AA0656220	921	CATTCCCAAAGAAATGGACTCTGC	AGTGGGTAAACATGGGAATTATGGAG	1320	N3A-T3B	N3B-T3D	N3D-T3A	60
UOX-3BL	3BL	TGACv1_scaffold_224491_3B	33739-37253	forward		TRIAE_CS42_3B_TGACv1_224491_AA0797170	921	CAAGCTAATATGAGCTACCATGCACAC	ACACGGCCTTTACTGTCTAATGCAGA	1201	N3A-T3B	N3B-T3D	N3D-T3A	60
UOX-3DL	3DL	TGACv1_scaffold_250245_3DL	7513-4074	reverse		TRIAE_CS42_3DL_TGACv1_250245_AA0864940	921	GACTCTTTTAGTAGATCAGCACGGTATCC	CCATCAGCAATAGAAGTGCTACTCCA	1794	N3A-T3B	N3B-T3D	N3D-T3A	60
AS-4AS	4AS	TGACv1_scaffold_306654_4AS	106421-103331	reverse		TRIAE_CS42_4AS_TGACv1_306654_AA1011590	996	GACAACTCCTGGGCATGGATTATTATC	CAATGTGATTTACACACAAATCGACACC	687	N4A-T4B	N4B-T4D	N4D-T4A	63

AS-4BL	4BL	TGACv1_scaffold_642345_U	29660-26851	reverse	4BL_6969139	TRIAE_CS42_U_TGACv1_642345_AA2116150	996	GCTACCAAGTCTGGAATTTCTGTTCTTC	ATTGAGGTCTCAGGAAGTGGCGAG	1748	N4A-T4B	N4B-T4D	N4D-T4A	63
AS-4DL	4DL	TGACv1_scaffold_345647_4DL	10169-12682	forward		TRIAE_CS42_4DL_TGACv1_345647_AA1154100	996	CACACAATAACTTTGTTTGCCTGAAGC	TCGAGGTCACCAATCATAAATAGCAAG	1059	N4A-T4B	N4B-T4D	N4D-T4A	63
ALN-2AL	2AL	TGACv1_scaffold_095387_2AL	16574-13248	reverse		TRIAE_CS42_2AL_TGACv1_095387_AA0310350	1515	ACCCAAGCAAATTGGAGAGTCTACTG	GCATTATGATCCTGCTGAAAGTGAAAC	1277	N2A-T2D	N2B-T2A	N2D-T2B	60
ALN-2BL	2BL	TGACv1_scaffold_129785_2BL	24458-27757	forward		TRIAE_CS42_2BL_TGACv1_129785_AA0395850	1515	CCTCTCTGAGCTGGAGAGTGCCATA	GGTTTGAATTGTAGGACTGTTGGAACC	1211	N2A-T2D	N2B-T2A	N2D-T2B	60
ALN-2DL	2DL	TGACv1_scaffold_158900_2DL	31798-28513	reverse		TRIAE_CS42_2DL_TGACv1_158900_AA0528270	1515	CAGGCTCAATCATTTCCTATCAAAGC	TTCGTTCATGCAAAGTTAGCCTAGTTG	975	N2A-T2D	N2B-T2A	N2D-T2B	60
AAH-7AL	7AL	TGACv1_scaffold_557474_7AL	44308-48620	forward		TRIAE_CS42_7AL_TGACv1_557474_AA1781750	1482	TGGTGAGAACTATCAACTACGTGTTTCC	GCTGGGCTTTAGATAGTGACATAGTGTG	1592	N7A-T7D	N7B-T7A	N7D-T7A	60
AAH-7BL	7BL	TGACv1_scaffold_577331_7BL	89268-84932	reverse		TRIAE_CS42_7BL_TGACv1_577331_AA1872370	1485	AGAGAAGGCTTTATTGATTGGATCTCAC	AAACGGCTAAGATAACGAAAGCCAAT	764	N7A-T7D	N7B-T7A	N7D-T7B	60
AAH-7DL	7DL	TGACv1_scaffold_604778_7DL	31979-27514	reverse		TRIAE_CS42_7DL_TGACv1_604778_AA2001660	1461	CAAAGAGGTGCATCTGGGTTTCTC	GTTAAGCAGAGCAACTTCGGTGATG	2646	N7A-T7D	N7B-T7A	N7D-T7B	60
UGAH-2AS	2AS	TGACv1_scaffold_112057_2AS	82627-102959	forward		TRIAE_CS42_2AS_TGACv1_112057_AA0329070	906	ATCTATACGCGGGCTTGACGAAGA	CGGCATCGACCAACTCACAAAC	1844	N2A-T2B	N2B-T2D	N2D-T2B	60
UGAH-2BS	2BS	TGACv1_scaffold_146051_2BS	139062-122243	reverse		TRIAE_CS42_2BS_TGACv1_146051_AA0454060	906	GGACTTTCAGTCAGGCGAGTATCTCA	AACTCTTCATAAGCTGGCATAGCTCTCT	991	N2A-T2B	N2B-T2D	N2D-T2B	60
UGAH-2DS	2DS	TGACv1_scaffold_178117_2DS	54209-24134	reverse	2DS_760574 (exon ii)	TRIAE_CS42_2DS_TGACv1_178117_AA0591060	906	GCCCATCCGTCAAACCAACTACT	CTCTATGTATGCGGCTATTTGTAAGTGC	1996	N2A-T2B	N2B-T2D	N2D-T2B	60
UAH-5AS	5AS	TGACv1_scaffold_393200_5AS	38479-43028	forward		TRIAE_CS42_5AS_TGACv1_393200_AA1269720	1395	GTATGGTTGGACCGTTGGAGTAATAGC	AACAGATGATTGCACAACACAGAGGAC	1909	N5A-T5D	N5B-T5B	N5D-T5A	60
UAH-5BS	5BS	TGACv1_scaffold_423557_5BS	23012-18479	reverse		TRIAE_CS42_5BS_TGACv1_423557_AA1379400	1395	ACTTGGTCTTAATTTGTGGGTTCTCATC	GGAAAACCACAGGAATTTGATACGC	1054	N5A-T5D	N5B-T5B	N5D-T5A	60
UAH-5DS	5DS	TGACv1_scaffold_458278_5DS	9741-14262	forward		TRIAE_CS42_5DS_TGACv1_458278_AA1493010	1395	GAGGCAATCAGTGTGCTCCAAAG	AATTGCACAACACAGAAAAACAGCAG	1517	N5A-T5D	N5B-T5B	N5D-T5A	60

Table S3-2. Percentage of nucleic acid identity between the coding sequences of wheat purine catabolism gene.

Gene	XDH1-1AL	XDH1-1BL	XDH1-1DL	XDH2-6AS	XDH2-6BS	XDH2-6DS	UOX-3AL	UOX-3B	UOX-3DL	AS-4AS	AS-4BL	AS-4DL	ALN-2AL	ALN-2BL	ALN-2DL	AAH-7AL	AAH-7BL	AAH-7DL	UGAH-2AS	UGAH-2BS	UGAH-2DS	UAH-5AS	UAH-5BS	UAH-5DS
XDH1-1AL	100																							
XDH1-1BL	98.8	100																						
XDH1-1DL	98.7	98.6	100																					
XDH2-6AS	89.7	89.6	89.9	100																				
XDH2-6BS	90.4	90.5	90.4	95.3	100																			
XDH2-6DS	90	90.1	90.2	93.5	93.9	100																		
UOX-3AL							100																	
UOX-3B							97.3	100																
UOX-3DL							96.3	95.6	100															
AS-4AS										100														
AS-4BL										97.8	100													
AS-4DL										95.9	96.8	100												
ALN-2AL													100											
ALN-2BL													95	100										
ALN-2DL													95	94.5	100									
AAH-7AL																100								
AAH-7BL																95	100							
AAH-7DL																93.1	92.9	100						
UGAH-2AS																			100					
UGAH-2BS																			99	100				
UGAH-2DS																			98.3	98	100			
UAH-5AS																						100		
UAH-5BS																						98.1	100	
UAH-5DS																						97.9	98.1	100

Table S3-3. Prediction of subcellular localisation of wheat purine catabolic protein. Protein sequences were obtained by translating the predicted CDSs associated with the gene models presented in Table S3-1. For the analysis were employed WoLF PSORT (wolfpsort.hgc.jp) and TargetP1.1 (www.cbs.dtu.dk/services/TargetP/) prediction tools. For the former, the numbers indicate the number of proteins in the WoLF PSORT training data that have the most similar localization features. For the latter, highest score indicates the most likely localisation; C, chloroplast; M. mitochondrion; SP, secretory pathway.

Protein name	Length	WoLF PSORT	TargetP 1.1					
			cTP	mTP	SP	other	Loc	RC
TaXDH1-1AL	**1369**	chlo: 4, nucl: 3.5, plas: 3, cysk_nucl: 2.5, vacu: 2, cyto: 1	0.119	0.167	0.134	0.747	_	3
TaXDH1-1BL	**1369**	plas: 10, E.R.: 2, nucl: 1, vacu: 1	0.116	0.181	0.09	0.826	_	2
TaXDH1-1DL	**1369**	plas: 11, E.R.: 2, nucl: 1	0.121	0.176	0.092	0.823	_	2
TaXDH2-6DS	**1372**	chlo: 4, plas: 3, nucl: 2.5, vacu: 2, cysk_nucl: 2, cyto: 1, E.R.: 1	0.065	0.235	0.079	0.754	_	3
TaUOX-3AL	**307**	pero: 6, chlo: 4, cyto: 3, nucl: 1	0.03	0.759	0.018	0.416	M	4
TaUOX-3BL	**307**	pero: 6, chlo: 4, cyto: 3, nucl: 1	0.034	0.775	0.02	0.35	M	3
TaUOX-3DL	**307**	pero: 7, cyto: 3, chlo: 2, nucl: 1, golg: 1	0.03	0.664	0.023	0.553	M	5
TaAS-4AS	**332**	chlo: 8, nucl: 3.5, nucl_plas: 3, plas: 1.5, mito: 1	0.071	0.552	0.025	0.322	M	4
TaAS-4BL	**332**	chlo: 7, nucl: 4, mito: 2, plas: 1	0.1	0.547	0.019	0.186	M	4
TaAS-4DL	**332**	chlo: 6, nucl: 5, mito: 2, plas	0.109	0.574	0.016	0.263	M	4
TaALN-2AL	**505**	extr: 3, vacu: 3, E.R.: 3, chlo: 2, nucl: 1, mito: 1, golg: 1	0.009	0.184	0.82	0.013	S	2
TaALN-2BL	**505**	chlo: 3, vacu: 3, E.R.: 3, extr: 2, nucl: 1, mito: 1, golg: 1	0.01	0.182	0.82	0.013	S	2
TaALN-2DL	**505**	chlo: 3, vacu: 3, E.R.: 3, extr: 2, nucl: 1, mito: 1, golg: 1	0.009	0.184	0.82	0.013	S	2
TaAAH-7AL	**494**	extr: 5, vacu: 4, golg: 2, chlo: 1, mito: 1, E.R.: 1	0.076	0.007	0.988	0.015	S	1
TaAAH-7BL	**495**	vacu: 6, chlo: 3, mito: 2, extr: 2, nucl: 1	0.131	0.067	0.773	0.045	S	2
TaAAH-7DL	**487**	extr: 6, vacu: 3, golg: 2, chlo: 1, mito: 1, E.R.: 1	0.033	0.006	0.989	0.022	S	1
TaUGAH-2AS	**302**	cyto: 5.5, cyto_E.R.: 4, extr: 3, chlo: 2, E.R.: 1.5, nucl: 1, vacu: 1	0.026	0.025	0.937	0.044	S	1
TaUGAH-2BS	**302**	cyto: 5.5, cyto_E.R.: 4, extr: 3, chlo: 2, E.R.: 1.5, nucl: 1, vacu: 1	0.021	0.025	0.938	0.071	S	1
TaUGAH-2DS	**302**	E.R.: 5.5, E.R._plas: 4.5, plas: 2.5, nucl: 2, mito: 2, cyto: 1, vacu: 1	0.031	0.017	0.927	0.079	S	1
TaUAH-5AS	**465**	E.R.: 4.5, extr: 4, E.R._plas: 3, mito: 2, chlo: 1, cyto: 1, vacu: 1	0.108	0.11	0.391	0.153	S	4
TaUAH-5BS	**465**	E.R.: 4.5, extr: 4, mito: 3, E.R._plas: 3, chlo: 1, cyto: 1	0.092	0.094	0.5	0.182	S	4
TaUAH-5DS	**465**	E.R.: 4.5, extr: 4, mito: 3, E.R._plas: 3, chlo: 1, cyto: 1	0.086	0.166	0.366	0.166	S	5

Table S3-4. Details of the genes used for phylogenetic analysis. Gene ID are specified for Arabidopsis, *Zea mays*, *Sorghum bicolor*, *Oryza sativa* and *Brachypodium distachyon*. Scaffold or contig ID where gene were identified for *Hordeum vulgare*, *Triticum urartu* and *Aegilops tauschii*, *Triticum durum* and *Triticum aestivum* cv. Cadenza.

Organism	Gene name		Gene or scaffold ID	Source
Arabidopsis thaliana	AtXDH1		At4g34890	Phytozome v11
	AtXDH2		At4g34900	Phytozome v11
	AtUOX		At2g26230	Phytozome v11
	AtAS		At5g58220	Phytozome v11
	AtALN		At4g04955	Phytozome v11
	AtAAH		At4g20070	Phytozome v11
	AtUGAH		At4g17050	Phytozome v11
	AtUAH		At5g43600	Phytozome v11
Sorghum bicolor	SbXDH		Sobic.001G325200	Phytozome v11
	SbURICASE		Sb03g040810	Phytozome v11
	SbAS		Sb01g033540	Phytozome v11
	SbALN		Sb06g033160	Phytozome v11
	SbAAH		Sb10g026590	Phytozome v11
	SbUGAH		Sb02g033970	Phytozome v11
	SbUAH		Sb08g020250	Phytozome v11
Zea mays	ZmXDH		GRMZM2G050984	Phytozome v11
	ZmURICASE		GRMZM2G164141	Phytozome v11
	ZmAS		LOC100273769	Phytozome v11
	ZmALN		GRMZM2G173413	Phytozome v11
	ZmAAH		LOC100284878	Phytozome v11
	ZmUGAH		LOC100217019	Phytozome v11
	ZmUAH		GRMZM2G130034	Phytozome v11
Oryza sativa	OsXDH		LOC_Os03g31550	Phytozome v11
	OsURICASE		LOC_Os01g64520	Phytozome v11
	OsAS		LOC_Os03g27320	Phytozome v11
	OsALN		LOC_Os04g58390	Phytozome v11
	OsAAH		LOC_Os06g45480	Phytozome v11
	OsUGAH		LOC_Os07g31270	Phytozome v11
	OsUAH		LOC_Os12g40550	Phytozome v11
Brachypodium distachyon	BdXDH		Bradi1g15800	Phytozome v11
	BdURICASE		Bradi2g55950	Phytozome v11
	BdAS		Bradi1g60700	Phytozome v11
	BdALN		Bradi5g26410	Phytozome v11
	BdAAH		Bradi1g31890	Phytozome v11
	BdUGAH		Bradi1g27060	Phytozome v11
	BdUAH1		Bradi4g02700	Phytozome v11
	BdUAH2		Bradi2g19950	Phytozome v11
Hordeum vulgare	HvXDH		morex_contig_567184; 1586860; 1561442; 1571616; 424532; 1559465; 432775;	Barley WGS Morex Assembly version 3
	HvUOX		morex_contig_98538; 175336	Barley WGS Morex Assembly version 3
	HvAS		morex_contig_42553	Barley WGS Morex Assembly version 3
	HvALN		morex_contig_66029	Barley WGS Morex Assembly version 3
	HvAAH		morex_contig_41890	Barley WGS Morex Assembly version 3
	HvUGAH		morex_contig_2548646; 1641800; 2546598; 1671223; 1576585	Barley WGS Morex Assembly version 3
	HvUAH		morex_contig_58442; 1557988	Barley WGS Morex Assembly version 3
Triticum urartu	uraXDH1		Tururtu_scaffold54488	Mayer et al., 2014
	uraXDH2	exon i-xiii	Tururtu_scaffold66678; 56329	Mayer et al., 2014

	uraUOX		TGAC_WGS_urartu_v1_contig_169717	Ling et al., 2013
	uraAS		TGAC_WGS_urartu_v1_contig_168051	Ling et al., 2013
	uraALN		TGAC_WGS_urartu_v1_contig_165835	Ling et al., 2013
	uraAAH		TGAC_WGS_urartu_v1_contig_40563; 238137	Ling et al., 2013
	uraUGAH		Tururtu_scaffold52848	Mayer et al., 2014
	uraUAH		TGAC_WGS_urartu_v1_contig_164271	Ling et al., 2013
Aegilops tauschii	tauXDH1		Atauschii_scaffold45203; 72244	Mayer et al., 2014
	tauXDH2		Atauschii_scaffold106251	Mayer et al., 2014
	tauUOX		Atauschii_scaffold58678	Mayer et al., 2014
	tauAS		Atauschii_scaffold74235	Mayer et al., 2014
	tauALN		TGAC_WGS_tauschii_v1_contig_118195	Jia et al., 2013
	tauAAH		Atauschii_scaffold23719	Mayer et al., 2014
	tauUGAH		Atauschii_scaffold25606	Mayer et al., 2014
	tauUAH		TGAC_WGS_tauschii_v1_contig_154375	Jia et al., 2013
Triticum durum (cv. Kronos)	TdXDH2-A (Kronos)	exon i-xiii	Triticum_turgidum_Kronos_EIv1_scaffold_062030	Clavijo et al., 2016
	TdXDH2-B (Kronos)	exon ii, iii, iv	Triticum_turgidum_Kronos_EIv1_scaffold_038360	Clavijo et al., 2016
Triticum aestivum (cv. Cadenza)	TaXDH2-6AS-Cad	exon ii-xiii	Triticum_aestivum_Cadenza_EIv1_scaffold_396064_6AS	Clavijo et al., 2016
	TaXDH2-6BS-Cad	exon ii, iii, iv	Triticum_aestivum_Cadenza_EIv1_scaffold_420391_6BS	Clavijo et al., 2016
	TaXDH2-6DS-Cad		Triticum_aestivum_Cadenza_EIv1_scaffold_442587_6DS	Clavijo et al., 2016

Table S3-5. Details of primers employed for quantitative reverse transcription PCR (qRT-PCR) analysis.

Gene name	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Size product (bp)	Tm (°C)
TaXDH1	ACAGTGAAAATCGTTGGAGGA	GCAACTTGGGCTATCTTTGTG	187	60
TaXDH2	CATTGGAATCGTTGTGGCAGAC	GCTCTTGACCTCCGACTTGAAC	233	60
TaUOX	GTGAAGAAGTCTGGAAGCC	CAGGAAGAAGAGTGTAGCG	115	60
TaAS	CTGGAGATGTGGAAGGAC	TGGAGGTATTGAAACTTATGC	172	60
TaALN	TCGTTTCAAGTGCTCTCC	AGTTGCCTTCCTCCATTAG	143	63
TaAAH	TCTCTGCTCTGAAGGTCTTG	CTCCTCGTCGCTGAATGC	86	60
TaUGAH	TTAGGAGCATATTTGATTAGCC	GCAGGAAGATAAGCATAAGAG	200	60
TaUAH	TGTAACTGCCATTGCTGCTC	CAACTGTTCCAACTGTATCTATCG	179	63
TaActin	GACAATGGAACCGGAATGGTC	GTGTGATGCCAGATTTTCTCCAT	236	60
TaGAPdH	TTCAACATCATTCCAAGCAGCA	CGTAACCCAAAATGCCCTTG	197	60
TaEFA2	CAGATTGGCAACGGCTACG	CGGACAGCAAAACGACCAAG	227	60

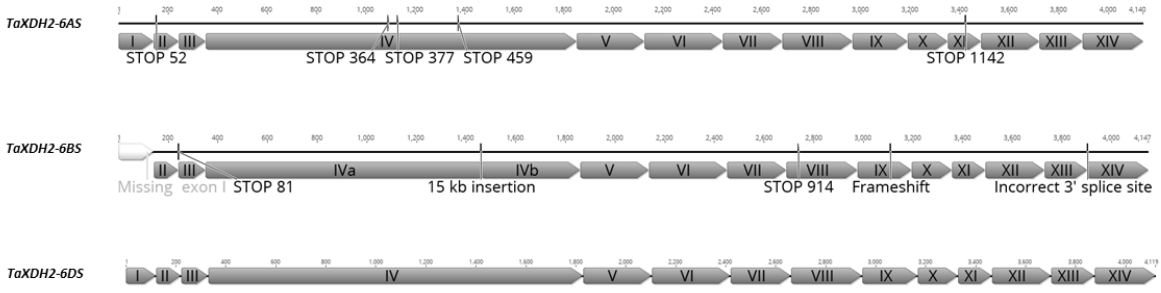


Fig. S3-1. TaXDH2 homeologs coding sequence (CDS) structure. At the top, putative CDS structure on wheat chromosome 6AS (*TaXDH2-6AS*). In the middle, putative CDS structure of *TaXDH2-6BS.* At the bottom, putative CDS structure of TaXDH2-6DS. Features that alter the open reading frame of the genes are indicated as vertical black bar and related details are provided below it. Figure were prepared with the support of Geneious 10.0.2 software.

Figure S3-2. **Genome A-XDH2 translated putative CDS alignment from hexaploid bread wheat cv. Chinese Spring and Cadenza, tetraploid durum wheat cv. Kronos and diploid *Triticum urartu* genomes**. Gene models of XDH2 genes associated with wheat A genome were translated into amino acids sequences and aligned with Geneious 10.0.2 software (for accessions details see Table S3-1 and Table S3-4). Premature stop codons are highlighted with red square lines with the purpose of remarking the evolutionary linkage of the XDH2-A gene degeneration. Figure was prepared with the support of Geneious 10.0.2 software.

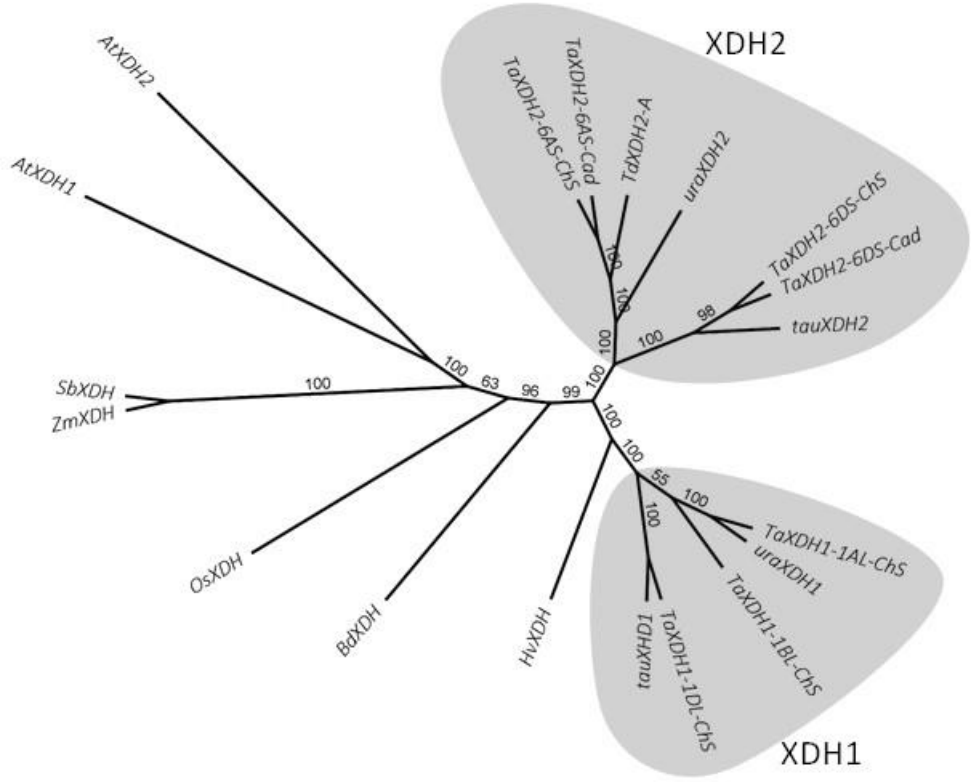


Figure S3-3. **Phylogenetic analysis of XDH among grass genomes.** Unrooted tree with cladogram transformed branches presenting XDH gene coding sequences (CDS) for the selected grasses genomes and *Arabidopsis*. CDS were aligned using Jalview2 MAFFT algorithm with default parameters (http://www.jalview.org/; FFT-NS-2 algorithm). The alignment was edited to adjust for the codon structure of the CDS and trimmed at the 5' and 3' ends. Phylogenetic trees were constructed using PhyML (http://www.atgc-montpellier.fr/phyml) with HKY85 as substitution model, SPR & NNI as tree improvement, and a boostrap value of 100. The output in Newick tree format was visualised as an unrooted tree with cladogram transformed branches using Geneious 10.0.2. Wheat and wheat ancestors subgroups are highlighted with the grey shade and divided into two clades: XDH1 and XDH2. Bootstrap values are reported on tree branches.

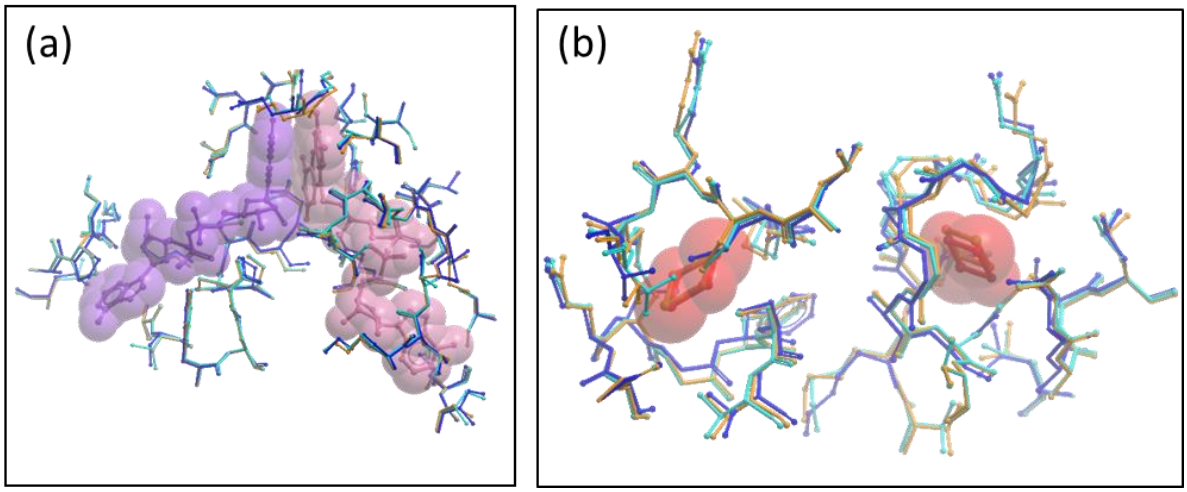


Figure S3-4. TaXDH1-1DL and XDH2-6DS have highly conserved FAD-binding site and 2Fe-2S clusters. (A) NADH (purple) and FAD (pink) are shown in sticks and spheres and residues that line the FAD-binding site are shown in sticks. Overlayed structures show the high conservation in the pocket with all hydrogen bonding interactions maintained in TaXDH1-1DL (cyan) and TaXDH2-6DS (orange) when compared to the 3UNI bovine XDH1 (blue). Structures are also conserved with rat XDH1 (not shown). (**B**) 2Fe-2S clusters are shown in red spheres. Residues that coordinate binding are represented in sticks. Overlayed structures show the high conservation in the pocket with all interactions maintained in TaXDH1-1DL (cyan) and TaXDH2-6DS (orange) when compared to the 3UNI bovine XDH (blue). Structures are also conserved with rat XDH1 (not shown).

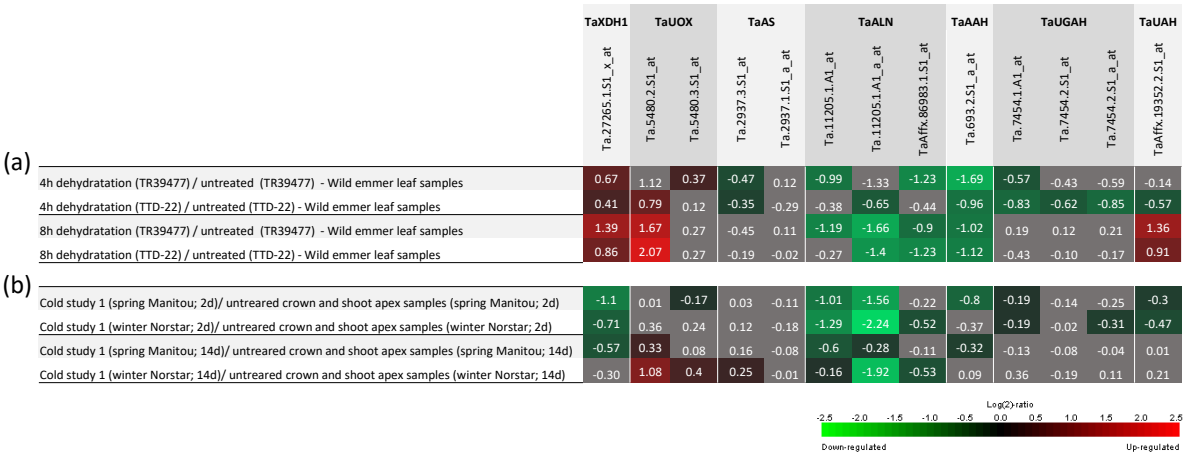


Figure S3-5. Expression data of purine catabolism genes from data mining of publicly available data bases. Wheat probes were identified for Affimetrix Wheat 61k Microarray Platform by blasting the gene CDS onto the "Find Your Gene" tool available on PLEXdb website (http://plexdb.org/). All the probes that were identified associated with a specific gene were used for data mining. For XDH genes, the probe a.27265.1.S1_x_at displayed >97% of identity with TaXDH1 and 90% identity with XDH2-6DS, so it was considered to be associated with TaXDH1. The probes were used to search Genevestigator 5-12-03 perturbation condition search tool and the fold change as Log(2) ratio of each probe for selected experiments is reported here. Colour-coding (green to red) indicates down- or up-regulation and p-value < 0.05 of a certain probe, while grey values indicate p-value > 0.5. **(A)** leaf samples of two lines of wild emmer dehydrated from hydroponic solution for 4 and 8 hours (Array Express database with accession E-MEXP-1488); **(B)** crown and shoot apex samples of spring and winter wheat after 2 and 14 days of cold treatment at 6 °C (NCBI Gene Expression Omnibus database with accession GSE23889).

Chapter 4: The dual role of the purine catabolic intermediate allantoin under stress and N homeostasis in bread wheat

Statement of Authorship

Title of Paper	**The dual role of the purine catabolic intermediate allantoin under stress and N homeostasis in bread wheat**
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Manuscript prepared in accordance with the guidelines for the Journal of Experimental Botany

Principal Author

Name of Principal Author (Candidate)	Alberto Casartelli		
Contribution to the Paper	Executed the study, performed analysis on all samples, interpreted data, wrote manuscript		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	

Co-Author Contributions

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

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr. Vanesa J. Melino		
Contribution to the Paper	Supervised development of work, help in data interpretation and manuscript evaluation and editing		
Signature		Date	10/10/17

Name of Co-Author	Dr. Matteo Riboni		
Contribution to the Paper	Performed qRT-PCR analysis and provided support with its interpretation		
Signature		Date	26/10/2017

Name of Co-Author	Himasha Mendis		
Contribution to the Paper	Provided support in the development of allantoin quantification method and inputs in the materials and method section		
Signature		Date	17/10/17

Name of Co-Author	Prof. Ute Roessner		
Contribution to the Paper	Provided support and interpretation of allantoin analyses and was involved in supervising the development of work		
Signature		Date	10/10/17

Name of Co-Author	Dr. Rainer Hoefgen		
Contribution to the Paper	Provided support for the metabolomics analyses and evaluation of the manuscript		
Signature		Date	25.10.2017

Name of Co-Author	Dr. Mamoru Okamoto		
Contribution to the Paper	Supervised development of work, help in data interpretation and manuscript evaluation and editing		
Signature		Date	26/10/2017

Name of Co-Author	Dr. Sigrid Heuer		
Contribution to the Paper	Supervised development of work, help in data interpretation and manuscript evaluation and editing		
Signature		Date	10/10/17

Title

The dual role of the purine catabolic intermediate allantoin under stress and N homeostasis in bread wheat.

Alberto Casartelli[1], Vanessa Melino[1], Matteo Riboni[1], Himasha Mendis[2], Ute Roessner[2], Rainer Hoefgen[3], Mamoru Okamoto[1] and Sigrid Heuer[1,4*]

Affiliations:

[1] School of Agriculture Food and Wine, The University of Adelaide, Urrbrae, SA 5064, Australia

[2]Metabolomics Australia, The University of Melbourne, Parkville, VIC 310, Australia

[3] Max Plank Institute for Molecular Plant Physiology, 14476 Potsdam - Golm, Germany

[4] Plant Biology and Crop Science department, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UK

*Corresponding author: sigrid.heuer@rothamsted.ac.uk

Abstract

Allantoin is an intermediate of the purine catabolic pathway that recycles the N present in the purine ring. Recent studies have revealed that purine catabolism is responsive to several abiotic stresses in many plant species. In certain tropical legumes allantoin is an important N form used for translocation and storage, however, its role in grasses such as wheat has been poorly characterised. Here we describe the analysis of two Australian bread wheat genotypes showing that allantoin accumulates in leaves under drought, up to 29-fold greater than leaves from plants grown under well-watered conditions. In contrast,

allantoin levels were reduced by up to 22-fold in plants grown under N deficiency as compared to plants grown with sufficient N supply. Tissue specific accumulation of allantoin was accompanied by transcriptional changes of key purine catabolic genes, namely *TaXDH1*, *TaUOX*, *TaALN* and *TaAAH*. In addition, we show that wheat plants can grow with xanthine and allantoin as the sole N source as efficiently as nitrate suggesting that the N supplied by purine degradation can support plant growth. These results indicate that accumulation or degradation of purines and more specifically allantoin in wheat is involved in maintaining optimal N homeostasis under drought and N deficiency.

Introduction

Plants can synthesise an enormous array of chemical compounds and it is estimated that a single plant can produce up to 25,000 metabolites at any given time (Verpoorte, 1998; Oksman-Caldentey and Inzé, 2004). With the advance of metabolomics techniques, that allow the simultaneous monitoring of thousands of compounds, it is becoming evident that small molecules previously believed as mere components of the plant central metabolism have additional roles as osmolytes, reactive oxygen species (ROS) scavengers and signalling molecules and, thus, to actively contribute toward plant survival under abiotic stresses (for example see, Bouche and Fromm, 2004; Szabados and Savoure, 2010; Li *et al.*, 2015). This metabolic plasticity reflects the highly responsive mechanisms of adaptation that plants, given their sessile nature, have evolved to cope with environmental changes, such as shortage of water or nutrients. Therefore, the beneficial role of specific metabolites under stress can potentially be exploited for crop improvements in order to increase the productivity of agricultural systems that are currently under threat of global warming (Porter *et al.*, 2014). Vanrensburg *et al.* (1993) proposed to use proline accumulation as a criteria to screen tobacco genotypes for drought tolerance; this was supported more recently by Vendruscolo et al. (2007) who showed that transgenic wheat that accumulated proline under the control of a stress-induced promoter had increased tolerance to drought.

Additionally, Griffith et al. (2016) showed that an analogue of the signalling molecule trehalose-6-phosphate employed as an agrochemical successfully improved drought tolerance and, thus, may represent an alternative approach to genetic improvement. However, metabolites are less robust than DNA-based markers and, thus, metabolites markers to be employed in breeding activities need to be thoroughly tested in large number of accessions and under different field conditions.

A small nitrogen (N) rich metabolite allantoin has recently gained attention by the scientific community as several metabolomics studies reported this metabolite to differentially accumulate in a broad range of plant species upon drought (Bowne *et al.*, 2011; Oliver *et al.*, 2011; Silvente *et al.*, 2012; Degenkolbe *et al.*, 2013; Yobi *et al.*, 2013; Casartelli *et al.*, 2018), high salt (Kanani *et al.*, 2010; Wu *et al.*, 2012; Nam *et al.*, 2015; Wang *et al.*, 2016), cold (Kaplan *et al.*, 2004) and nutrient deficiency (Nikiforova *et al.*, 2005; Amiour *et al.*, 2012; Rose *et al.*, 2012; Coneva *et al.*, 2014). Allantoin is an intermediate of the purine catabolic pathway that mediates the degradation of molecules containing the purine heterocyclic ring (Fig. 4-1). Overall, the oxidation of one molecule of xanthine, the first intermediate of the pathway, liberates one molecule of glyoxylate, three molecules of CO_2 and four molecules of ammonium (NH_4^+). The liberated ammonium is presumed to be assimilated by glutamine synthetase (GS, EC 6.3.1.2) and glutamate synthase (GOGAT, EC 1.4.7.1), operating in the GS/GOGAT cycle (Fig. 4-1). Therefore, the N remobilised from the purine ring may contribute to internal N recycling under stress or senescence. (for comprehensive reviews see Zrenner *et al.*, 2006; Werner and Witte, 2011). Allantoin and its degradation product allantoate, collectively referred to as ureides, are well characterised in tropical legumes given their role as the main transport forms of symbiotically fixed dinitrogen (N_2) from root nodules to the shoots (Herridge *et al.*, 1978; Pate *et al.*, 1980; Schubert, 1986; Pèlissier and Tegeder, 2007). Ureides also translocate from storage organs to leaves and vice versa according to seasonal changes in deciduous trees, such as maple and comfrey, (Reinbothe and Mothes, 1962). Ureide accumulation under water stress in tropical legumes was first characterised in nodulated and more recently in non-nodulated plants

(Sinclair and Serraj, 1995; deSilva *et al.*, 1996; King and Purcell, 2005; Alamillo *et al.*, 2010; Coleto *et al.*, 2014).

The relevance of allantoin, and generally purine catabolism, under abiotic stress have been well characterised in *Arabidopsis thaliana*. Early works showed that loss of *AtXDH*, the gene coding for the xanthine-degrading enzyme xanthine dehydrogenase, disrupted allantoin production and induced early onset of senescence and reduced tolerance to drought stress (Nakagawa *et al.*, 2007; Watanabe *et al.*, 2010; Watanabe *et al.*, 2014a). Interestingly, Brychkova *et al.* (2008a) showed that the detrimental effects of prolonged dark exposure on Arabidopsis *xdh* mutants could be attenuated with exogenous application of allantoin and allantoate. On the other hand, Arabidopsis knock-out mutants in *AtALN*, the gene coding for the allantoin-degrading enzyme allantoinase, constitutively accumulate high levels of allantoin. These mutants showed better performances than wild type plants under different stresses such as desiccation, drought, osmotic, salt and heavy metal exposure. (Watanabe *et al.*, 2014b; Irani and Todd, 2016; Lescano *et al.*, 2016; Nourimand and Todd, 2016). Particularly, Irani and Todd (2016) showed that *Ataln* mutants accumulated lower amounts of ROS under water stress. However, since allantoin does not possess any *in-vitro* antioxidant activity (Wang *et al.*, 2012), it is likely to act indirectly. Watanabe *et al.* (2014b) and Takagi *et al.* (2016) showed that increasing allantoin levels in the plant by either loss of function of the *ALN* gene or by exogenously applying allantoin to the media activated the abscisic acid (ABA) and jasmonic acid (JA) pathways, which are well known for their role in plant stress-responses (Wasternack and Hause, 2013; Yoshida *et al.*, 2014).

Details on the role of allantoin and the purine catabolic pathway remain mostly unexplored in wheat. Early reports showed that *XDH* is involved in biotic stress responses to wheat leaf rust (Montalbini, 1992) and our recent work showed that wheat purine catabolic genes are differentially regulated under abiotic stress and wheat plants accumulated less allantoin in grain when grown under N restriction (Casartelli *et al*., Chapter 3). Given the interplay of allantoin between stress tolerance and N metabolism, the aim of the present study was to characterise purine catabolism under water and N restriction in two distinct

Australian bread wheat genotypes (RAC875 and Mace). Metabolic and transcriptomic analyses revealed that allantoin differentially accumulates in plant tissues under N and water stress conditions and that the purine catabolic genes act synergistically to regulate allantoin accumulation. We further show that the N recycled from the purine catabolism can support growth by supplying xanthine and allantoin as a sole N source to N starved wheat seedlings. Monitoring of N-metabolites pools in seeds undergoing grain filling revealed that allantoin progressively accumulated in grains and genotypic differences between the selected genotypes exists for allantoin and other important N-containing metabolites.

Materials and Methods

Plant material

Two semi-dwarf South Australian wheat genotypes, RAC875 and Mace, were evaluated in this study. RAC875 (RAC655/3/Sr21/4*LANCE//4*BAYONET), a breeding line from Roseworthy Agricultural Campus that has never been commercially released. RAC875 is high-yielding under drought and heat-prone South Australian environments (Izanloo *et al.*, 2008; Bennett *et al.*, 2012) and RAC875 seedlings have higher net N uptake efficiency than older CIMMYT genotypes (1966-1985) (Melino *et al.*, 2015). Mace (WYALKATCHEM/STYLET//WYALKATCHEM), was bred and released by Australian Grain Technologies (AGT) in 2008, preliminary studies suggest that Mace has high N use efficiency across different southern Australian environments (Mahjourimajd *et al.*, 2016).

Plant growth and sample collection

Plant growth in potting mixture

The experiment was conducted in a controlled environment with day/night cycle of 12h/12h at a flux density at canopy level of 300 µmol m^{-2} s^{-1}, 20°C/15°C day/night temperature and 82% average humidity.

Potting mixture was composed of river sand and coco-peat and prepared according to Melino *et al.* (2015). Granular urea was provided as basal N application with rates of 150 and 75 mg N kg^{-1} for high and low N treatments, respectively. At stem elongation (39 days after sowing) a third of the basal urea rates for each treatment were applied by soil drenching. Soil water retention curve was constructed by measuring the pre-dawn leaf water potential of three-week old seedlings under progressive drought stress with a Scholander-type pressure chamber (Soil Moisture Equipment Corp., Santa Barbara, USA) (Supplementary Fig. S4-1). Overall, the experiment was comprised of two cultivars (RAC875 and Mace), three treatments (high nitrogen well-watered (HN-WW), low nitrogen well-watered (LN-WW) and high nitrogen drought (HN-D)), six sampling time points and six biological replicates for a total of 216 experimental units. Individual seeds were germinated in black square plastic pots (11x11 cm area, 14 cm height) containing 1.5 kg of soil mix arranged in a randomised complete block design. WW conditions were maintained by watering every day the pots to weight at 20% soil water content (SWC), calculated as SWC = $[(m_{wet\ soil} – m_{dry\ soil}): m_{dry\ soil}]\times100$. Drought was induced by withholding water until signs of leaf rolling appeared (approximately 6.5% SWC; Table 4-1). Samples collection during reproductive stages were performed according to the anthesis date of individual plants. In particular, water was withheld starting from two weeks after anthesis. Details on the sample collection throughout the experiment are given in Table 4-1. Samples for molecular analysis were snap-frozen on liquid nitrogen and stored at -80°C until further use.

Growth in hydroponics

The hydroponics system was situated in a controlled environment room with a 12h/12h day/night cycle at 20°C/15°C and a flux density at canopy level of 300 µmol m^{-2} s^{-1}. Seedling germination and growth in 120L ebb-and-flow hydroponics system was according to Melino *et al.* (2015). Wheat seedlings (genotype Mace) were grown in completely randomised design under constant supply of KNO_3^- (1 mM N) for two weeks and then N starved (0 mM N, –N) for 24 h , while another set of replicate plants were maintained at sufficient N (1 mM N). After 24 hours the –N plants were transferred to 4 L hydroponic units (6 plants

each) where N was resupplied either as nitrate (rsNO_3^-), xanthine (rsXanth) or allantoin (rsAlnt), while a set of replicate plants were maintained under –N conditions. The nutrient solutions used in this experiment were a modified Johnson's solution (Johnson *et al.*, 1957). Specifically, the NO_3^- nutrient solutions were composed of 1 mM N as NO_3^-, 1.05 mM K, 0.25 mM Ca, 0.5 mM Mg, 0.50 mM S and 0.5 mM of P. The remaining treatments were composed of 1.05 mM K, 0.25 mM Ca, 0.5 mM Mg, 0.75 mM S and 0.5 mM and 0.5 mM of P plus 0.25 mM xanthine (rsXanth) or 0.25 mM allantoin (rsAlnt). All solutions contained: 2 µM Mn, 2 µM Zn, 25 µM B, 0.5 µM Cu, 0.5 µM Mo and 0.1 mM Fe (as FeEDTA). Nutrient solutions were maintained at pH 5.9 and replaced every 7 days when plants were growing in 120L ebb-and-flow hydroponic system and every 2 days when grown in 4L hydroponic units. Two independent experiments were carried out and an overview of the sample collection is provided in Table 4-1.

Allantoin GC-QqQ-MS analysis

Metabolite extraction

Metabolites were extracted from 10 mg of homogenised, freeze-dried tissue with 500 µl of 100% (v/v) methanol containing 12.5 µM $^{13}C^{15}N$-allantoin, expect mature grains samples for which the concentration used was 25 µM $^{13}C^{15}N$-allantoin. Samples were vortexed and incubated in an Eppendorf Thermomixer at 1400 rpm and 30°C for 15 minutes followed by a 15 min centrifugation at 13,000 rpm (4°C).The supernatant was transferred to a new tube. 500 µl of milli-Q Water was added to the remaining sample pellet, vortexed and centrifuged for 15 min at 15,000 rpm. The supernatant was combined with the previous one, vortexed for 30 sec and centrifuged for 15 min at 15,000 rpm. The resulting supernatant was transferred to a new tube and 300 µl of 100% (v/v) chloroform was added, vortexed and centrifuged for 5 min at 15,000 rpm. 800 µl of the top (polar) phase was transferred into a new tube for allantoin analysis.

Sample derivatisation

200 µl of the upper polar phase was aliquoted in a glass insert and dried under vacuum (RVC 2-33 CD plus, John Morris Scientific Australia) set at ambient temperature. All samples were re-constituted in 10 µl of methoxyamine hydrochloride (30 mg ml^{-1} pyridine) and derivatised at 45°C for 60 min at 500 rpm before adding 20 µl of N-methyl-N-(tert-butyldimethylsilyl)trifluoroacetamide (MTBSTFA) with 1% (w/v) trimethyl chlorosilane (TMCS) and incubated at 45°C for 45 min.

GC-MS instrument conditions

1 µl of derivatised sample was injected into a GC-QqQ-MS system comprised of a Gerstel 2.5.2 Autosampler, a 7890A Agilent gas chromatograph and a 7000 Agilent triple-quadrupole MS (Agilent, Santa Clara, USA) with an electron impact (EI) ion source. The GC was operated in constant flow mode with helium as the carrier gas. The MS was adjusted according to the manufacturers' recommendations using tris-(perfluorobutyl)-amine (CF43). A J&W Scientific VF-5MS column (30 m long with 10 m guard column, 0.25 mm inner diameter, 0.25 µm film thickness) was used. The injection temperature was set at 250°C, the MS transfer line at 290°C, the ion source was adjusted to 230°C and the quadrupole at 150˚C. Helium was used as the carrier gas at a flow rate of 1 ml min^{-1}. Nitrogen (UHP 5.0) was used as the collision cell gas at a flow rate of 1.5 ml min^{-1}. Helium (UHP 5.0) was used as the quenching gas at a flow rate of 2.25 ml min^{-1}. Gain factor for the triple axis detector set for 2. Derivatised sample was injected into the column at 100°C followed by 1 min hold followed by a ramp of 25°C min^{-1} to 325 ˚C.

Method optimization

The allantoin standard and the internal standard ($^{13}C^{15}N$-Allantoin) were purchased from Sigma Aldrich (Australia). Standards were analysed on the GC-QqQ-MS to obtain retention times and to identify a corresponding unique, precursor ion. For each precursor ion, product ion scans were carried out to identify distinct product ions. For each standard, the collision energy was optimized using a series of collision energies between 0 and 20 for the identified precursor to product ion transitions (Multiple Reaction

Monitoring (MRM). Once collision energies were optimized for each MRM transition, a product ion was selected as the Target ion (T) and the other subsequent MRM transition was set as the qualifier ion (Q) (Table S4-2). Linearity of the allantoin MRM transition was evaluated by preparing serial dilutions of the calibration standard ranging from 1 mM to 0.98 µM and showed a linear calibration range from 0.98 µM to 250 µM. The calibration curve created for allantoin was fitted using linear regression and the linear correlation coefficient for the calibration curve (R^2) was 0.99 for the target allantoin MRM (398 → 171). The limit of detection (LOD) for the allantoin MRM, was 0.294 µM based on a signal to noise ratio of 3.

Data was processed using Agilent MassHunter QQQ Quantitative Analysis software (B.07.00). Allantoin was quantified by single point calibration based on the relative response (the response area of allantoin (MRM 398→171) divided by the response area of $^{13}C^{15}N$-allantoin (MRM 400→173)) and $^{13}C^{15}N$-allantoin concentration. Sample dry weight and extraction volume were taken into consideration when calculating the final allantoin concentration.

Metabolome-wide analysis

Metabolites were extracted from 60 mg dry weight (DW) of freeze-dried wheat samples and polar phase aliquots were analysed by GC-time of flight (TOF)-MS as described previously (Lisec *et al.*, 2006; Erban *et al.*, 2007). Ion chromatography (IC) analysis were performed according to (Moschen *et al.*, 2016). Free spermidine was quantified according to Do *et al.* (2013).

Amino acid analysis

Amino acids were extracted from 50 mg DW of freeze-dried wheat samples with 1 ml of 10 mM sodium acetate containing 250 nmol ml^{-1} Norvaline (internal standard). Amino acids were quantified on a Waters Acquity™ UPLC system using the Waters AccQ-Tag Ultra Chemistry Kit following the manufacturer's

instructions (Waters Corp., USA). Chromatograms were analysed with Empower® 3 Software (Waters Corp., USA).

Total N analysis

Total N concentration was determined using the Elementar rapid N exceed® (Elementar Analysensysteme GmbHanalyser, Germany) using 75-100 mg DW of homogenised wheat grain and leaf samples. Aspartic acid (250 mg) was used as standard for calibration: the theoretical aspartic acid N% (10.52) was divided by the N% measured by the instrument generating a N factor. The N factor was then used to correct the N% measured for each sample. Shoot N content was calculated as follow: Shoot N% × Shoot DW (g).

RNA preparation and qRT-PCR analysis

RNA extraction, cDNA synthesis and qRT-PCR analysis were performed according to Casartelli *et al*. (Chapter 3). Change in gene expression were calculated using qBASE+ software and reported as calibrated normalised relative quantities (CNRQ) that represents the relative gene expression level between different samples for a given target gene:

$$NRQ = \frac{E_{goi}^{\Delta Ct, goi}}{\sqrt[f]{\prod_{o}^{f} E_{ref_o}^{\Delta Ct, ref_o}}}$$

E: efficiency
ΔCt: delta-Ct
Ct: cycle threshold
goi: gene of interest
ref: reference

NRQ is then divided by a calibration factor (CF) (Hellemans *et al.*, 2007). Four reference genes were quantified by qRT-PCR: *TaActin*, *TaGAPdH*, *TaCyclophilin* and *TaEFA2*. CNRQ values were calculated using the most stable genes within a specific tissues (selected by qBASE+ software). The primer sequences are provided in Supplementary Table S4-1.

Photosynthetic rate and chlorophyll measurements

Photosynthetic rates were measured using a portable LI-6400XT gas analysis system with a fluorescence chamber (*Li-Cor*, Lincoln, NE, USA) set with the following parameters: 25°C leaf temperature, 300 µmol s^{-1} CO_2 flow rate to the sample cell, 400 µmol CO_2 mol^{-1} reference cell CO_2, 1500 PAR (µmol m^{-2} s^{-1}). Measurements were performed on the youngest fully emerged leaf (YFEL) of the main stem between 10:00 AM and 02:00 PM. Photosynthetic rate values were calculated with an adjusted area that accounted for leaf width that was estimated from scanned leaf pictures using ImageJ software (imagej.nih.gov/ij/). Chlorophyll content was measured using a portable chlorophyll meter (SPAD-502, Konica Minolta, Tokyo, Japan). Measures were taken by averaging reads at the bottom, the centre and the tip of the YFEL.

Data analysis

Statistical analyses were performed using GraphPad Prism version 7.00 for Windows (GraphPad Software, La Jolla California USA. www.graphpad.com) and GenStat for Windows 18th Edition (VSN International, Hemel Hempstead, UK. GenStat.co.uk). All data are reported as mean ± SEM. Significant differences between means of two groups of data were tested by Student's *t*-test. Significant differences between means of more than two groups were tested by either one-way or two-way ANOVA. For metabolomics data, the normality of each metabolite distribution were assessed and transformed were required prior to performing statistical tests. Transcriptional data was reported as log2 ratio of the fold-change between treatment (drought or low N) and control conditions.

Results

Phenotypic effect of drought stress and N deficiency on plant development

For this study we selected two South Australian bread wheat genotypes, namely RAC875 and Mace (see M&M for details). Plants were initially grown under well-watered (WW) conditions divided in two subsets supplied with either high N (HN) or low N (LN). During the course of the experiment a subset of plants grown under HN was subjected to drought (D) at tillering and grain filling. (Table 4-1). Therefore, the experiment comprised of three treatments: control (high N, well-watered; HN-WW), low N (low N, well-watered; LN-WW) and drought (high N, drought; HN-D). Agronomic traits at plant maturity were measured and analysed by two-way ANOVA to assess the impact of the different treatments and genotypes on plant growth and yield (Table 4-2). Changes in above ground biomass were significant for treatment, in fact this trait was reduced in both genotypes by about 20% and 15% under low N (LN-WW) and drought (HN-D), respectively, as compared to control (HN-WW) (Table 4-2). Changes in grain yield (GY) were significant for treatment and genotype × treatment. In particular, both RAC875 and Mace showed 25% reduction of grain yield under low N (LN-WW) conditions as compared to control (HN-WW). Under drought (HN-D), Mace showed higher grain yield than RAC875 ($p<0.02$ by Student's *t*-test) that corresponded to 16% and 27% reduction for Mace and RAC875, respectively, relative to control conditions (HN-WW) (Table 4-2). The harvest index (HI) was reduced by 4.2% in RAC875 and by 2.4% in Mace under low N (LN-WW), whilst under drought (HN-D) RAC875 HI was reduced by 13% and Mace HI only marginally reduced by 2%. (Table 4-2).

Allantoin accumulation under low N and drought stress

To assess whether N and water stress alter allantoin levels in the two wheat genotypes, allantoin concentration was quantified in the youngest fully emerged leaf (YFEL) during vegetative growth and in flag leaf, stem, spikelet and developing grain during reproductive growth in plants grown under control,

low N and drought conditions (Fig. 4-2A, B). Nitrogen deficiency significantly reduced allantoin concentration at most of the time points across different tissues and in both genotypes when compared to control conditions (Fig. 4-2A). In both genotypes grown under low N conditions allantoin levels were generally below 50 nmol g^{-1}DW in YFEL, flag leaf and stem, and much higher (50 to 100 nmol g^{-1}DW) in spikelet and developing grain. The highest reduction of allantoin in both genotypes was measured in the stem under low N conditions as compared to HN control, specifically 22-fold and 10-fold reduction in RAC875 and Mace, respectively (Fig. 4-2A). Interestingly, under control conditions Mace flag leaves accumulated high levels of allantoin (652 nmol g^{-1}DW) when compared to RAC875 (47 nmol g^{-1}DW), whilst a relatively smaller difference was recorded under low N (52 and 18 nmol g^{-1}DW for Mace and RAC875 respectively) (Fig. 4-2A). Contrasting to the reduced accumulation of allantoin under nitrogen deficiency, allantoin accumulated under drought in all tissues assessed (Fig. 4-2B). During the tillering stage, allantoin levels were 2-fold greater in RAC875 YFEL at mild drought stress (11.4% soil water content; SWC) as compared to control, whilst YFEL allantoin levels were 3- and 2.5-fold greater in droughted RAC875 and Mace plants showing signs of leaf rolling (day 8, 6.3% SWC) as compared to the WW control (Fig. 4-2B). Similarly, during grain filling, allantoin significantly accumulated in flag leaves and spikelets in both genotypes at day 5 and 8 (8.6% and 6.6% SWC, respectively) (Fig. 4-2B). Allantoin accumulated in stem and developing grain only at day 8, when drought stress was most severe (Fig. 4-2B). Interestingly, flag leaves of RAC875 grown under drought conditions, progresssively accumulated increasing amounts of allantoin from 136 nmol g^{-1} DW at day 5 to 1197 nmol g^{-1} DW at day 8, representing an almost 30-fold increase relative to WW control. On the other hand, Mace accumulated 1614 nmol of allantoin per gram DW after only 5 days of drought treatment and levels remained constant until day 8. However, allantoin levels only increased 2.5-fold in Mace under drought as compared to WW control at day 8 (Fig. 4-1B). The large difference in the magnitude of allantoin accumulation at day 8 between RAC875 and Mace flag leaf was associated with the higher allantoin levels in Mace present already under control conditions ($p<0.05$ by Student's *t*-test).

To identify metabolites that correlated with allantoin and further corroborate that allantoin accumulation occurred as a consequence of drought stress, the metabolic signatures from the developing grains of RAC875 and Mace were assessed using metabolomics techniques (see M&M for details) and visualised by hierarchical clustering and PCA (Supplementary Fig. S4-2). This analysis revealed that allantoin accumulation under drought (HN-D) conditions at day 8 correlated with twelve additional metabolites forming a distinct cluster (Supplementary Fig. S4-2A). This cluster I contained metabolites previously reported to be associated with drought responses, such as a signalling molecule 4-amino-butanoic acid (GABA) and several amino acids including proline, aspartate and asparagine (Bouche and Fromm, 2004; Planchet and Limami, 2015). Interestingly, a second cluster (cluster II) containing 14 metabolites displayed higher levels in RAC875 than in Mace, regardless of the treatment. This group included N-containing metabolites such as the polyamine putrescine, 5-methyltioadenosine (MTA) and N,N-dimethylglycine (DMG), as well as organic acids, such as malic and malonic acids. A third cluster (cluster III) was identified containing seven metabolites that were generally lower under drought stress compared to control conditions. This cluster included the amino acids tryptophan and glutamate, as well as the sugars sucrose, fructose-6-phosphate and glucose-6-phophate.

Purine catabolic genes transcription under low N and drought stress

Transcription of the purine catabolic genes were measured in all of the RAC875 samples collected during the course of the experiment (Fig. 4-3) and on selected Mace samples (Supplementary Fig. S4-3). The data for each gene and treatment is presented in Fig. 4-3 as log2 ratio of the calibrated normalised relative quantities (CNRQ) (Hellemans *et al.*, 2007) between treatment (LN or D) and control conditions (HN-WW).

Analysis of the transcriptional regulation of purine catabolism under low nitrogen revealed that purine catabolic genes were up-regulated in RAC875 at most developmental stages and tissues analysed (Fig. 4-3). In particular, the genes *TaALN* and *TaAAH*, putatively coding for the key enzymes involved in

allantoin and allantoate degradation, respectively, were highly responsive to the low N treatment. Interestingly, at 19 days after anthesis (DAA), all purine catabolic genes were significantly up-regulated in flag leaf and, with a lower magnitude, also in stem and spikelet. This may be a developmental effect given the plants were transitioning from late milk to early dough stages (Fig. 4-3).

In contrast to the co-ordinated transcriptional response evident in plants exposed to N deficiency, RAC875 droughted plants displayed a differential regulation of specific sets of purine catabolic genes (Fig. 4-3). Generally, *TaALN* was down-regulated in all analysed tissues of droughted plants relative to plants grown under control conditions, except in the spikelets where no significant changes in transcription of *TaALN* were detected. Interestingly, a decrease in *TaALN* transcript abundance in YFEL and flag leaf already occurred under mild drought conditions, suggesting that this gene is particularly drought responsive in these tissues. It is worth noting that in YFEL *TaALN* down-regulation at day 2 corresponded to the accumulation of allantoin (Fig. 4-2). In contrast, in flag leaves, despite that *TaALN* transcription was down-regulated at day 3, allantoin accumulation was observed only from day 5 of the drought (HN-D) treatment (Fig. 4-2). The *TaXDH1* and *TaUOX* genes, putatively coding for the enzymes that synthetise allantoin, showed up-regulation under drought across the analysed tissues, with the exception of transcriptional profiles in the grain. Interestingly, the *TaXDH2* gene showed the opposite transcriptional regulation to its paralog, *TaXDH1*. In fact, *TaXDH2* was down-regulated under drought in the YFEL, the flag leaf, the stem and spikelets, with a similar pattern to *TaALN*. *TaAAH* and *TaUGAH* presented different drought responses depending on the tissue analysed (e.g. down-regulation in stem and up-regulation in spikelet).

The data obtained in Mace YFEL are largely in agreement to those obtained from RAC875, showing that N deficiency increases mRNA levels of *TaXDH1*, *TaALN* and *TaAAH* in YFEL; whilst under drought *TaALN* was down-regulated whilst *TaXDH1* was up-regulated (Supplementary Fig. S4-3).

Wheat seedlings grown on xanthine and allantoin as a sole N source

The results presented above showed that the reduction of allantoin in wheat tissues grown under N deficiency was paralleled by an up-regulation of the ureide-degrading genes *TaALN* and *TaAAH* (Fig. 4-2; Fig. 4-3), suggesting enhanced rates of allantoin degradation through the purine catabolic pathway. To assess whether the N recycled from this pathway can be effectively assimilated to support growth in wheat, we performed a hydroponics study in which xanthine and allantoin were provided as the sole source of N after a short period of nitrogen starvation (Table 4-1). The N efficient bread wheat genotype Mace was employed for this study. Visual observation of plants after N had been re-supplemented either as NO_3^- ($rsNO_3^-$), xanthine (rsXant) or allantoin (rsAlnt) for eight days, did not reveal obvious differences to control plants grown under constant NO_3^- supply. In contrast, –N plants appeared smaller, with yellowing of the older leaves (Fig. 4-4A). Physiological analysis of YFEL of the main stem at eight days after N resupply showed that the chlorophyll content and the photosynthetic rate of plants resupplied with xanthine and allantoin were comparable with plants grown on NO_3^-, whilst plants deprived of N had significantly reduced both, chlorophyll content and photosynthetic rate (Fig. 4-4B, C). Plants re-supplemented with xanthine and allantoin showed slightly higher levels of chlorophyll than plants growing on NO_3^-. However, a second experiment aimed to confirm these data showed no statistically differences in chlorophyll levels among N supplied plants (Exp. 2, Table 4-1; Supplementary Fig. S4-4). Analysis of the shoot to root dry weight ratio (S:R) showed that, regardless of the type of N source, plants supplemented with N increased their S:R with similar trends over the eight days after N was re-supplemented; whilst plants kept under constant N deprivation during the same period did not display any significant trends (Fig. 4-4D). Interestingly, linear regression analysis highlighted slightly lower S:R of plants resupplied with allantoin than those resupplied with xanthine (Fig. 4-4D). Shoot N content and shoot dry weight (DW) increased over time in plants supplemented with N and displayed similar slopes, whilst –N plants had smaller shoots and showed no increase in shoot N over time (Fig. 4-4E). The same

experiment was repeated using wheat seedlings harvested eight days after N was resupplied with similar results (Exp. 2, Table 4-1; Supplementary Fig. S4-4).

Allantoin accumulation in wheat grain

To assess the contribution of allantoin to the overall N pools in the grain under control and stress conditions (drought and low N), the remaining tillers of plants harvested during grain filling stages (Table 4-1) were grown until maturity and total N, allantoin and free amino acids concentrations were determined in grain harvested from the tallest remaining tiller (Fig. 4-5). Analysis of grain N% highlighted significant differences between treatments (two-way ANOVA, $p<0.0001$), in particular, it showed a 20% reduction in grain N% of plants grown under low N (LN-WW) conditions with no evident genotypic differences between RAC875 and Mace (Fig. 4-5A). Allantoin concentration was significantly changed because of the applied treatments and genotypes (two-way ANOVA, Genotype × Treatment, $p = 0.0028$) (Fig. 4-5B). Particularly, allantoin was reduced under low N conditions by 44% and 49% in RAC875 and Mace plants, respectively. In contrast, under drought (HN-D), allantoin in RAC875 grains increased by 39% relatively to control (HN-WW). Interestingly, the largest difference in allantoin concentration between genotypes was observed under control conditions (HN-WW), in fact, Mace grains accumulated 66% more allantoin than RAC875, whilst there were no significant differences between the genotypes under low N and drought conditions, as mentioned above (Fig. 4-5B).

Analysis of grain free amino acids content in plants grown under high N and well-watered conditions revealed that allantoin accumulated to a similar concentration as certain amino acids, such as alanine and histidine (Fig. 4-5C). When considering that one allantoin molecule contains four N atoms (1:1 nitrogen to carbon ratio, N:C), the overall N stored in allantoin in grain was larger than the N present in glutamate and aspartate, which were the most concentrated free amino acids identified but have only one N atom (1:4 N:C) (Fig. 4-5C). However, the N stored in allantoin was lower compared to N-rich amino

acids especially with respect to arginine (2:3 N:C) that retained approximately double the N content of allantoin in both RAC875 and Mace grains. The analysis also showed significant genotypic differences in certain amino acids concentrations between RAC875 and Mace. In particular, Mace had 70% and 31% higher concentration levels of arginine and alanine than RAC875, whilst RAC875 had 179%, 43% and 35% higher concentration of tryptophan, aspartate and serine than Mace, respectively (Fig. 4-5C).

To further characterise trends of allantoin and other N-metabolites in wheat grain, metabolomics analyses were performed using RAC875 and Mace developing grain collected between 17 and 22 days after anthesis (DAA) from plants grown under control conditions. Different metabolomics techniques were employed (see M&M for details), and the significance of changes in selected metabolites of interest was assessed via two-way ANOVA analysis (Fig. 4-6). Analysis of metabolites related to N and amino acid metabolisms revealed that the majority of significant changes ($p<0.05$) were related to the sampling time points (T; 17, 19 and 22 DAA). This group contained eighteen metabolites including amino acids, such as threonine and leucine, the polyamines putrescine and spermidine and the inorganic N ions nitrate and ammonium. All of these showed reduced levels at 22 DAA compared to 17 DAA. Eleven metabolites were detected as significantly changed because of the genotype (G) and this group included, for example, glutamine, aspartate and N,N-dimethylglycine (DMG). Overall, RAC875 appeared to accumulate these metabolites to a higher level than Mace (G, $p<0.05$). Four metabolites were detected that accumulated differently between RAC875 and Mace across the analysed time points (G×T, $p<0.05$). For example, levels of lysine and valine were reduced in Mace but appeared constant over time in RAC875. Interestingly, allantoin showed an opposite trend, displaying constant levels in RAC875 between 17 and 22 DAA but increased levels at 22 DAA in Mace.

Discussion

Allantoin accumulation upon drought is mediated by coordinated changes of purine catabolic genes transcription in bread wheat.

The ureide allantoin has been reported to accumulate in response to abiotic stresses, particularly under drought and dehydration (see Introduction). Irani and Todd (2016) and Lescano *et al.* (2016) have recently shown that allantoin accumulation upon drought and salt stress in Arabidopsis leaves was coordinated with changes in transcription of purine catabolic pathway genes. Our data showed that allantoin accumulation under drought was paralleled by up-regulation of genes putatively coding for allantoin-synthesising enzymes (*TaXDH1* and *TaUOX*) and/or down-regulation of *TaALN* putatively coding for the allantoin-degrading enzyme. The data also revealed that leaf tissues and spikelet presented earlier response to drought stress than stem and grain (Fig. 4-2B; Fig. 4-3; Supplementary Fig. S4-3). These observations may indicate that fluxes of the pathway's intermediates are transcriptionally controlled under stress, leading to allantoin accumulation. However, additional levels of regulation (e.g. post-translational) are likely to occur as proposed for AAH in common bean and soybean (Alamillo *et al.*, 2010; Gil-Quintana *et al.*, 2013). For example, post-translational regulation mechanisms may explain the discrepancy observed in the timing of allantoin accumulation and down-regulation of *TaALN* in flag leaf (Fig. 4-2B; Fig. 4-3). Another intriguing observation is that transcription of *TaXDH2* showed opposite regulation to its paralog *TaXDH1* under drought (Fig. 4-3). Casartelli *et al.* (Chapter 3) showed that TaXDH2-6DS protein, the only *TaXDH2* homeologs that is expressed, has a predicted inactive xanthine binding site, prompting the question whether the function of*TaXDH2* has deviated from xanthine degradation.

While accumulation of allantoin was proven to stimulate ABA and JA metabolism (Watanabe *et al.*, 2014b; Takagi *et al.*, 2016), its physiological role in regard to N metabolism is still poorly characterised. Under drought, nutrient remobilisation caused by premature leaf senescence (Munné-Bosch and Alegre, 2004) and increased photorespiration (Wingler *et al.*, 1999) are likely sources ammonium, which is a by-product of these processes. Under drought, the inhibition of photosynthesis reduces the rates of CO_2 assimilation and this has an impact on the supply of C skeletons required by GS/GOGAT to recycle ammonium.

Concentrations of ammonium in plants are tightly regulated as an excess of it can cause toxicity (Britto *et al.*, 2001; Miller *et al.*, 2001; Britto and Kronzucker, 2002). Moreover, ammonia emission from plant tissues is closely associated with photorespiration and GS activity, especially with the chloroplastic isoform GS2 (Mattsson *et al.*, 1997; Kumagai *et al.*, 2011). Interestingly, Watanabe *et al.* (2014b) provided evidence that high levels of allantoin in Arabidopsis leads to reduced stomatal aperture, which is an ABA-regulated process under drought (Leckie *et al.*, 1998; Bright *et al.*, 2006). Stomatal aperture is important for CO_2 assimilation and synthesis of C-skeletons, required for recycling ammonium. Furthermore, soybean leaf homogenates were shown to generate ammonia from ^{14}C-allantoin with rates of about 10 μmol g^{-1} FW h^{-1} (Todd and Polacco, 2004). In this scenario, the increased accumulation of allantoin observed (i.e. 30-fold increase in RAC875 flag leaf under drought; Fig. 4-2) may be involved in both N homeostasis, by preventing N losses as ammonia, and with C assimilation, by regulating stomatal aperture via ABA. However, to what extent purine catabolism affects N and C economy under stress still needs to be characterised in details.

Interestingly, allantoin concentration in flag leaf and mature grain between RAC875 and Mace was most different under control conditions, rather than under stress. In fact, Mace accumulated about 1000% and 66% more allantoin than RAC875 in flag leaf and mature grain, respectively (Fig. 4-2B; Fig. 4-5B). Genotypic differences in other metabolites under control conditions were also observed in the grain of the two cultivars examined (see Fig. 4-5C; Fig. 4-6). Casartelli *et al.*, (2018) (Appendix 1) reported that drought tolerant aus-type rice varieties accumulated allantoin and other metabolites with higher magnitude than sensitive indica varieties under drought and that the differences were evident under well-watered conditions suggesting that it was genetically determined. If accumulation of allantoin in pre-germinated seeds or in young seedlings grown under control conditions can be associated with increased adaptability to stress later in development then this would facilitate the screening of a large number of accessions. Our study indicated that Mace performed better in terms of grain yield under drought than RAC875 ($p<0.02$ by Student's *t*-test), providing the first evidence of a possible involvement of allantoin in drought

tolerance in wheat. Nonetheless, RAC875 was previously described as a highly drought tolerant genotype (Izanloo *et al.*, 2008; Bowne *et al.*, 2011), therefore, the analysis of a larger and more diverse panel of wheat accessions is required to further characterise this association.

Allantoin represents an internal source of recycled N under nutrient restrictions in bread wheat.

Recent metabolomics studies reported that allantoin is reduced in vegetative leaves in maize plants grown under N starvation (Amiour *et al.*, 2012) and also in rice roots grown under limiting N conditions (Coneva *et al.*, 2014). Furthermore, it was shown by Casartelli *et al.* (Chapter 3) that field-grown wheat plants supplemented with higher N levels accumulated more allantoin in the grain than plants grown under N restrictions. Yang and Han (2004), who first cloned a functional *ALN* gene from *Arabidopsis thaliana* and the leguminous tree *Robinia pseudoacacia,* described that both *AtALN* and *RpALN* were transcriptionally activated under N limitation. The results reported in the present study are the first that comprehensively showed that reduction of allantoin under low N (Fig. 4-2A) was paralleled by increased transcription of purine catabolic genes (Fig. 4-3). In particular, *TaALN* and *TaAAH* were the most responsive genes to the low N treatment. Given that the physiological role of purine catabolism in senescing tissues is recycling the N from the purine ring (Zrenner *et al.*, 2006; Werner and Witte, 2011), it is worth questioning whether increasing the rates of N remobilised from the purine ring could provide an alternative source of N that the plant can access when grown under N restriction.

Arabidopsis seedlings were shown to be able to use allantoin and xanthine as the sole source of N in agar media, however, they displayed reduced growth as compared to seedlings supplemented with inorganic N (Desimone *et al*., 2002; Brychkova *et al.*, 2008b). The assumption that intermediates of purine catabolism can provide nutrient supply was demonstrated in the present study as wheat seedlings grew with xanthine and allantoin as a sole source of N (Fig. 4-4). However, in contrast to the findings in Arabidopsis, growth with xanthine or allantoin as the sole source of N was comparable to seedlings

supplied with nitrate (Fig. 4-4). These differences between Arabidopsis and wheat may be explained by the fact that our study included a pre-starvation phase with NO_3^- supply, whilst in the Arabidopsis studies allantoin and xanthine were supplied as basal application since germination. However, in our study the comparison of control plants kept at –N conditions with N-resupplied plants suggests that, despite NO_3^- reserves may have been used during the starvation period (–N plants showed a marginal shoot growth between day 3 and 8 of N starvation), supplying allantoin and xanthine effectively recovered normal growth in wheat (Fig. 4-4E). Another possible explanations is that allantoin and xanthine availability in agar media, used to grow Arabidopsis seedlings, is lower than for liquid media employed in the present study. It is worth noting that wheat seedlings resupplied with allantoin showed generally lower S:R than plants resupplied with xanthine (Fig. 4-4D). Wheat roots increase their surface area when foraging for N (Melino *et al.*, 2015) resulting in increased biomass allocation to below ground organs causing the reduction of shoot to root ratio. Therefore, this can suggests that allantoin is taken up or utilised less efficiently than xanthine. Further studies are required to dissect this in more details.

Allantoin is a significant N pool not bound to proteins in bread wheat grains.

Our previous work (Casartelli *et al.,* Chapter 3) had provided evidence that ureides from purine catabolism may be exported from source tissues, such as flag leaf, to developing wheat grain during natural senescence. The data reported in this study are in support of this showing that grain allantoin progressively accumulates during grain filling. In fact, allantoin concentration at three weeks after anthesis (very late milk; Z79) was approx. 100 nmol g^{-1} DW (Fig. 4-2) and increased to up to 1440 nmol g^{-1}DW at maturity in grain of the wheat genotype Mace (Fig. 4-5B). Consistently, earlier reports have shown allantoin concentration of approx. 700 nmol g^{-1} in grain of the Italian bread wheat genotype Mentana (Montalbini, 1992).

Although allantoin accounted for only 0.3% of the total N measured in wheat grains, our data suggests that allantoin represents a significant non-protein bound N pool, comparable to free amino acids (Fig. 4-5C). Howarth *et al.* (2008) showed that total free amino acid concentration in wheat grain decreased from 300 to 50 μmol g^{-1}DW between 7 and 28 days after anthesis, respectively, likely due to incorporation into grain protein. Interestingly, several free amino acids and polyamines generally declined in RAC875 and Mace grains between 17 and 22 DAA. However, allantoin did not display such a trend, especially in Mace (Fig. 4-6). Previous studies have shown that allantoin stored in wheat grain is quickly utilised starting at one day after germination (Montalbini, 1992), suggesting that it can be used as a readily available N substrate. In addition, a study in rice reported a positive link between grain allantoin and rice seedling survivals in seedbeds under abiotic stress and a beneficial effect on plant growth when rice grains were primed with allantoin solution before germination (Wang *et al.*, 2012).The authors also showed that allantoin concentration varied among rice cultivars, being generally higher in japonica compared with indica varieties. The data reported here suggest that genotypic differences exists also among bread wheat genotypes, in fact, Mace accumulated 66% more allantoin in grain than RAC875, independent of total grain N, which was similar in the two genotypes (Fig. 4-5A, B). In support of this, allantoin analysis also showed Mace to accumulate allantoin about 1000% more than RAC875 in flag leaf under control conditions ($p<0.05$ by Student's *t*-test) (Fig. 4-2). Interestingly, Mace mature grain accumulated 70% more arginine than RAC875 (Fig. 4-5C). Arginine is the proteinogenic amino acid with highest N:C ratio (3:2), known to represent a major N form for storage and transport and its catabolism liberates two molecules of ammonium via urease (Winter *et al.*, 2015). Our data suggest that the two analysed wheat genotypes use different N compounds for transport and grain loading. Preliminary studies on Mace suggested that this cultivar has high N use efficiency (Mahjourimajd *et al.*, 2016). The possibility that Mace superior N use efficiency may be related to the preferential use of metabolites with high N:C ratio, which are more energy effective forms for transporting and storing N, is an intriguing hypothesis that would be interesting to dissect in more details. Nonetheless, is worth to mention that grain allantoin and arginine may also play

a regulatory role in plant growth and development as they are known components of ABA- and nitric oxide (NO)-signalling, respectively (Watanabe *et al*., 2014; Winter *et al*., 2015).

Conclusion

While much attention has been paid in characterising the role of allantoin in tropical legumes and in stress responses of the model plant Arabidopsis, the function of allantoin and purine catabolic genes remained mostly uncharacterised in monocot cereals. Our study is the first to reveal a possible dual role of allantoin accumulation or degradation in bread wheat. Under drought, the reduced photosynthetic capacity results in reduced C skeletons supply that would be required to assimilate the ammonium derived from allantoin degradation. In this scenario, accumulation of allantoin might prevent the liberation of ammonium, which could otherwise lead to cellular toxicity and N losses by ammonia volatilisation. Therefore, allantoin may be a temporary N sink under drought and a N source during recovery of growth. On the other hand, under N restrictions, plants suffer from a shortage of ammonium assimilated from inorganic N sources that can be alleviated by increasing the amounts of ammonium derived from the recycling of N-rich metabolites, such as allantoin. Therefore, the role of allantoin degradation when external N is limiting may reside in providing the plant with an ideal alternative N source, given its high N:C ratio, to be used for maintaining growth. In addition, we demonstrate, for the first time, that allantoin accumulates during grain filling and the transcriptional data reported by Casartelli *et al* (Chapter 3) suggest that allantoin is translocated to the grain from source tissues. Genotypic differences in allantoin accumulation were observed between RAC875 and Mace in flag leaves and grains. Screening a larger number of wheat accessions may support the predicted relationship between allantoin and drought tolerance or N use efficiency. Specific purine catabolic genes were particularly responsive to either drought and/or low N treatments (*TaXDH1*, *TaALN*, *TaAAH*) and, therefore, represent ideal candidates for targeted genetic engineering approaches.

Supplementary data

Table S4-1. Primers employed for quantitative reverse transcription PCR (qRT-PCR) analysis.

Table S4-2. Optimisation of allantoin detection method.

Fig. S4-1. Soil water retention curve.

Fig. S4-2. Metabolite response of RAC875 and Mace developing grain at day 8 of the drought treatment.

Fig. S4-3. Comparison of relative expression of *TaXDH1*, *TaALN* and *TaAAH* in youngest fully emerged leaf (YFEL) of RAC875 and Mace under low N and drought.

Fig. S4-4. Analysis of wheat N starved seedlings resupplied with different N sources for eight days.

Acknowledgments

We thank Akiko Enju, Pia Müller and Jessey George (University of Adelaide) for supporting with sample collections; Sanjiv Satija (University of Adelaide) for technical support with total N analysis; Larissa Chirkova (University of Adelaide) for technical support with the amino acid analysis; Julian Taylor (University of Adelaide) for support in experimental design and statistical analysis; Yuan Li (University of Adelaide) for technical support with qRT-PCR analyses; Nirupama S. Jayasinghe, Siria Natera, Gina Barossa and Veronica Liu (University of Melbourne) for technical support with allantoin GS-QqQ-MS quantification; Elmien Heyneke, Mutsumi Watanabe, Alexander Erban, Ines Fehrle, Astrid Basner and Ellen Zuther (Max Plank Institute, Potsdam-Golm) for support in performing metabolomics analysis. This work was supported by an Australian Research Council linkage grant LP1400100239 in partnership with DuPont-Pioneer.

Figure Legends

Fig. 4-1. Outline of the purine catabolic pathway and its role in N remobilisation. The degradation reactions are reported beginning with xanthine, the first intermediate of the pathway. Abbreviations of enzymes are indicated in capital letters. The ammonium (NH_4^+) generated can be assimilate by GS/GOGAT cycle. The chemical structure of xanthine and allantoin, key pathway's intermediates, are

reported in a box. Abbreviations include XDH, xanthine dehydrogenase; UOX, urate oxidase; AS, allantoin synthase; ALN, allantoinase; AAH, allantoate amidohydrolase; UGAH, ureidoglycine aminohydrolase; UAH, ureidoglycolate amidohydrolase; GS, glutamine synthetase; GOGAT, glutamate synthase.

Fig. 4-2. Allantoin concentration in bread wheat genotypes RAC875 and Mace under N deficiency and drought stress. Allantoin concentration was quantified by GC-QqQ-MS under (A) low N and (B) drought conditions with respect to control conditions in youngest fully emerged leaf, flag leaf, stem, spikelet and developing grain. YFEL samples were collected during tillering stages at 23, 24 and 27 days after sowing, corresponding to 2, 3 and 6 days of drought stress (days after last watering), respectively. Flag leaf, stem, spikelet and developing grain samples were collected during grain filling stages at 17, 19 and 22 days after anthesis (DAA), corresponding to 3, 5 and 8 days of drought (days after last watering), respectively (see Table 1). Data are mean ±SEM of four to six biological replicates and asterisks indicate significant differences between treatment (high N, drought) and control conditions per genotype per time point (*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001 by Student's *t*-test).

Fig. 4-3. qRT-PCR analysis of the transcriptional profile of purine catabolic genes in various tissues of the bread wheat genotype RAC875 under stress. RNA was extracted from the youngest fully emerged leaf (YFEL), flag leaf, stem, spikelet and developing grain of plants grown under control (HN-WW), low N (LN) and drought (D) conditions collected at different time points (for details see Table 4-1). Levels of gene expression were determined by quantitative real-time PCR (qRT-PCR) employing gene-specific primers and the calibrated normalised relative quantity (CNRQ) was calculated utilising the most stable reference genes across tissues: *TaActin* and *TaGAPdH* for YFEL and spikelet; *TaCyclophilin* and *TaGAPdH* for flag leaf and stem; *TaCyclophilin* and *TaEFA2* for developing grain. Data is expressed as log2 calibrated normalised relative quantity (CNRQ) of gene transcription under N deficiency/control or CNRQ gene transcription under drought/control. Colour denotes significant differences between treatment (LN, D) and control conditions (WW, HN) ($p<0.05$ by Student's *t*-test) with red indicating up-regulation and blue down-regulation with respect of control. Asterisk indicates significant change in allantoin concentration occurred at the specific time point.

Fig. 4-4. Growth of bread wheat cv. Mace supplemented with different N sources after N starvation. Mace seedlings were grown in a hydroponic system supplied with 1 mM nitrate for two weeks and then starved of N for 24 hours. N was resupplied as 1 mM nitrate ($rsNO_3^-$), 0.25 mM xanthine (rsXanth), 0.25 mM allantoin (rsAlnt) or no resupply (−N) for eight days; a subsets of plants were kept at constant nitrate supply throughout the experiment (NO_3^-). (A) Photographs of the shoots; (B) chlorophyll content of the youngest fully emerged leaf (YFEL) of the main stem and (C) photosynthetic rate of the YFEL eight days

after N was resupplied. Data are mean of six (B) and five (C) biological replicates ± SEM and letters indicate significant differences between treatments in a single time point by one-way ANOVA with Tukey's correction ($p<0.05$). (D) Relationship between plant dry weight (DW) and shoot to root DW ratio (S:R) and (E) between shoot DW and shoot N content of seedlings at 3, 5 and 8 days after N was resupplied. Linear regression lines were fitted to the data and in (D) slopes were significantly non-zero with $p<0.005$ for all treatments, except for −N (non-significant). In (E) slopes were significantly non-zero with $p<0.0001$ for all treatments, except −N ($p = 0.0101$).

Fig. 4-5. Total N and N-containing metabolites in mature grain of bread wheat genotypes RAC875 and Mace. Comparison of (A) N concentration (%) and (B) allantoin concentration (nmol g^{-1}DW) in mature grain of RAC875 and Mace plants grown under control conditions (HN-WW), drought (HN-D) and low N (LN-WW) conditions. Letters indicate significant differences between genotypes and treatments by two-way ANOVA with Tukey's test ($p<0.05$); (C) free amino acids levels in mature grain of RAC875 and Mace grown under control conditions (HN-WW). Amino acid concentration is reported in black, whilst N atoms present in each amino acid in grey, values are expressed as nmol g^{-1}DW. Allantoin and corresponding N concentration are reported in lighter colours to allow comparison. Asterisks denote significant differences between RAC875 (R) and Mace (M) by Student's *t*-test (*$p<0.05$; **$p<0.01$; ****$p<0.0001$). All data are mean ± SEM of 16-18 biological replicates.

Fig. 4-6. Changes of primary metabolism in developing grains of bread wheat genotypes RAC875 and Mace grown under control (HN-WW) conditions. Metabolic analysis using GC-TOF-MS, GC-QqQ-MS (allantoin), HPLC (spermidine) and ion chromatography techniques were performed on developing grain collected between 17 and 22 days after anthesis (DAA) from heads of the main stem (see M&M). Data from selected metabolites was analysed and reported as mean ± 5-6 biological replicates of scaled values (%) to the maximum value of corresponding metabolite across the experiment. Significant differences were assessed by analysis of variance with $p<0.05$. G×T, significant interaction of genotype and time point; G/T, genotype and time point are both significantly different but not their interaction; G significant differences between genotypes; T, significant differences between time points. DMG, N,N-dimethyl glycine; MAT, 5-Methylthio-adenosine.

References

Alamillo JM, Diaz-Leal JL, Sanchez-Moran MV, Pineda M. 2010. Molecular analysis of ureide accumulation under drought stress in *Phaseolus vulgaris* L. *Plant, Cell & Environment* **33**, 1828-1837.

Amiour N, Imbaud S, Clément G*, et al.* 2012. The use of metabolomics integrated with transcriptomic and proteomic studies for identifying key steps involved in the control of nitrogen metabolism in crops such as maize. *Journal of Experimental Botany* **63**, 5017-5033.

Bennett D, Izanloo A, Reynolds M, Kuchel H, Langridge P, Schnurbusch T. 2012. Genetic dissection of grain yield and physical grain quality in bread wheat (*Triticum aestivum* L.) under water-limited environments. *Theoretical and Applied Genetics* **125**, 255-271.
Bouche N, Fromm H. 2004. GABA in plants: just a metabolite? *Trends in Plant Science* **9**, 110-115.
Bowne JB, Erwin TA, Juttner J, Schnurbusch T, Langridge P, Bacic A, Roessner U. 2011. Drought responses of leaf tissues from wheat cultivars of differing drought tolerance at the metabolite level. *Molecular Plant* **5**, 418-429.
Bright J, Desikan R, Hancock JT, Weir IS, Neill SJ. 2006. ABA-induced NO generation and stomatal closure in Arabidopsis are dependent on H_2O_2 synthesis. *The Plant Journal* **45**, 113-122.
Britto DT, Kronzucker HJ. 2002. NH_4^+ toxicity in higher plants: a critical review. *Journal of Plant Physiology* **159**, 567-584.
Britto DT, Siddiqi MY, Glass AD, Kronzucker HJ. 2001. Futile transmembrane NH_4^+ cycling: a cellular hypothesis to explain ammonium toxicity in plants. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4255-4258.
Brychkova G, Alikulov Z, Fluhr R, Sagi M. 2008a. A critical role for ureides in dark and senescence-induced purine remobilization is unmasked in the *Atxdh1* Arabidopsis mutant. *The Plant Journal* **54**, 496-509.
Brychkova G, Fluhr R, Sagi M. 2008b. Formation of xanthine and the use of purine metabolites as a nitrogen source in Arabidopsis plants. *Plant Signaling & Behavior* **3**, 999-1001.
Casartelli A, Riewe D, Hubberten HM, Altmann T, Hoefgen R, Heuer S. 2018. Exploring traditional aus-type rice for metabolites conferring drought tolerance. *Rice* **11**, 9.
Coleto I, Pineda M, Rodino AP, De Ron AM, Alamillo JM. 2014. Comparison of inhibition of N_2 fixation and ureide accumulation under water deficit in four common bean genotypes of contrasting drought tolerance. *Annals of Botany* **113**, 1071-1082.
Coneva V, Simopoulos C, Casaretto JA*, et al.* 2014. Metabolic and co-expression network-based analyses associated with nitrate response in rice. *BMC Genomics* **15**, 1056.
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.
deSilva M, Purcell LC, King CA. 1996. Soybean petiole ureide response to water deficits and decreased transpiration. *Crop Science* **36**, 611-616.
Desimone M, Catoni E, Ludewig U, Hilpert M, Schneider A, Kunze R, Tegeder M, Frommer WB, Schumacher K. 2002. A Novel Superfamily of Transporters for Allantoin and Other Oxo Derivatives of Nitrogen Heterocyclic Compounds in Arabidopsis. *The Plant Cell* **14**, 847-856.
Do PT, Degenkolbe T, Erban A, Heyer AG, Kopka J, Köhl KI, Hincha DK, Zuther E. 2013. Dissecting Rice Polyamine Metabolism under Controlled Long-Term Drought Stress. *PLoS ONE* **8**, e60325.
Erban A, Schauer N, Fernie AR, Kopka J. 2007. Nonsupervised construction and application of mass spectral and retention time index libraries from time-of-flight gas chromatography-mass spectrometry metabolite profiles. *Metabolomics: methods and protocols*, 19-38.
Gil-Quintana E, Larrainzar E, Seminario A, Diaz-Leal JL, Alamillo JM, Pineda M, Arrese-Igor C, Wienkoop S, Gonzalez EM. 2013. Local inhibition of nitrogen fixation and nodule metabolism in drought-stressed soybean. *Journal of Experimental Botany* **64**, 2171-2182.
Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology* **8**, R19.
Herridge DF, Atkins CA, Pate JS, Rainbird RM. 1978. Allantoin and Allantoic Acid in the Nitrogen Economy of the Cowpea (*Vigna unguiculata* [L.] Walp.). *Plant Physiology* **62**, 495-498.
Howarth JR, Parmar S, Jones J*, et al.* 2008. Co-ordinated expression of amino acid metabolism in response to N and S deficiency during wheat grain filling. *Journal of Experimental Botany* **59**, 3675-3689.
Irani S, Todd CD. 2016. Ureide metabolism under abiotic stress in *Arabidopsis thaliana*. *Journal of Plant Physiology* **199**, 87-95.
Izanloo A, Condon AG, Langridge P, Tester M, Schnurbusch T. 2008. Different mechanisms of adaptation to cyclic water stress in two South Australian bread wheat cultivars. *Journal of Experimental Botany* **59**, 3327-3346.
Johnson CM, Stout PR, Broyer TC, Carlton AB. 1957. Comparative chlorine requirements of different plant species. *Plant and Soil* **8**, 337-353.
Kanani H, Dutta B, Klapa MI. 2010. Individual vs. combinatorial effect of elevated CO_2 conditions and salinity stress on *Arabidopsis thaliana* liquid cultures: Comparing the early molecular response using time-series transcriptomic and metabolomic analyses. *BMC Systems Biology* **4**, 177-177.
Kaplan F, Kopka J, Haskell DW, Zhao W, Schiller KC, Gatzke N, Sung DY, Guy CL. 2004. Exploring the Temperature-Stress Metabolome of Arabidopsis. *Plant Physiology* **136**, 4159-4168.
King CA, Purcell LC. 2005. Inhibition of N_2 fixation in soybean is associated with elevated ureides and amino acids. *Plant Physiology* **137**, 1389.
Kumagai E, Araki T, Hamaoka N, Ueno O. 2011. Ammonia emission from rice leaves in relation to photorespiration and genotypic differences in glutamine synthetase activity. *Annals of Botany* **108**, 1381-1386.
Leckie CP, McAinsh MR, Allen GJ, Sanders D, Hetherington AM. 1998. Abscisic acid-induced stomatal closure mediated by cyclic ADP-ribose. *Proceedings of the National Academy of Sciences* **95**, 15837-15842.

Lescano C, Martini C, González C, Desimone M. 2016. Allantoin accumulation mediated by allantoinase downregulation and transport by Ureide Permease 5 confers salt stress tolerance to *Arabidopsis* plants. *Plant Molecular Biology* **91**, 581-595.
Li X, Lawas LMF, Malo R*, et al.* 2015. Metabolic and transcriptomic signatures of rice floral organs reveal sugar starvation as a factor in reproductive failure under heat and drought stress. *Plant, Cell & Environment* **38**, 2171-2192.
Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. 2006. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nature protocols* **1**, 387.
Mahjourimajd S, Kuchel H, Langridge P, Okamoto M. 2016. Evaluation of Australian wheat genotypes for response to variable nitrogen application. *Plant and Soil* **399**, 247-255.
Mattsson M, Häusler RE, Leegood RC, Lea PJ, Schjoerring JK. 1997. Leaf-Atmosphere NH_3 Exchange in Barley Mutants with Reduced Activities of Glutamine Synthetase. *Plant Physiology* **114**, 1307-1312.
Melino VJ, Fiene G, Enju A, Cai J, Buchner P, Heuer S. 2015. Genetic diversity for root plasticity and nitrogen uptake in wheat seedlings. *Functional Plant Biology* **42**, 942.
Metsalu T, Vilo J. 2015. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Research* **43**, W566-W570.
Miller AJ, Cookson SJ, Smith SJ, Wells DM. 2001. The use of microelectrodes to investigate compartmentation and the transport of metabolized inorganic ions in plants. *Journal of Experimental Botany* **52**, 541.
Montalbini P. 1992. Ureides and enzymes of ureide synthesis in wheat seeds and leaves and effect of allopurinol on *Puccinia recondita* f. sp. tritici infection. *Plant Science* **87**, 225-231.
Moschen S, Bengoa Luoni S, Di Rienzo JA, Caro MdP, Tohge T, Watanabe M, Hollmann J, González S, Rivarola M, García-García F. 2016. Integrating transcriptomic and metabolomic analysis to understand natural leaf senescence in sunflower. *Plant Biotechnology Journal* **14**, 719-734.
Munné-Bosch S, Alegre L. 2004. Die and let live: leaf senescence contributes to plant survival under drought stress. *Functional Plant Biology* **31**, 203-216.
Nakagawa A, Sakamoto S, Takahashi M, Morikawa H, Sakamoto A. 2007. The RNAi-mediated silencing of xanthine dehydrogenase impairs growth and fertility and accelerates leaf senescence in transgenic Arabidopsis plants. *Plant & Cell Physiology* **48**, 1484-1495.
Nam M, Bang E, Kwon T, Kim Y, Kim E, Cho K, Park W, Kim B, Yoon I. 2015. Metabolite Profiling of Diverse Rice Germplasm and Identification of Conserved Metabolic Markers of Rice Roots in Response to Long-Term Mild Salinity Stress. *International Journal of Molecular Sciences* **16**, 21959-21974.
Nikiforova VJ, Kopka J, Tolstikov V, Fiehn O, Hopkins L, Hawkesford MJ, Hesse H, Hoefgen R. 2005. Systems rebalancing of metabolism in response to sulfur deprivation, as revealed by metabolome analysis of Arabidopsis plants. *Plant Physiology* **138**, 304.
Nourimand M, Todd CD. 2016. Allantoin increases cadmium tolerance in Arabidopsis via activation of antioxidant mechanisms. *Plant & Cell Physiology* **57**, 2485-2496.
Oksman-Caldentey K-M, Inzé D. 2004. Plant cell factories in the post-genomic era: new ways to produce designer secondary metabolites. *Trends in Plant Science* **9**, 433-440.
Oliver MJ, Guo L, Alexander DC, Ryals JA, Wone BW, Cushman JC. 2011. A sister group contrast using untargeted global metabolomic analysis delineates the biochemical regulation underlying desiccation tolerance in *Sporobolus stapfianus*. *The Plant Cell* **23**, 1231-1248.
Pate JS, Atkins CA, White ST, Rainbird RM, Woo KC. 1980. Nitrogen Nutrition and Xylem Transport of Nitrogen in Ureide-producing Grain Legumes. *Plant Physiology* **65**, 961-965.
Pèlissier HC, Tegeder M. 2007. PvUPS1 plays a role in source–sink transport of allantoin in French bean (*Phaseolus vulgaris*). *Functional Plant Biology* **34**, 282.
Planchet E, Limami AM. 2015. Amino Acid Synthesis under Abiotic Stress. In: D'Mello JPF, ed. *Amino Acids in Higher Plants*. Oxfordshire and Boston: CABI Publishers, 262-276.
Porter JR, Xie L, Challinor AJ, Cochrane K, Howden SM, Iqbal MM, Lobell DB, Travasso MI. 2014. Food security and food production systems. In: Field CB, Barros VR, Dokken DJ, Mach KJ, Mastrandrea MD, Bilir TE, Chatterjee M, Ebi KL, Estrada YO, Genova RC, Girma B, Kissel ES, Levy AN, MacCracken S, Mastrandrea PR, White LL, eds. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel of Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 485-533.
Reinbothe H, Mothes K. 1962. Urea, Ureides, and Guanidines in Plants. *Annual Review of Plant Physiology* **13**, 129-149.
Rose MT, Rose TJ, Pariasca-Tanaka J, Yoshihashi T, Neuweger H, Goesmann A, Frei M, Wissuwa M. 2012. Root metabolic response of rice (*Oryza sativa* L.) genotypes with contrasting tolerance to zinc deficiency and bicarbonate excess. *Planta* **236**, 959-973.
Schubert KR. 1986. Products of biological nitrogen fixation in higher plants: synthesis, transport, and metabolism. *Annual Review of Plant Physiology*, 539-574.
Silvente S, Sobolev AP, Lara M. 2012. Metabolite Adjustments in Drought Tolerant and Sensitive Soybean Genotypes in Response to Water Stress (Metabolite Adjustments of Soybean to Drought). *PLoS ONE* **7**, e38554.
Sinclair TR, Serraj R. 1995. Legume nitrogen fixation and drought. *Nature* **378**, 344-344.

Szabados L, Savoure A. 2010. Proline: a multifunctional amino acid. *Trends in Plant Science* **15**, 89-97.
Takagi H, Ishiga Y, Watanabe S*, et al.* 2016. Allantoin, a stress-related purine metabolite, can activate jasmonate signaling in a MYC2-regulated and abscisic acid-dependent manner. *Journal of Experimental Botany* **67**, 2519.
Todd CD, Polacco JC. 2004. Soybean cultivars 'Williams 82' and 'Maple Arrow' produce both urea and ammonia during ureide degradation. *Journal of Experimental Botany* **55**, 867.
Vanrensburg L, Kruger GHJ, Kruger H. 1993. Proline accumulation as drought-tolerance selection criterion - its relationship to membrane integrity and chloroplast ultrasrtucture in *Nicotiana tabacum* L. *Journal of Plant Physiology* **141**, 188-194.
Verpoorte R. 1998. Exploration of nature's chemodiversity: the role of secondary metabolites as leads in drug development. *Drug Discovery Today* **3**, 232-238.
Wang P, Kong CH, Sun B, Xu XH. 2012. Distribution and function of allantoin (5-ureidohydantoin) in rice grains. *Journal of Agricultural and Food Chemistry* **60**, 2793-2798.
Wang W-S, Zhao X-Q, Li M, Huang L-Y, Xu J-L, Zhang F, Cui Y-R, Fu B-Y, Li Z-K. 2016. Complex molecular mechanisms underlying seedling salt tolerance in rice revealed by comparative transcriptome and metabolomic profiling. *Journal of Experimental Botany* **67**, 405-419.
Wasternack C, Hause B. 2013. Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany. *Annals of Botany* **111**, 1021-1058.
Watanabe S, Kounosu Y, Shimada H, Sakamoto A. 2014a. Arabidopsis xanthine dehydrogenase mutants defective in purine degradation show a compromised protective response to drought and oxidative stress. *Plant Biotechnology* **31**, 173-178.
Watanabe S, Matsumoto M, Hakomori Y, Takagi H, Shimada H, Sakamoto A. 2014b. The purine metabolite allantoin enhances abiotic stress tolerance through synergistic activation of abscisic acid metabolism. *Plant, Cell & Environment* **37**, 1022-1036.
Watanabe S, Nakagawa A, Izumi S, Shimada H, Sakamoto A. 2010. RNA interference-mediated suppression of xanthine dehydrogenase reveals the role of purine metabolism in drought tolerance in Arabidopsis. *FEBS Letters* **584**, 1181-1186.
Werner AK, Witte CP. 2011. The biochemistry of nitrogen mobilization: purine ring catabolism. *Trends in Plant Science* **16**, 381-387.
Wingler A, Quick WP, Bungard RA, Bailey KJ, Lea PJ, Leegood RC. 1999. The role of photorespiration during drought stress: an analysis utilizing barley mutants with reduced activities of photorespiratory enzymes. *Plant, Cell & Environment* **22**, 361-373.
Winter G, Todd CD, Trovato M, Forlani G, Funck D. 2015. Physiological implications of arginine metabolism in plants. *Frontiers in Plant Science* **6**.
Wu H, Liu X, You L, Zhang L, Zhou D, Feng J, Zhao J, Yu J. 2012. Effects of Salinity on Metabolic Profiles, Gene Expressions, and Antioxidant Enzymes in Halophyte *Journal of Plant Growth Regulation* **31**, 332-341.
Yang J, Han KH. 2004. Functional characterization of allantoinase genes from Arabidopsis and a nonureide-type legume black locust. *Plant Physiology* **134**, 1039-1049.
Yobi A, Wone BW, Xu W, Alexander DC, Guo L, Ryals JA, Oliver MJ, Cushman JC. 2013. Metabolomic profiling in *Selaginella lepidophylla* at various hydration states provides new insights into the mechanistic basis of desiccation tolerance. *Molecular Plant* **6**, 369-385.
Yoshida T, Mogami J, Yamaguchi-Shinozaki K. 2014. ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. *Current Opinion in Plant Biology* **21**, 133-139.
Zrenner R, Stitt M, Sonnewald U, Boldt R. 2006. Pyrimidine and purine biosynthesis and degradation in plants. *Annual Review of Plant Biology* **57**, 805.

Table 4-1. Details of plant sample collection during the course of the experiments.

Pot experiment

Age at collection	Developmental stage	Zadoks growth scale		Well-watered Soil water content(%)	Drought Soil water content (%)	Days after last watering	Tissue harvested
23 DAS	Tillering	Z22	main shoot and two tillers	18.19	11.4	Day 2	YFEL (main stem)
24 DAS	Tillering	Z22	main shoot and two tillers	19.28	10.0	Day 3	YFEL (main stem)
27 DAS	Tillering	Z22	main shoot and two tillers	18.11	6.3	Day 6	YFEL (main stem)
17 DAA	Grain filling	Z79	very late milk, half solid/half liquid	17.60	12.3	Day 3	Flag leaf, stem, spikelet, developing grain (main stem)
19 DAA	Grain filling	Z79	very late milk, half solid/half liquid	18.2	8.6	Day 5	Flag leaf, stem, spikelet, developing grain (main stem)
22 DAA	Grain filling	Z79	very late milk, half solid/half liquid	18.75	6.6	Day 8	Flag leaf, stem, spikelet, developing grain (main stem)
Maturity	Ripening	Z92	harvest ripe				Above ground biomass; mature (grain first tiller)

Hydroponic experiment

Experiment	Age at collection	Developmental stage	Zadoks growth scale		Days after N resupply	Tissue harvested	
Exp.1	18 DAG	Tillering	Z21	main stem and one tiller	3	Shoot, root	
Exp. 1	20 DAG	Tillering	Z22	main shoot and two tillers	5	Shoot, root	
Exp. 1	23 DAG	Tillering	Z22	main shoot and two tillers	8	Shoot, root	
Exp. 2	23 DAG	Tillering	Z23	main stem and three tillers	8	Shoot, root	(data presented in Supplementary Fig. S5)

DAS, days after sowing; DAA, days after anthesis; DAG, days after germination; YFEL, youngest fully emerged leaf

Table 4-2. Mean value of agronomic traits of plants used for molecular analysis at maturity.

Cultivar, treatment	Above ground biomass g (DW) per plant	Grain yield (GY) g per plant	Harvest Index (HI) GY : Biomass
RAC875 HN-WW	10.6 ± 0.377 (a)	4.2 ± 0.171 (a)	0.399 ± 0.006
RAC875 LN-WW	8.4 ± 0.289 (b)	3.2 ± 0.101 (b)	0.382 ± 0.003
RAC875 HN-D	9.0 ± 0.408 (b)	3.1 ± 0.147 (b)	0.347 ± 0.009
Mace HN-WW	10.3 ± 0.399 (a)	4.1 ± 0.154 (a)	0.396 ± 0.006
Mace LN-WW	7.9 ± 0.234 (b)	3.1 ± 0.101 (b)	0.387 ± 0.009
Mace HN-D	8.8 ± 0.371 (b)	3.4 ± 0.145 (b)	0.390 ± 0.004
p-value			
Genotype	n.s.	n.s.	n.d.
Treatment	p<0.0001	p<0.0001	n.d.
Genotype × Treatment	n.s.	p<0.05	n.d.

HN-WW, high nitrogen well-watered (control condition); LN-WW, low nitrogen well-watered; HN-D, high nitrogen drought. Each value represent the mean ± SEM of 16-18 biological replicates. Two-way ANOVA analysis was performed with Tukey's correction. Statistical significance is indicated by p-values, n.s. indicates not significant, n.d. indicates not determined and letters indicate significant differences between genotypes and treatments at $p<0.05$.

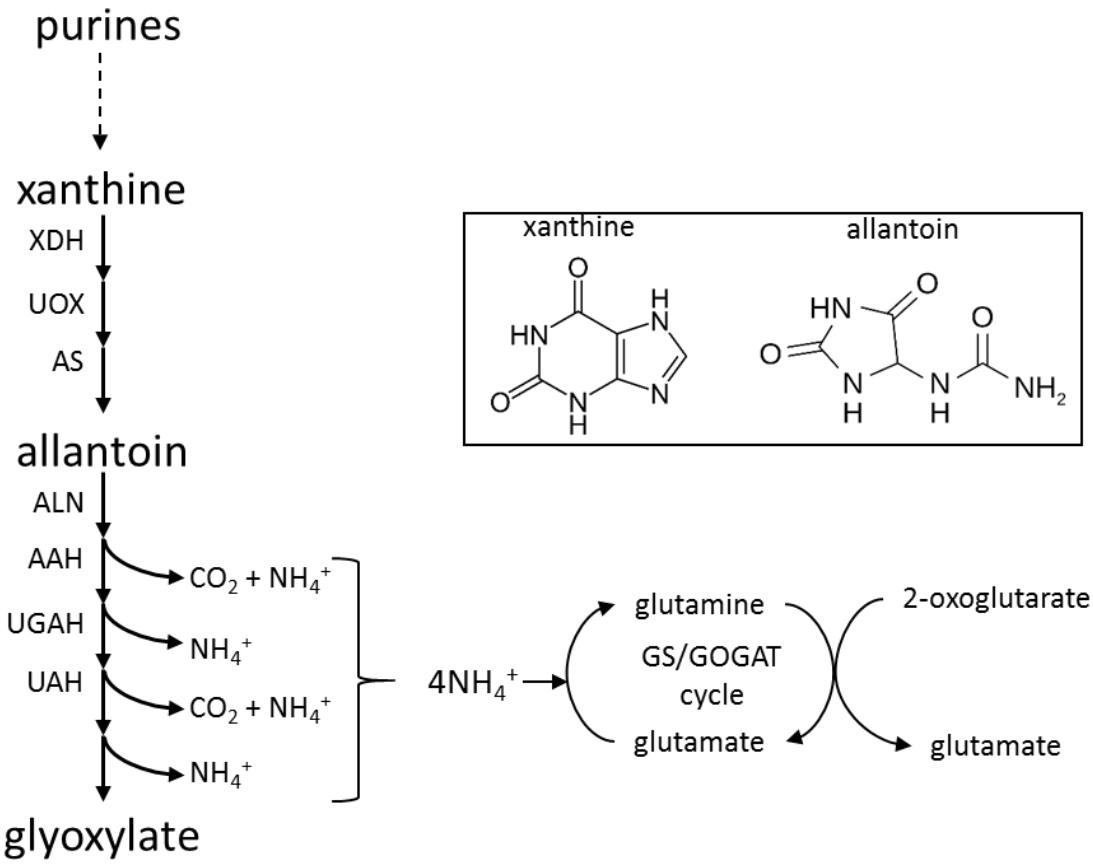


Fig. 4-1. Outline of the purine catabolic pathway and its role in N remobilisation.

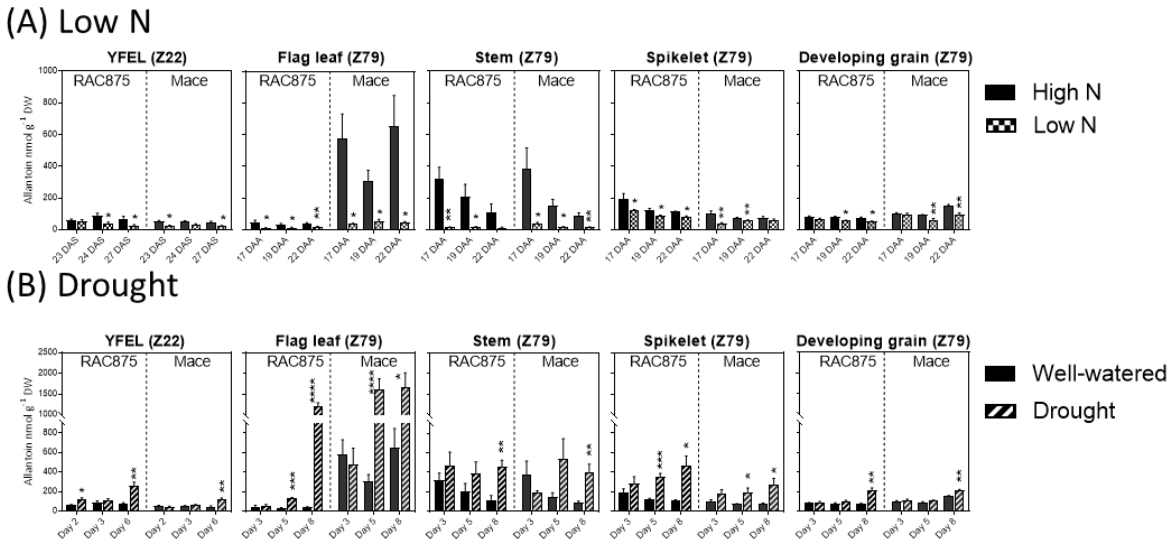


Fig. 4-2. Allantoin concentration in bread wheat genotypes RAC875 and Mace under N deficiency and drought stress.

Low N	YFEL (Z22)			Flag leaf (Z79)			Stem (Z79)			Spikelet (Z79)			Developing grain (Z79)		
Log2 CNRQ (LN/HN)	23 DAS	24 DAS*	27 DAS*	17 DAA*	19 DAA*	22 DAA*	17 DAA*	19 DAA*	22 DAA	17 DAA*	19 DAA*	22 DAA*	17 DAA	19 DAA*	22 DAA*
TaXDH1	0.12	0.33	0.38	0.94	2.03	-0.05	0.42	0.40	0.54	0.01	0.29	0.09	0.52	-0.05	0.79
TaXDH2	-0.07	-0.08	0.40	-0.49	1.39	0.86	0.19	0.57	0.45	0.06	0.40	0.21	1.94	0.17	0.81
TaUOX	0.22	0.27	0.26	0.52	1.94	0.44	0.08	0.44	0.28	0.09	0.18	0.02	0.63	-0.17	0.12
TaAS	0.11	0.28	0.05	0.82	1.65	0.45	-0.08	0.30	0.24	0.01	0.17	-0.07	0.30	0.39	0.33
TaALN	1.78	1.00	1.12	0.59	1.42	1.14	1.91	1.16	0.45	1.33	0.82	0.60	0.60	0.26	0.25
TaAAH	0.60	1.90	0.85	1.50	2.37	2.74	1.24	1.08	1.39	0.79	0.54	0.71	0.50	-0.01	0.15
TaUGAH	-0.02	-0.09	-0.15	0.21	0.61	0.02	0.09	0.48	0.32	-0.09	0.22	0.06	0.84	0.26	-0.25
TaUAH	-0.03	0.16	0.09	0.27	1.72	0.45	0.36	0.26	0.31	-0.08	0.30	-0.14	0.44	-0.11	0.12

Drought	YFEL (Z22)			Flag leaf (Z79)			Stem (Z79)			Spikelet (Z79)			Developing grain (Z79)		
Log2 CNRQ (D/WW)	Day 2*	Day 3	Day 6*	Day 3	Day 5*	Day 8*	Day 3	Day 5	Day 8*	Day 3	Day 5*	Day 8*	Day 3	Day 5	Day 8*
TaXDH1	0.16	0.46	0.63	-0.32	0.36	2.62	0.01	-0.06	0.89	-0.11	0.92	2.06	0.06	-0.14	0.14
TaXDH2	-0.44	-0.90	-0.22	-1.48	-1.04	-0.74	-0.23	-0.29	-1.37	-0.52	-0.14	-0.14	0.96	0.06	0.09
TaUOX	0.16	0.19	0.66	-0.22	0.78	1.06	-0.12	-0.06	0.14	-0.14	0.81	1.49	0.31	-0.27	-0.54
TaAS	0.20	0.41	0.11	-0.07	0.53	-1.04	-0.14	-0.13	-0.45	-0.01	0.51	0.50	-0.04	0.16	-0.85
TaALN	-0.78	-1.57	-1.53	-1.18	-1.88	-0.07	0.60	-1.10	-2.07	0.49	-0.26	-0.09	0.03	-0.79	-2.86
TaAAH	-0.02	0.80	0.21	0.08	-0.49	1.38	0.14	-0.55	-1.21	0.01	0.20	0.43	0.16	-0.64	-1.19
TaUGAH	-0.30	-0.14	0.31	-0.44	-0.64	-2.89	-0.19	-0.44	-1.36	-0.21	0.42	-0.08	0.66	0.19	-1.01
TaUAH	-0.12	0.38	0.85	-0.37	0.24	1.71	0.08	0.07	0.58	-0.06	0.86	1.70	-0.01	-0.27	-0.72

Fig. 4-3. qRT-PCR analysis of fold-change transcription of purine catabolic genes in bread wheat genotypes RAC875 under N deficiency and drought stress.

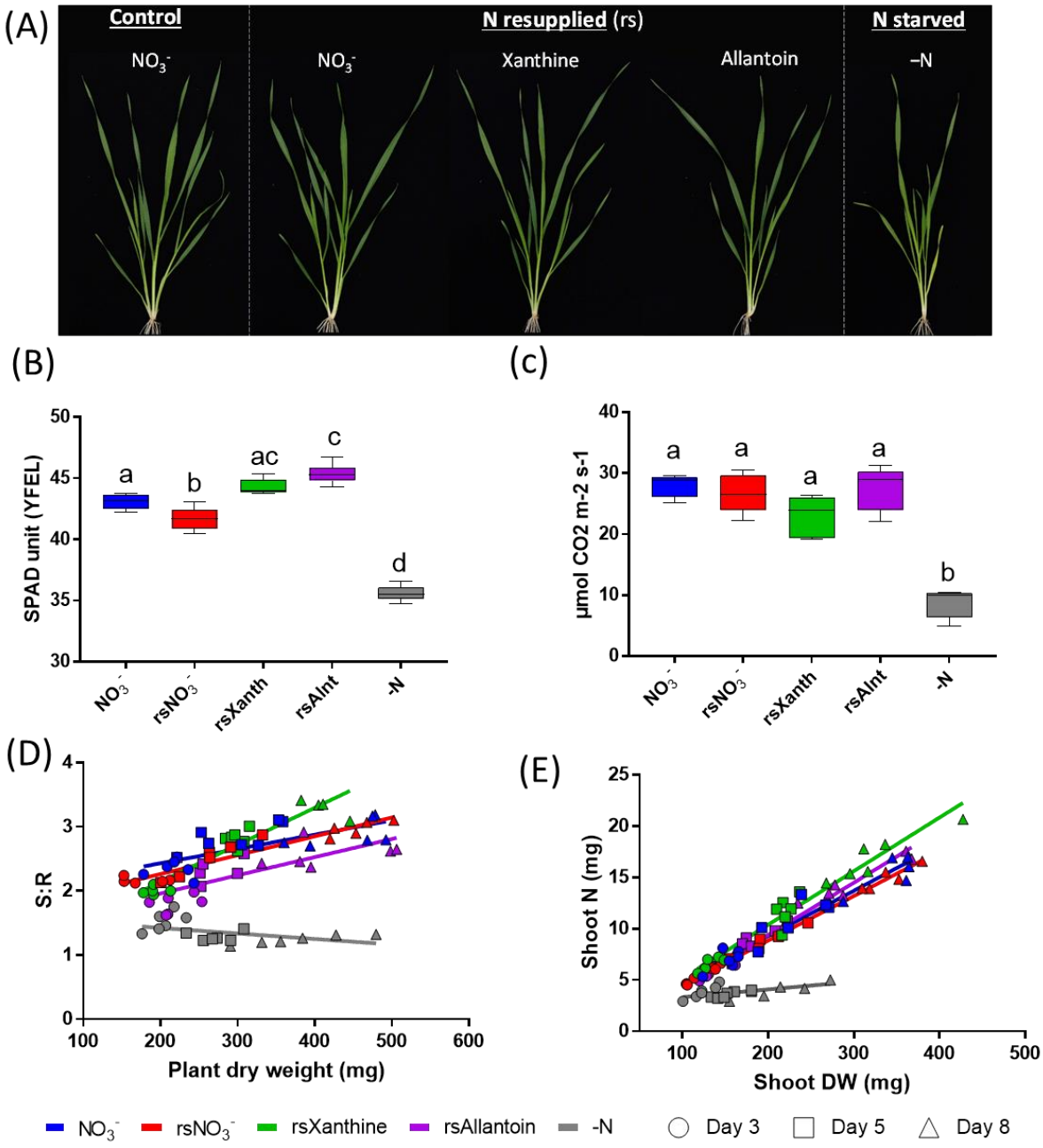


Fig. 4-4. Growth of bread wheat cv. Mace supplemented with different N sources after N starvation.

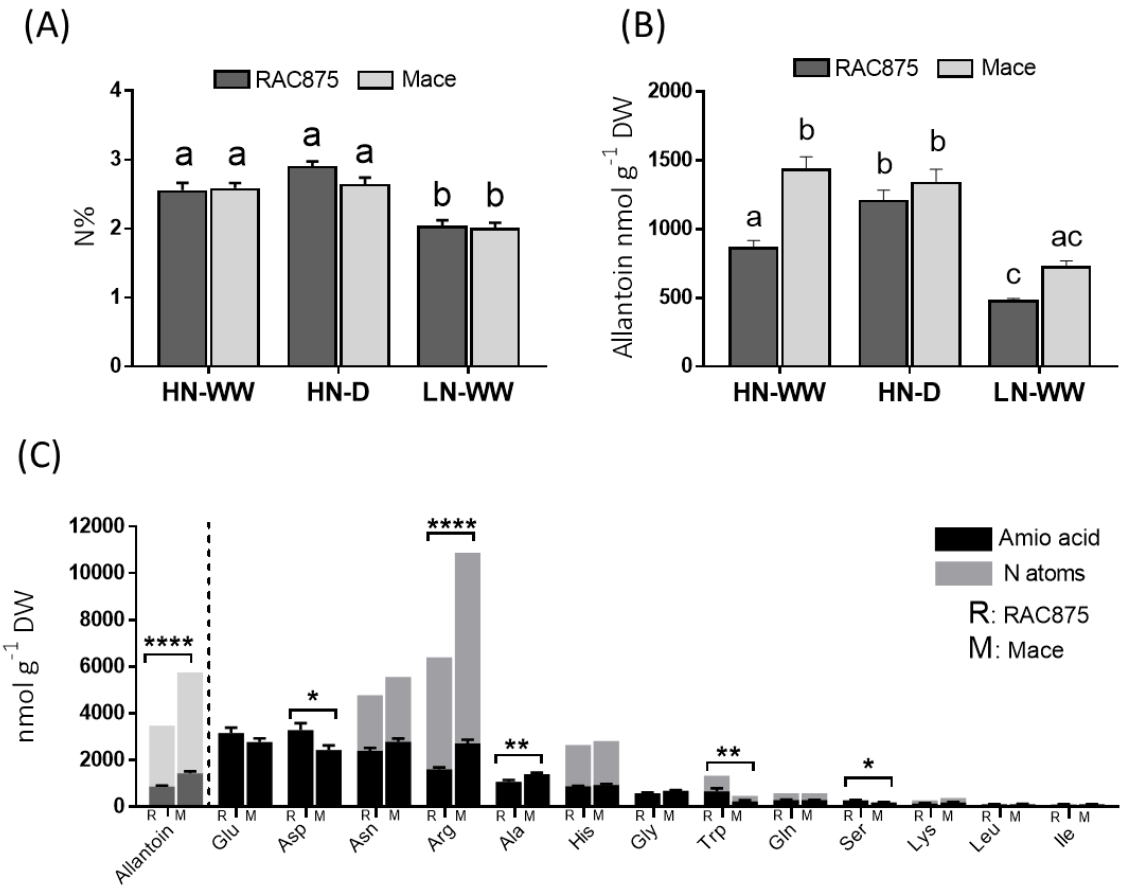


Fig. 4-5. Total N and N-containing metabolites in mature grain of bread wheat genotypes RAC875 and Mace.

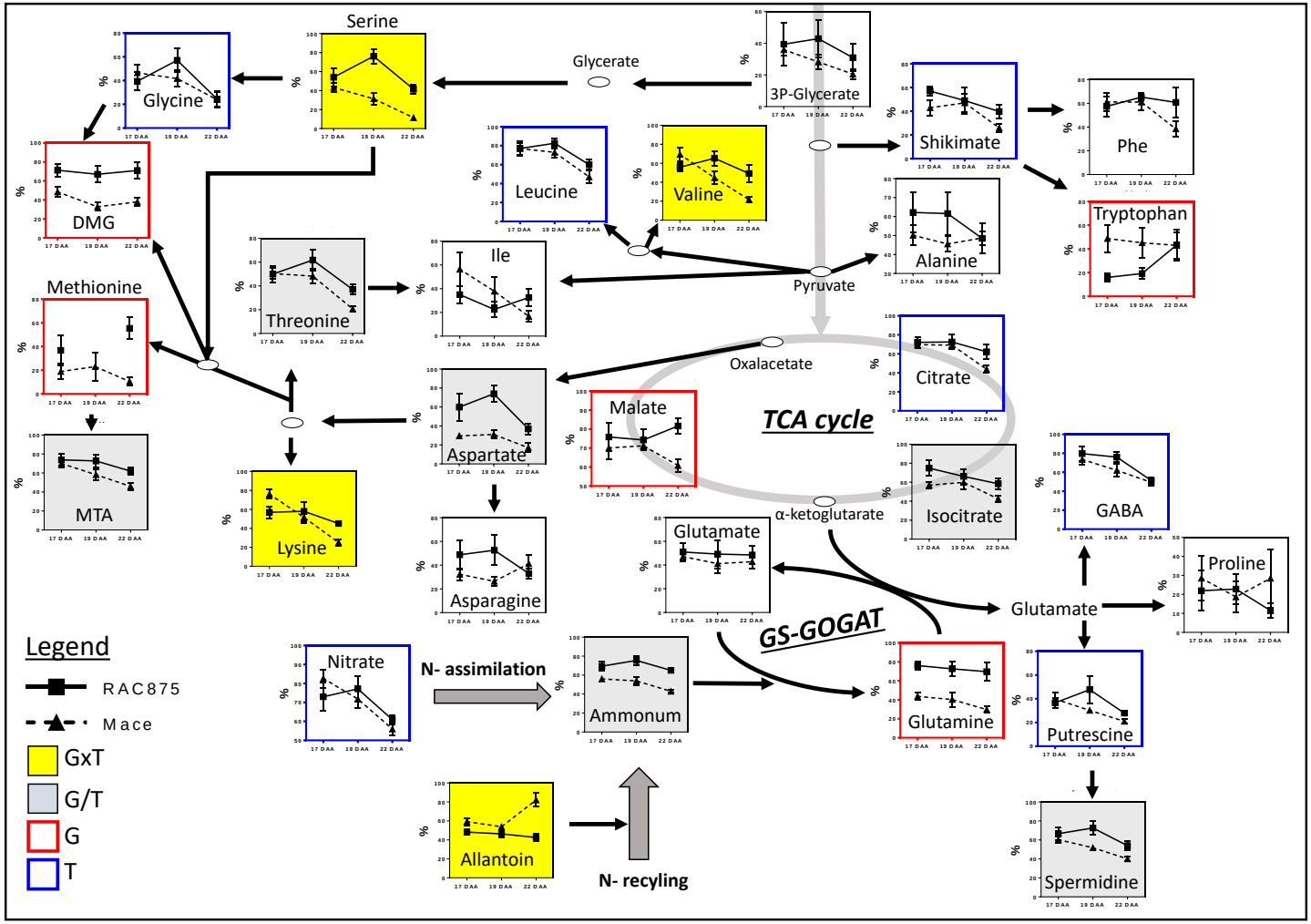


Fig. 4-6. Changes of primary metabolism during grain filling stages of bread wheat genotypes RAC875 and Mace grown under control (HN-WW) conditions.

Chapter 4
Supplementary material

Table S4-1. Primers employed for quantitative reverse transcription PCR (qRT-PCR) analysis.

Gene name	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Product size (bp)	Tm (°C)
TaXDH1	ACAGTGAAAATCGTTGGAGGA	GCAACTTGGGCTATCTTTGTG	187	60
TaXDH2	CATTGGAATCGTTGTGGCAGAC	GCTCTTGACCTCCGACTTGAAC	233	60
TaUOX	GTGAAGAAGTCTGGAAGCC	CAGGAAGAAGAGTGTAGCG	115	60
TaAS	CTGGAGATGTGGAAGGAC	TGGAGGTATTGAAACTTATGC	172	60
TaALN	TCGTTTCAAGTGCTCTCC	AGTTGCCTTCCTCCATTAG	143	63
TaAAH	TCTCTGCTCTGAAGGTCTTG	CTCCTCGTCGCTGAATGC	86	60
TaUGAH	TTAGGAGCATATTTGATTAGCC	GCAGGAAGATAAGCATAAGAG	200	60
TaUAH	TGTAACTGCCATTGCTGCTC	CAACTGTTCCAACTGTATCTATCG	179	63
TaActin	GACAATGGAACCGGAATGGTC	GTGTGATGCCAGATTTTCTCCAT	236	60
TaGAPdH	TTCAACATCATTCCAAGCAGCA	CGTAACCCAAAATGCCCTTG	197	60
TaCyclophilin	CAAGCCGCTGCACTACAAGG	AGGGGACGGTGCAGATGAA	227	60
TaEFA2	CAGATTGGCAACGGCTACG	CGGACAGCAAAACGACCAAG	227	60

Table S4-2. Optimisation of allantoin detection method.

Compound Name	Precursor Ion	Product Ion	Transition*	Collision Energy (V)
Allantoin	398	171	T	15
Allantoin	398	284	Q	15
Allantoin	398	213	Q	15
$^{13}C^{15}N$-Allantoin	400	173	T	15
$^{13}C^{15}N$-Allantoin	400	286	Q	15
$^{13}C^{15}N$-Allantoin	400	215	Q	15

Optimised GC-QqQ Multiple Reaction Monitoring (MRM) transition and collision energy for the quantification of Allantoin. Dwell time 20. *Target (T) or Qualifier (Q) transition.

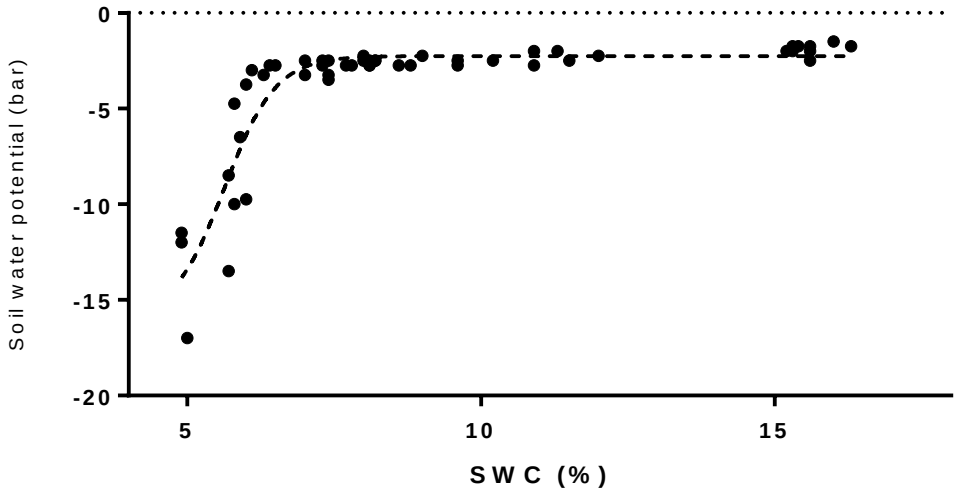
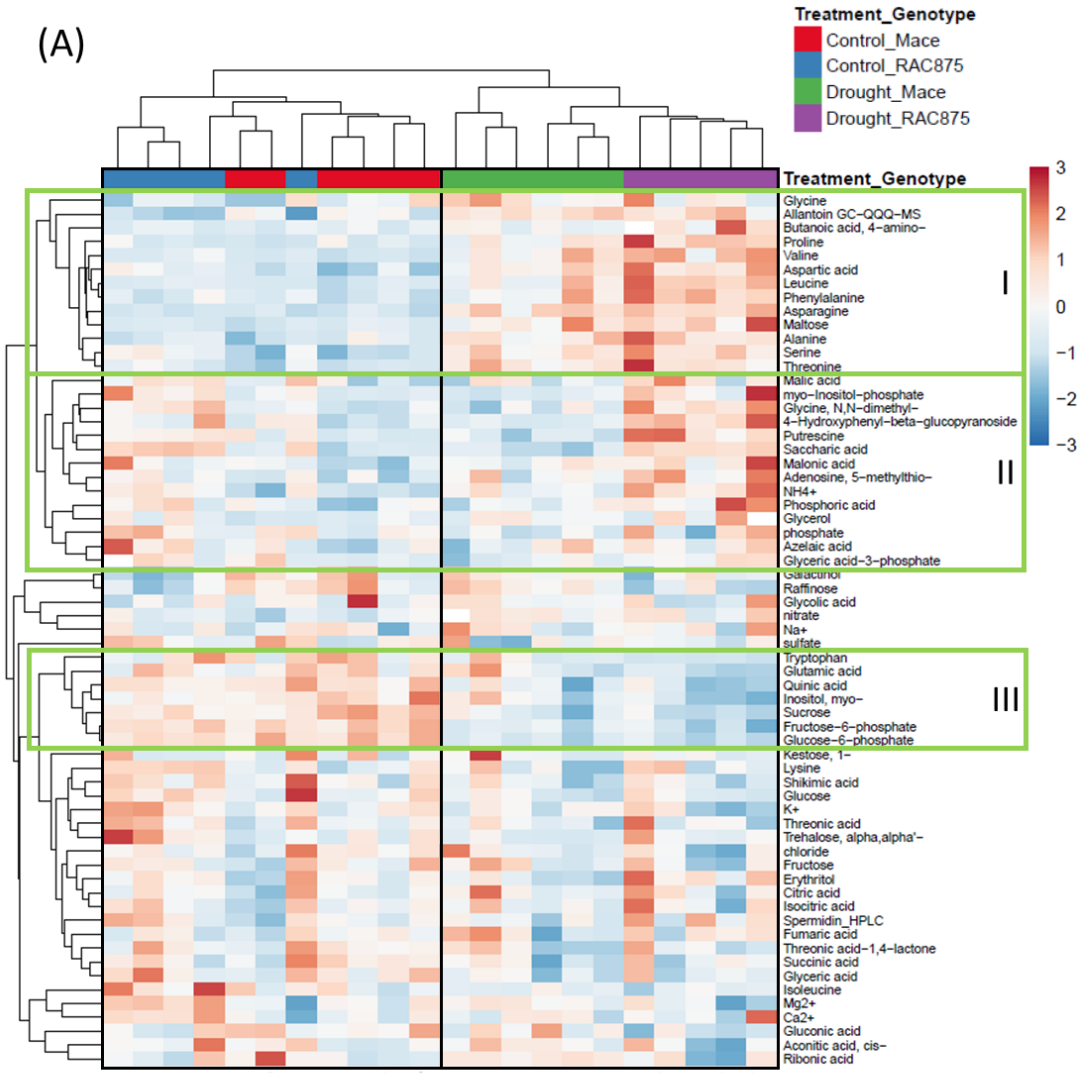


Fig. S4-1. Soil water retention curve. Soil water content (SWC) was measured by weight. Predawn leaf water potential was measured on YFEL using a Scholander-type pressure chamber (Soil Moisture Equipment Corp., Santa Barbara, USA).


(A)
Treatment_Genotype
Control_Mace
Control_RAC875
Drought_Mace
Drought_RAC875
Treatment_Genotype
Glycine
Allantoin GC−QQQ−MS
Butanoic acid, 4−amino−
Proline
Valine
Aspartic acid
Leucine
Phenylalanine
Asparagine
Maltose
Alanine
Serine
Threonine
I
Malic acid
myo−Inositol−phosphate
Glycine, N,N−dimethyl−
4−Hydroxyphenyl−beta−glucopyranoside
Putrescine
Saccharic acid
Malonic acid
Adenosine, 5−methylthio−
NH4+
Phosphoric acid
Glycerol
phosphate
Azelaic acid
Glyceric acid−3−phosphate
II
Raffinose
Glycolic acid
nitrate
Na+
sulfate
Tryptophan
Glutamic acid
Quinic acid
Inositol, myo−
Sucrose
Fructose−6−phosphate
Glucose−6−phosphate
III
Kestose, 1−
Lysine
Shikimic acid
Glucose
K+
Threonic acid
Trehalose, alpha,alpha'−
chloride
Fructose
Erythritol
Citric acid
Isocitric acid
Spermidin_HPLC
Fumaric acid
Threonic acid−1,4−lactone
Succinic acid
Glyceric acid
Isoleucine
Mg2+
Ca2+
Gluconic acid
Aconitic acid, cis−
Ribonic acid
3
2
1
0
−1
−2
−3

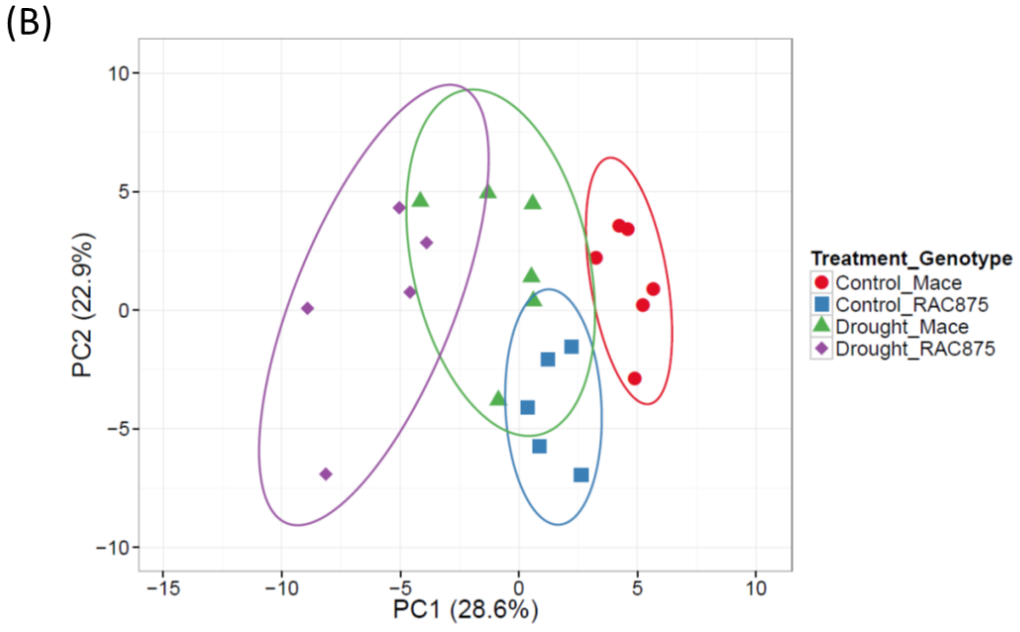

(B)
PC2 (22.9%)
PC1 (28.6%)
10
5
0
−5
−10
−15
−10
−5
0
5
10
Treatment_Genotype
Control_Mace
Control_RAC875
Drought_Mace
Drought_RAC875

Fig. S4-2. Metabolite response of RAC875 and Mace developing grain at day 8 of the drought treatment. Metabolite levels were log transformed and scaled by subtracting the median metabolite value in each metabolite distribution. (A) Hierarchical clustering and heatmap of metabolite levels of RAC875 and Mace developing grain at day 8 of the drought treatment. Pearson's correlation distance of scaled data was used for the hierarchical clustering. Representative clusters are noted with a green line. (B) Principle component analysis (PCA). Mace n = 6, RAC875 n = 5. Figures were prepared with the support of ClustVis web tool (Metsalu and Vilo, 2015).

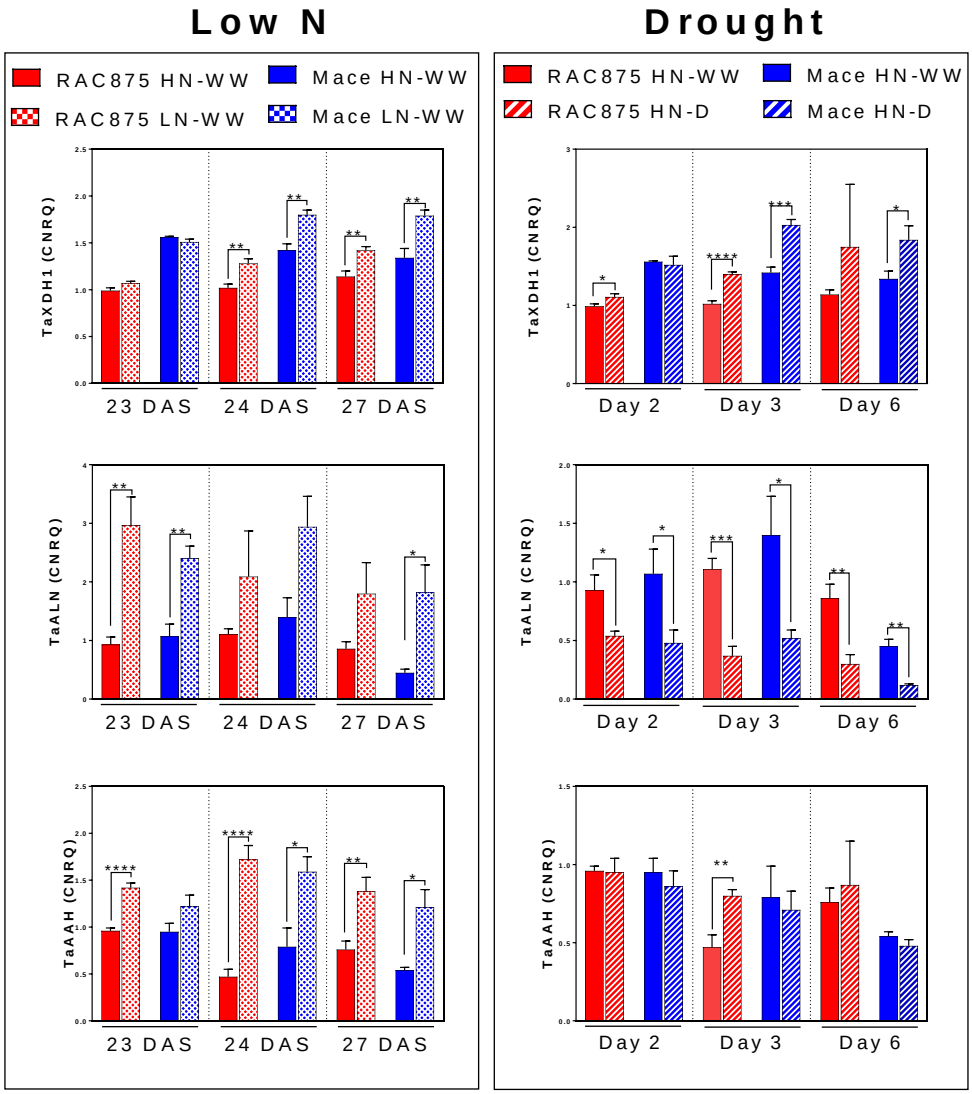


Fig. S4-3. **Comparison of relative expression of *TaXDH1*, *TaALN* and *TaAAH* in youngest fully emerged leaf (YFEL) of RAC875 and Mace under low N and drought.** RNA was extracted from YFEL collected from the main stem at 23, 24 and 27 days after sowing (DAS) (Zadoks growth stage Z22) , corresponding to 2, 3 and 6 days of drought stress, respectively. Levels of gene expression were determined by quantitative real-time PCR (qRT-PCR) employing gene-specific primers and are represented as calibrated normalised relative quantity (CNRQ) with respect to the most stable genes *TaActin* and *TaGAPdH*. Data are mean ±SEM of four to six biological replicates and asterisks indicate significant differences between control (HN,WW) and treatment (LN,WW and HN,D) per genotype per time point (*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001 by Student's t-test).

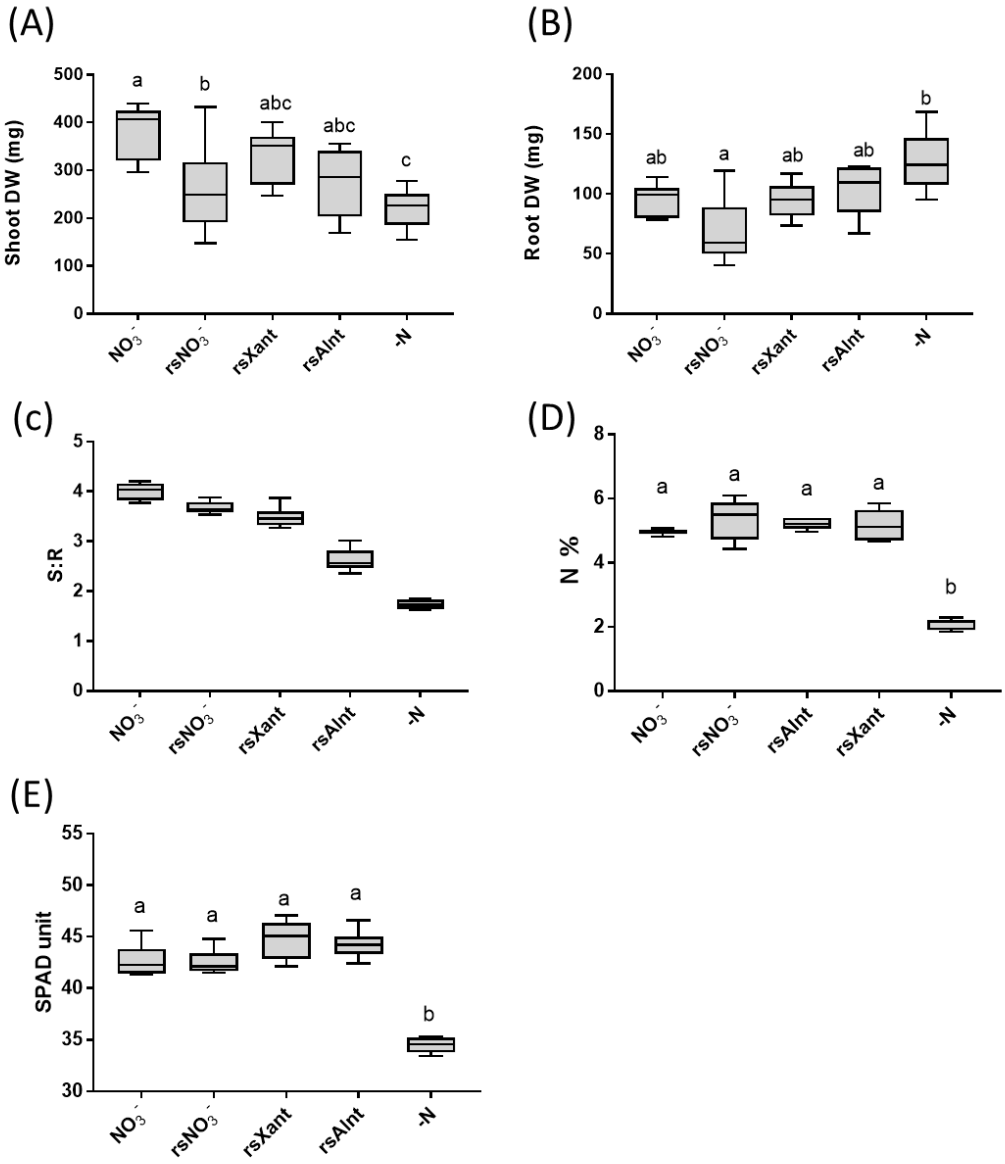


Fig. S4-4. Analysis of wheat N starved seedlings resupplied with different N sources for eight days. Mace seedlings were grown in a hydroponic system supplied with 1 mM nitrate for two weeks and then starved of N for 24 hours. N was resupplied as 1 mM nitrate ($rsNO_3^-$), 0.25 mM xanthine (rsXant), 0.25 mM allantoin (rsAlnt) or not (−N) for eight days; a subsets of plants were kept at constant nitrate supply throughout the experiment (NO_3^-). (A) shoot dry weight (DW), (B) root DW, (C) shoot to root DW ratio, (D) shoot N concentration (%) and (E) chlorophyll content expressed in SPAD unit at eight (8 d) after N resupply and without N. Letters indicate significant different between treatments per time point by one-way ANOVA with Tuckey's correction ($p<0.05$).

Chapter 5: Allantoinase (*ALN*) genetic diversity in bread wheat

External contributions to the chapter "*Allantoinase (ALN) genetic diversity in bread wheat*"

The *TaALN* sequence capture data was generated within large targeted capture experiment that was conducted in our department at the School of Agriculture, Food and Wine (The University of Adelaide, UoA). *TaALN* was kindly included to the pipeline established for other targeted genes.

The author acknowledges that the following personnel was involved in the experimental steps that led to the generation of the data used in this thesis chapter:

Dr Ute Baumann (UoA) designed and supervised the development of the capture experiment and provided support in data interpretation and chapter evaluation.

Mr Juan Carlos Sanchez Ferrero (UoA) performed the read mapping and provided support in data interpretation and chapter evaluation.

Ms Priyanka Kalambettu (UoA) performed DNA extraction from the bread wheat accessions used in this study.

The student, **Alberto Casartelli**, performed the promoter study, the post-read mapping analyses, data interpretation, figures preparation and wrote the present chapter. In addition, the student assembled the *TaALN* gene model that was used for probe designing (Chapter 3; supplementary information).

The student's supervisors, **Dr Mamoru Okamoto**, **Dr Vanessa Melino** and **Dr Sigrid Heuer** helped in data interpretation, evaluation and editing of the chapter.

Allantoinase (*ALN*) genetic diversity in bread wheat

Abstract

The nitrogen-rich metabolite allantoin is an intermediate of the purine catabolic pathway. Recent large scale metabolomics studies reported allantoin to accumulate upon abiotic stress in a broad range of plant species with a generally higher accumulation in stress-tolerant than intolerant genotypes. Mutation of the Arabidopsis allantoinase gene, which codes for the allantoin-degrading enzyme, led to improved plant performance under different abiotic stress conditions. However, the *ALN* gene remains mostly uncharacterised in the economically important crop bread wheat (*Triticum aestivum*). In this chapter the identification of highly conserved putative *cis*-regulatory elements located in the promoters of *ALN* genes from bread wheat and other grass genomes is described. These motifs are potential targets of transcription factors involved in dehydration and cold stress responses and also under limiting nitrogen (N stress) supply. In addition, a targeted *TaALN* sequence capture analysis revealed that several bread wheat accessions have highly divergent or absent *TaALN* homeologs as compared to the cultivar Chinese Spring reference genome. This suggests that genetic diversity of *ALN* exists among wheat cultivars and landraces, Accessions with putative functional polymorphisms were identified and could be used for future genetic studies.

Introduction

Allantoin and allantoate, collectively referred to as ureides, are small-nitrogen (N) rich metabolites that play an important role in plant N metabolism. For example, ureides are the major N form resulting from the symbiotic dinitrogen (N_2) fixation performed by certain tropical legumes, such as common bean and soybean (Pate *et al.*, 1980; Schubert, 1986). Ureides are also intermediates of the oxidative purine catabolic pathway. Complete degradation of ureides generates glyoxylate and CO_2, as well as ammonium which is recycled into central metabolism (Zrenner *et al.*, 2006; Werner and Witte, 2011). The enzymes

and underlying genes of purine catabolism have been identified in bacteria, fungi and animals, including insects (Vogels and Van der Drift, 1976; Ramazzina *et al.*, 2006; Scaraffia *et al.*, 2008) indicating that this pathway is ubiquitous.

In Arabidopsis the ureide allantoin has been recently reported to stimulate abscisic acid (ABA) and jasmonic acid (JA) metabolism that are important components of plant abiotic stress signalling and adaptation (Watanabe *et al.*, 2014; Takagi *et al.*, 2016). Loss of function mutants (knock-out, KO) of allantoinase (*AtALN*), the gene coding for the allantoin-degrading enzyme, constitutively accumulated about four-fold more allantoin than wild type lines (Col-0) (Watanabe *et al.*, 2014). Interestingly, *Ataln* mutants proved to be more tolerant than Col-0 lines to different abiotic stress conditions such as desiccation, drought, osmotic, salt and heavy metal. (Watanabe *et al.*, 2014; Irani and Todd, 2016; Lescano *et al.*, 2016; Nourimand and Todd, 2016). The ALN enzyme has been localised in the endoplasmic reticulum (ER) in plants (Hanks *et al.*, 1981; Werner *et al.*, 2008). However, it is still unclear how allantoin, which is synthesised in the peroxisome, is transported to the ER (Werner and Witte, 2011). Early biochemical studies suggested *ALN* is a member of the amidohydrolase family (Kim *et al.*, 2000; Mulrooney and Hausinger, 2003), despite presenting low sequence similarities with other amidohydrolases found in plants and yeast (Yang and Han, 2004). Kim *et al.* (2009), who first reported the ALN crystal structure from *Escherichia coli*, showed that the enzyme consists of a homotetramer that presents a binuclear metal centre in its active site. The metal ions were demonstrated to ligate six amino acid residues that are highly conserved across organisms as previously suggested by Yang and Han (2004). In addition, Ho *et al.* (2011) reported that manganese cations among other divalent cations activated the purified recombinant EcALN the most, in accordance with early reports showing plant ALN to be activated by manganese (Vogels and Van der Drift, 1966).

Recent comparative metabolomics studies showed allantoin to accumulate with higher magnitude upon abiotic stress in genotypes reported as tolerant under stress (Bowne *et al.*, 2011; Oliver *et al.*, 2011; Degenkolbe *et al.*, 2013; Yobi *et al.*, 2013; Nam *et al.*, 2015; Casartelli *et al.*, 2018). Allantoin levels

accumulated by plants under stress can be increased by either enhancing its synthesis and/or by inhibiting its degradation. Irani and Todd (2016) reported that allantoin accumulated in Arabidopsis leaves after five days of drought treatment, which was paralleled by up-regulation of purine catabolic genes involved in allantoin synthesis, namely xanthine dehydrogenases (*AtXDH1* and *AtXDH2*), urate oxidase (*AtUOX*) and allantoin synthase (*AtAS*). On the other hand, the authors reported that under salt and osmotic (mannitol) stress, *AtALN* was down-regulated after four days of treatment in comparison to control conditions. Similar observations were reported by Lescano *et al.* (2016) who showed that *uida* gene activity (GUS staining) under the control of *AtALN* promoter decreased with the increase of NaCl concentration. In bread wheat (*Triticum aestivum* L.) Casartelli *et al*. (Chapter 4) reported that allantoin accumulated upon drought in leaves during vegetative growth and in flag leaves during grain filling stages. This was accompanied by down-regulation of *TaALN* at mild drought stress and up-regulation of *TaXDH1*, *TaUOX* and *TaAS* under medium to severe drought stress conditions. In addition, allantoin levels in wheat plants grown under N deficiency were significantly reduced and, at the same time, *TaALN* was strongly up-regulated (Casartelli *et al*., Chapter 4). Similarly, Yang and Han (2004) reported that *ALN* mRNA levels increased under N deficiency in both in Arabidopsis and the non-ureide type legume tree Black Locust (*Robina pseudoacacia*).

Taken together, these pieces of evidence suggest that under stress allantoin levels are associated with transcriptional changes in purine catabolic genes. The fact that Arabidopsis *aln* KO mutants showed improved performances under abiotic stress than wild type lines (Watanabe *et al*., 2014; Irani and Todd, 2016; Lescano *et al*., 2016; Nourimand and Todd, 2016) further suggests that *TaALN* is a key player in stress responses. The main aim of the present study was to further characterise *TaALN* which was first reported in wheat by Casartelli *et al.* (Chapter 3). Phylogenetic analysis of the *ALN* promoters in grasses allowed the identification of putative *cis*-regulatory elements presenting putative binding sites for transcription factors involved in stress responses. Targeted capture and sequencing of *TaALN* homeologs from a bread wheat diversity panel revealed many *TaALN* sequence variations in relation to the wheat

Chinese Spring reference genome. In addition, the analysis allowed us to identify accessions carrying possible *TaALN* homeologs number variations or functional polymorphisms. The ultimate goal is to test these accessions in relation to allantoin accumulation under stress and identify possible association of exotic *TaALN* alleles with improved performances.

Materials and Methods

Phylogenetic promoter analysis

A 2 kb region upstream of the translational start site (ATG) was considered as the putative promoter for each target gene. In particular, the *ALN* promoter regions of rice (*Oryza sativa*), maize (*Zea mays*), sorghum (*Sorghum bicolor*) and Brachypodium (*Brachypodium distachyon*) genomes were retrieved from the Phytozome BioMart tool (http://phytozome.jgi.doe.gov). ALN promoter regions of barley (*Hordeum vulgare*), *Aegilops tauschii*, *Triticum urartu* and bread wheat (*Triticum aestivum*) were retrieved from genomic databases that are listed in Casartelli *et al.* (Chapter 3; Supplementary material). The retrieved promoter sequences were analysed using MEME Suite 4.12.0, with the following parameters: zero or one occurrence per sequence, maximum number of motif = 30, 0-order as background model and optimum motif width 2-30 base pair (Bailey *et al.*, 2006). Top candidate motifs were selected based on the following criteria: E-value ≤ 0.01, SEM of the start site ≤ 100 bp and number of sequence 'Hits'≥ 9. Motifs were ranked according to E-value. A second list of candidate motifs was generated allowing number of 'Hits' ≥ 5 and then ranked according to E-values. To aid the comparison with the sequence capture experiment presented below, motif positions are reported as bp distance from position 1 that corresponded to -2000 bp upstream the ATG site. Orthologous pALNphy sequences related to top candidate motifs were submitted to a database of Plant Cis-acting Regulatory DNA Elements (PLACE) (Higo *et al.*, 1999). Putative transcription factor (TF) binding sites were considered as relevant if they were found in at least

50% of the orthologous sequences and occurred at the same position within the specific pALNphy motif where they were identified.

Multiple sequence alignment

Sequence alignments of *ALN* promoter regions, CDSs and proteins were performed using Geneious 10.0.2 software (www.geneious.com). For nucleotide alignments, the identity matrix was constructed using ClustalW on Geneious with default parameters (gap open cost: 15; gap extend cost: 6.66) and phylogenetic trees were generated with Geneious Tree Builder with default parameters (substitution model: HKY; tree build method: Neighbor-Joining; number of replicates: 1000). ALN protein alignment was performed using the MUSCLE algorithm (Edgar, 2004) with Geneious 10.0.2 default parameters (distance measure: kmer6_6 and pcitid_kimura; clustering method: UPGMA; tree rooting method: pseudo; sequence weighting scheme: ClustalW). Analysis of polymorphism profiles was performed with Jalview 2.10.2b1 software (Waterhouse *et al.*, 2009) using the percentage of identity (PID) method.

Capture sequencing experiment

Targeted capture sequencing experiment was designed and performed by the Plant Genomics Centre (PGC) Bioinformatics group (University of Adelaide). DNA was extracted from single leaf material using an extraction buffer composed of 1% w/v sarkosyl, 100 mM Tris-HCl, 100 mM NaCl, 10 mM EDTA and 2%w/v polyvinyl-polypyrrolidone (PVPP) adjusted to pH 8.5 and a phenol/chloroform/iso-amyl alcohol (25:24:1) solution as a separating agent. DNA was sourced from 135 hexaploid bread wheat accessions that are part of a wheat diversity panel collected and available for research purposes at the University of Adelaide (Fleury *et al.,* unpublished). *TaALN* homeolog genomic regions, including 2kb upstream and downstream the translation starting site (ATG) and end (STOP), were provided to MyBaits (Michigan, USA) for the design of suitable capture probes. The gene models (exon/intron boundaries) were based

on the IWGSC Chinese Spring WGA V0.3 release (https://wheat-urgi.versailles.inra.fr/Seq-Repository/Assemblies). Preparation of DNA libraries of the selected bread wheat accessions, hybridisation and sequencing (HiSeq 100 bp paired end) was performed by MyBaits according to their established pipeline.

Sequence analysis

Reads that were generated by the sequencing of the captured samples were checked using FastQC (Babraham Bioinformatics, UK; www.bioinformatics.babraham.ac.uk). Reads were hard-clipped at the 5' end (4 bp) and 3' end (3 bp) using Trimmomatic (Bolger *et al.*, 2014) and paired-end mapped using Bowtie2 (Langmead and Salzberg, 2012). Bowtie2 parameters were optimised for reads mapping to Chinese Spring reference and allowed: (i) two mismatches, (ii) one mismatch and one insertion/deletion (INDEL) or (iii) two-three INDELs per read (93 bp). In order to identify polymorphisms in the TaALN genomic regions, a file containing the information of each position (pileup file) was generated using SAMtools software (Li *et al.*, 2009). The subsequent SNP calling was performed by an in-house software (Fruzangohar et al., unpublished, University of Adelaide) that considered only positions with at least 10X coverage and a minimum of 70% of same nucleotide in order to minimise sequencing errors.

Results and discussion

***ALN* promoter regions contain possible cis-elements responsible for regulation under stress**

Higher magnitudes of allantoin accumulation during abiotic stress have been reported to occur in genotypes that were defined as more tolerant. However, a question that needs to be addressed is whether allantoin and *ALN* are factors directly involved in determining stress tolerance or are components that lie downstream of the tolerant response. As a first step, a phylogenetic analysis of *ALN* promoter (pALN)

regions among grasses genomes was performed in order to collect some preliminary evidence of stress responsive regulation. The grass species taken into account in this study were: *Triticum aestivum*, *Aegilops tauschii*, *Triticum urartu*, *Hordeum vulgare*, *Brachypodium distachyon*, *Oryza sativa, Zea mays* and *Sorghum bicolor.* Global alignment of the 2 kb upstream region of *ALN* revealed 27-57% sequence identity between these species, whilst *ALN* CDS alignment showed higher sequence identity of 82-96% (Fig. S5-1C). The phylogenetic relationship between the promoter sequences was different from the relationship shown by the CDS alignment (Fig. S5-1B, D). For example, the barley *ALN* promoter sequence (pHvALN) clustered together with pTaALN-2AL, in contrast to the CDS phylogenetic tree, in which *HvALN* positioned before wheat *ALN* CDSs (Supplementary Fig. S5-1D). As sequence conservation of promoter regions is expected to be lower than for CDS it is not surprising that the evolutionary relationship of the former deviated from the latter.

Identification of highly conserved motifs in grass ALN gene promoters

Putative conserved *cis*-elements in the promoters of orthologous *ALN* genes (pALNphy) were identified using MEME that discovers novel and ungapped motifs in a set of query sequences. The list of identified motifs was reduced to five top candidates based on the following selection criteria: E-value ≤ 0.01, SEM of the start site ≤100 and number of sequence (Hits) ≥ 9 (Table 5-1). This allowed us to select motifs that were highly conserved amongst species occurring in similar positions (low SEM scores) within the analysed putative promoter regions. An additional list of candidate motifs was composed by selecting Hits ≥ 5, while keeping the remaining criteria as above (Table S5-1). With this, we could select motifs that were conserved in closely related species (e.g. bread wheat homeologs and wheat ancestors diploid genomes). Among the top candidates, pALNphy1 and pALNphy8 motifs were found highly conserved across all species (Table 5-1). In particular, of the top candidate motifs, pALNphy1, contained a highly conserved short 10 bp sequence (5'-GCCA[C/G]TTGGA-3') with almost 100% identity between all species

(Table 5-1, Fig. S5-2). pALNphy5 and pALNphy7 were found conserved in all promoter regions expect for *Oryza sativa* (pOsALN), whilst pALNphy19 was absent in *Zea mays* (pZmALN). Motif pALNphy5 contained a short 5 bp sequence (5'-CCAAT-3') with 100% sequence identity among all analysed species (Table 5-1, Fig. S5-2).

Table 5-1. Candidate motifs identified in the ALN promoter region of grass genomes with MEME web tool

Motif ID	Motif sequence	Hits	E-value	Motif start position	
				Mean	SEM
pALNphy1	[T/C]C[C/T]C[C/A][G/A][T/A][G/A][G/A][G/T][C/T][C/T][C/T][C/A][G/A]GCCA[C/G]TTGGA[G/T][G/C][G/C][A/C][T/G/C]	10	4.60E-51	1659.80	9.44
pALNphy5	G[G/C][G/A/C][G/C][C/A][A/G][G/A/C]CCAAT[G/C][A/G][A/C][G/A][A/G/C][G/C][C/G]C[G/T][T/C]C[A/G][C/A][T/C/A][G/T][A/C][G/C]	9	2.20E-32	1469.67	59.87
pALNphy8	[C/G/A][G/C/A][C/A/T/G][A/T/G][C/G/A][G/C/A]C[A/G][G/C][A/G][C/G]GC[A/G]G[A/C][C/A][G/A][G/C/A][G/A][G/C]A[C/G][G/A][G/A/C][G/C/A][A/T][C/G][T/G][C/A/T]	10	9.80E-25	1919.60	4.80
pALNphy19	[T/G][G/A/T][A/G][A/C][G/T/A][G/A/T][G/A]A[T/C/G][G/T][G/A][A/T][T/G][T/C/A][A/G][A/G][C/T/G][C/G][A/T/G][T/G]TT[G/T][G/C][G/T][T/G/C/A][T/G][G/C][G/C/A]	9	2.20E-08	1290.22	80.97
pALNphy7	[C/G][A/C][G/C/A][T/C][G/C]C[G/A/C]C[C/T]G[C/A/G]G[T/C/A][C/G/T][C/G][A/C/G]C[G/C][G/C][C/T/A][G/C]C[G/A/C]G[T/C][G/T][C/G][C/AG][C/T/A][C/G]	9	2.3e-029	1804.56	23.01

ALN promoter regions (2 Kb upstream of the translational start site ATG) of *Triticum aestivum*, *Aegilops tauschii*, *Triticum urartu*, *Hordeum vulgare*, *Brachypodium distachyon*, *Oryza sativa*, *Zea mays* and *Sorghum bicolor* were analysed with a phylogenetic approach using MEME v4.12.0 and conserved short regions were identified (pALNphy). Motif sequence is reported in the second column and in brackets are reported nucleotide ambiguities for a specific position within the motif. Hits indicates the number of promoter orthologous sequences in which a specific pALNphy was identified. Selection criteria were applied as follow: E-value ≤ 0.01, SEM of the motif start site position ≤100 and number of sequence "Hits" ≥9. A list of top motifs was generated. Hits refers to number of orthologous promoter regions containing a specific motif (total orthologous sequence = 10).

Putative regulatory function of pALNphy motifs

To determine whether any published TF binding sites are associated with the candidate pALNphy motifs identified using MEME, the Plant Cis-acting Regulatory DNA Elements (PLACE) database (Higo *et al.*, 1999) was interrogated. PLACE is a collection of previously published regulatory elements with information in regard to the species in which the element was discovered and characterised. It also provides information on the transcription factor (TF) that is associated with a specific element. Within individual pALNphy motifs multiple TF binding sites were identified and in Table 5-2 are those that were represented in at least 50% of the pALNphy orthologous sequences and started at the same position. A graphical representation of putative TF binding sites occurring in the identified conserved pALNphy motifs is provided in Fig. 5-1. Overall, a high degree of conservation of the top candidate motifs was observed among orthologous sequences with respect to their position from the transcriptional start site (ATG) and their order of occurrence. This suggests that a certain degree of selective pressure exists and that the identified motifs might be important for regulation of *TaALN* expression. In particular, PLACE identified a 6 bp regulatory element (5'-CANNTG-3') in pALNphy1, pALNphy19 and paALNphy5. Interestingly, in pALNphy1 the 5'-CANNTG-3' was present in the 10 bp sequence (5'-GCCA[C/G]TTGGA-3'; underlined bases) highly conserved across the species mentioned above. Two functional annotations were associated with the 5'-CANNTG-3' element, namely EBOXBNNAPA and MYCCONSENSUSAT. The former was identified in the promoter of napA storage protein of oilseed rape (*Brassica napus* L.) and was considered an E-box/ABRE(ABA-responsive element)-like sequence (Stalberg *et al.*, 1996). The latter was identified in *Arabidopsis thaliana* as a MYC recognition site in the promoter of the dehydration-responsive gene *rd22* (Abe *et al.*, 2003) and other promoters of genes involved in abiotic stress tolerance, such as the cold responsive gene *ICE1* (Chinnusamy *et al.*, 2003). The 5'-CANNTG-3' regulatory element was identified in abscisic acid (ABA)-regulated genes. This is of particular interest as allantoin, the substrate of the ALN enzyme, has been shown to activate both, ABA and JA pathways in Arabidopsis (Watanabe *et al.*, 2014; Takagi *et al.*, 2016). Conversely, ABA was shown to induce *ALN* expression in common bean (Alamillo *et al.*, 2010), suggesting the existence of a regulatory cross-talk. Recently, Li *et*

al. (2015) identified a CRT/DRE 5'-CCGAC-3' regulatory element in the rice core promoter of the purine catabolic gene *OsUAH* conferring transcriptional up-regulation under cold stress. The authors showed that the CRT/DRE element was activated *in vivo* by rice C-Repeat-Binding Factor 3 (OsCBF3). The grass *ALN* promoters reported here were therefore interrogated specifically for this element but, although the short motif had overall 26 hits, there was no apparent conservation with regard to their relative position (data not shown). This lack of conservation suggests the CRT/DRE 5'-CCGAC-3' regulatory element is not evolutionary related with *ALN* genes and may be specific to the Os*UAH* gene.

Within pALNphy5, a 5 bp motif (5'-CCAAT-3') was identified in all orthologous promoter sequences (Table 5-2; Fig. 5-3). This regulatory element, also called CCAAT-box, is considered part of the core promoter and found in about 30% of eukaryote gene promoters as a single-copy element (Bucher, 1990). This short element is recognised by CCAAT-box binding factor family (also known as Nuclear Factor Y, NF-Y), a type of TFs present in all eukaryotes. In plants, *NF-Y* genes have evolved into two distinct groups: one involved in the regulation of more general processes and one regulating pathways with more specific roles, such as ABA-related signal transduction pathway (for a review see Laloum *et al.*, 2013). NF-Y proteins are complexes of three subunits (NF-YA, NF-YB, NF-YC), each encoded by a multigene family, resulting in a large heterogeneity of possible combinations. Several members of these families have been reported to be involved in various processes, such as flowering time, N-nutrition and drought (see references within Laloum *et al.*, 2013). In particular, Zhao *et al.* (2011) showed that Arabidopsis NF-YA family members were induced upon N starvation, while the microRNA miR169, which targets NF-YA members, was down-regulated.

The 5'-TGAC-3' sequence found in pALNphy5 was identified by PLACE as the W-box element, WRKY71OS, known to be recognised by WRKY TFs. WRKY involvement in abiotic stress responses is well established and was reviewed by Rushton *et al.* (2010). Interestingly, Bi *et al.* (2007) reported a WRKY TF (*At1g80840*) to be induced under mild N limitations. This evidence suggests that the identified regulatory elements may be recognised and bound by transcription factors also during *TaALN* regulation

under low N, which supports the data reported by Casartelli *et al.* (Chapter 4) showing increased *TaALN* expression under N deficiency. Taken together, these lines of evidence support that the *TaALN* promoter region of grass genomes contains motifs that may be recognised by TFs with functions in abiotic stress and response to N supply.

Table 5-2. Known cis-acting regulatory elements contained in conserved motifs of ALN promoter in grass genomes.

pALNphy ID	Occurrence (%)	Site Name	Position (bp)	Regulatory Element	Key words
pALNphy1_A	100	EBOXBNNAPA	18	CANNTG	napA; storage protein; ABRE; E-box; seed;
pALNphy1_A	100	MYCCONSENSUSAT	18	CANNTG	MYC; rd22BP1; ABA; leaf; seed; stress; CBF3; cold; CBF/DREB1;ICE1; RRE;
pALNphy1_A	80	CACTFTPPCA1	18	YACT	mesohpyll; CACT;
pALNphy1_B	80	SORLIP1AT	16	GCCAC	phyA; phytochrome; light;
pALNphy1_C	60	SITEIIATCYTC	7	TGGGCY	cytochrome; TCP-domain; meristem; oxidative phosphorylation;
pALNphy1_D	60	SORLIP2AT	8, 9	GGGCC	phyA; phytochrome; light;
pALNphy19_A	56	EBOXBNNAPA	18	CANNTG	napA; storage protein; ABRE; E-box; seed;
pALNphy19_A	56	MYCCONSENSUSAT	18	CANNTG	MYC; rd22BP1; ABA; leaf; seed; stress; CBF3; cold; CBF/DREB1;ICE1; RRE;
pALNphy19_B	56	SEF3MOTIFGM	22	AACCCA	SEF3; beta-conglycinin; 7S; globulin; seed;
pALNphy5_A	100	CCAATBOX1	8	CCAAT	HSE (Heat shock element); CCAAT box;
pALNphy5_A	100	CAATBOX1	9	CAAT	CAAT; legA; seed;
pALNphy5_B	67	ASF1MOTIFCAMV	20	TGACG	TGACG; root; leaf; CaMV; 35S; promoter; auxin; salicylic acid; light; as-1; TGA1a, TGA1b; CREB; ASF1; TGA6; shoot; xenobiotic stress; SAR; SA; Disease resistance;
pALNphy5_C	67	WRKY71OS	21	TGAC	WRKY; GA; MYB; W box; TGAC; PR proteins;
pALNphy5_D	67	GTGANTG10	22	GTGA	g10; pollen; pectate lyase;
pALNphy5_E	56	EBOXBNNAPA	23	CANNTG	napA; storage protein; ABRE; E-box; seed;
pALNphy5_E	56	MYCCONSENSUSAT	23	CANNTG	MYC; rd22BP1; ABA; leaf; seed; stress; CBF3; cold; CBF/DREB1;ICE1; RRE;
pALNphy7_A	56	CACTFTPPCA1	2	YACT	mesohpyll; CACT;
pALNphy7_B	56	CGCGBOXAT	8	VCGCGB	calmodulin

Top motifs discovered with MEME and presented in Table 1 were searched on the Plant Cis-acting Regulatory DNA Elements (PLACE) database that contains data from previously published reports (sogo.dna.affrc.go.jp/cgi-bin/sogo.cgi?lang=en&pj=640&action=page&page=newplace). Within a specific conserved motif (pALNphy), a regulatory element was selected if represented in at least 50% of the orthologous motifs (occurrence) and if identified at same location within the pALNphy motif (position).

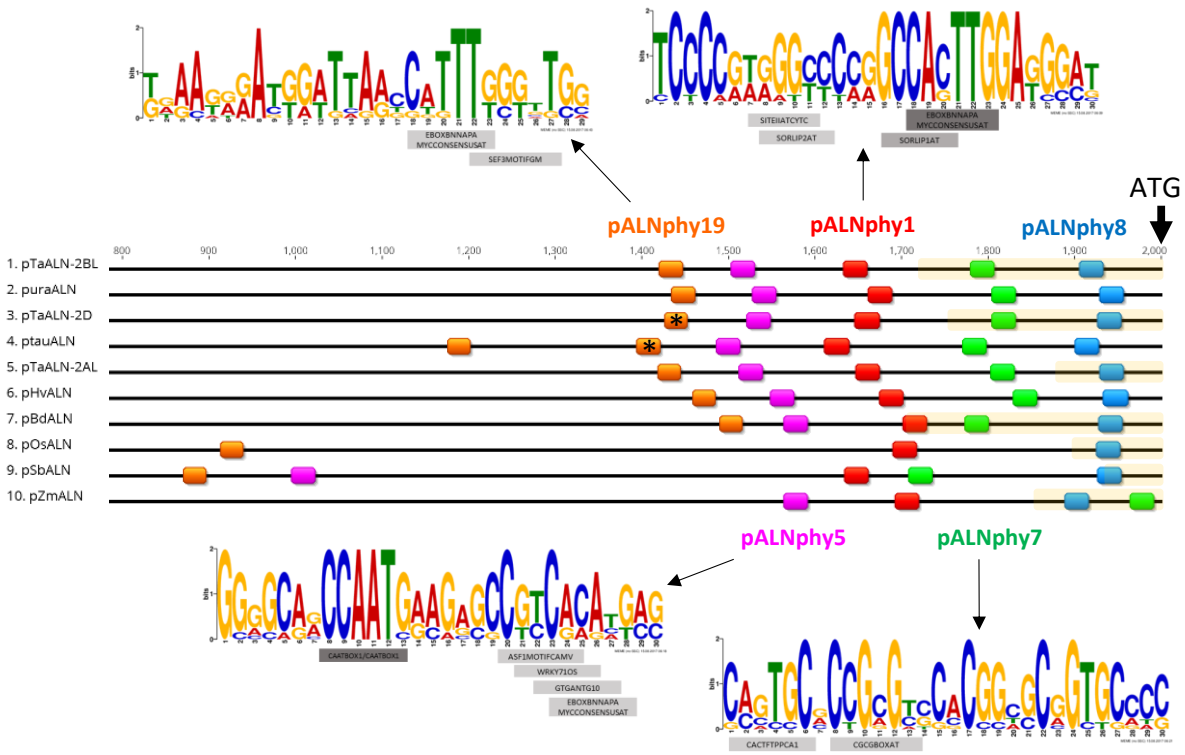


Fig. 5-1. Graphical summary of ALN promoter analysis. Top conserved motifs (pALNphy) discovered with MEME web tool (Table 1) are represented in the promoter region of each organism as boxes, colour indicates different motifs and the motif ID is reported accordingly. Yellow shade indicates the putative 5' untranslated region (5'UTR) identified by searching Phytozome or EnsemblPlants databases for expressed sequence upstream the ATG site (start of translation). Nucleotide details of the identified conserved motifs are represented as logos and the size of each letter (A, T, G, C) is proportional to the frequency of occurrence of a certain nucleotide at a specific location. Regulatory elements identified with PLACE web tool are represented below the specific motif logo as grey boxes. Names of regulatory elements are provided and darker shade indicates higher occurrence (hits) among orthologous sequences. Note that for motif pALNphy8 no regulatory elements satisfying the selection criteria (Table 5-2) were found and, therefore, is not represented here. Note that MEME discovered the pALNphy19 motif in ptauALN sequence approximately at 1200 bp, however an additional motif at 1400 bp presenting 100% identity with the motif discovered on pTaALN-2DL sequence was found and indicated with an asterisk.

Characterisation of *TaALN* haplotypes among a bread wheat diversity panel

In recent times, wheat breeding programs have seen a depletion of genetic diversity in their germplasms. The limited number of elite cultivars that are used as progenitor material is believed to be a major bottleneck. Germplasm exchange and introduction of wheat landraces, defined as traditional varieties with higher tolerance to abiotic and biotic stresses, are of great value for breeders (Lopes *et al.*, 2015). The discovery of allelic variants of genes involved in stress adaptation from wheat landraces could be used for DNA marker development and employed in large scale screenings. Identification of candidate wheat accessions carrying allelic variants of *TaALN* homeologs is therefore interesting for studying the association of allantoin accumulation with enhanced performances under stress. For this reason, the genetic diversity in *TaALN* homeologs was assessed across a bread wheat diversity panel of 135 accessions, which included commercial cultivars and landraces from different countries (Table S5-2). Briefly, *TaALN* probes were designed by MyBaits (Michigan, USA) to target the genomic regions of *TaALN* homeologs based on the Chinese Spring reference genome. The target genomic sequence captured 2 kb upstream and downstream of the *TaALN* ATG start and STOP codons, respectively. DNA from the wheat accessions was hybridised with the *TaALN* probes, the enriched genomic DNA fragments were sequenced and the reads mapped to the Chinese Spring reference genome (see M&M for details).

As a general observation, large differences in read coverage were observed among *TaALN* homeologs. In particular, long stretches of *TaALN* homeologs were not represented with any reads in several accessions. However, it is important to consider that poor read coverage for a specific accession may indicate that: (*i*) a specific *TaALN* homeolog is absent and the mapped reads are the results of unspecific hybridisation (also called cross-hybridisation) of closely related genes (e.g. amidohydrolase gene family), repetitive elements or transposons; (*ii*) the *TaALN* homeolog in certain accessions may be highly divergent compared to the reference sequence (Chinese Spring) and reads did not meet the mapping criteria (e.g. more than two mismatches per read; see M&M).

Classification of polymorphisms into allele groups revealed large differences among TaALN homeologs. Wheat accessions that are closely related are expected to present similar polymorphism and, thus, analysis of the identity could provide an insight into their pedigree. The polymorphisms identified for each *TaALN* homeolog and accessions by the SNP calling script (see M&M) were merged into nucleotide strings (polymorphism profiles) and analysed by distance matrix. For each homeolog the results were visualised as a circular tree and accessions were classified in major allele groups (ag-A1-4, ag-B1-3 and ag-D1-3) based on the tree branch structure (Fig. 5-2A; Fig. 5-3A; Fig. 5-4A). It is worth remarking that the grouping presented here depended solely on the polymorphisms that were detected by the targeted capture experiment, therefore, it should be considered as a preliminary classification of putative *TaALN* haplotypes. *TaALN-2AL* polymorphism profiles revealed four allele groups, namely ag-A1 (including 124 accessions), ag-A2 (nine accessions), ag-A3 (single accession SL232135) and ag-A4 (single accession SL232199) (Fig. 5-2A). Polymorphism profiles of the *TaALN-2BL* homeologs were grouped into ag-B1, ag-B2 and ag-B3 containing 66, 45 and 24 accessions, respectively (Fig. 5-3A). Polymorphism profiles of the *TaALN-2DL* homeologs were grouped into ag-D1, ag-D2 and ag-D3 containing 106, 27 and two accessions, respectively (Fig. 5-4A). A graphical visualisation of the full read mapping of representative accessions belonging to different *TaALN* homeologs allele groups is reported in Fig. 5-2B; 5-3B and 5-4B. The polymorphism profile of Chinese Spring (SL232061), which is the reference used in the experiment, was found in the largest allele groups of *TaALN-2BL* and *-2DL* homeologs (ag-B1 and ag-D1). In contrast, the polymorphism profile of Chinese Spring *TaALN-2AL* homeolog grouped in ag-A2 that only contained eight other wheat accessions. In addition, accessions that belonged to the same allele group as Chinese Spring had high read coverage, whilst those classified in other groups had moderate to low coverage and a high number of polymorphisms (Fig. 5-2B; Fig. 5-3B; Fig. 5-4B). This is not surprising and indicates that the results of the sequence capture experiment are highly depended on sequence similarity with the Chinese Spring, as aforementioned. In addition, mapped reads of Chinese

Spring also showed segments where the coverage was low (e.g. about 2000 bp and 6000 bp of *TaALN-2AL*; Fig. 5-2B). This may be due to a high GC% content which could reduce sequencing efficiency.

Whole-genome next generation sequencing data confirm low TaALN reads coverage when Chinese Spring is used as reference.

Visualisation of accessions belonging to ag-A1 (Fig. 5-2B), ag-B2 and ag-B3 (Fig. 5-3B) and ag-D2 and ag-D3 (Fig. 5-4B) revealed that only short portions of the targeted *TaALN* genomic regions had reads aligned. This may indicate that a *TaALN* homeolog is absent in that specific accession and the mapped reads are the result of cross-hybridisation or that the reads did not meet the mapping criteria in relation to the reference Chinese Spring, as aforementioned. In order to corroborate this, the *TaALN* genomic regions were analysed with the in-house software DAWN (Diversity Among Wheat geNomes) (Baumann et al., unpublished) (Fig. S5-3). This tool combines IWGSC Chinese Spring WGA V0.3 genome release data with the BPA (Bioplatform Australia) whole genome shotgun sequencing data of different wheat accessions and IWGSC Chinese Spring RNA-Seq data (Choulet *et al.*, 2014). Poor coverage within the *TaALN-2AL* genomic region was observed for two Australian bread wheat genotypes, RAC875 and Wyalkatchem (a pedigree of cv. Mace; Casartelli *et al.,* Chapter 4), and additional cultivars originating from other countries (Fig. S5-3). On the other hand, coverage for *TaALN-2DL* and *-2BL* homeologs appeared higher, although with some exceptions (e.g. Wyalkatchem) and thus, is in support of what was observed in the present study (Fig. 5-2; Fig. 5-3; Fig. 5-4).

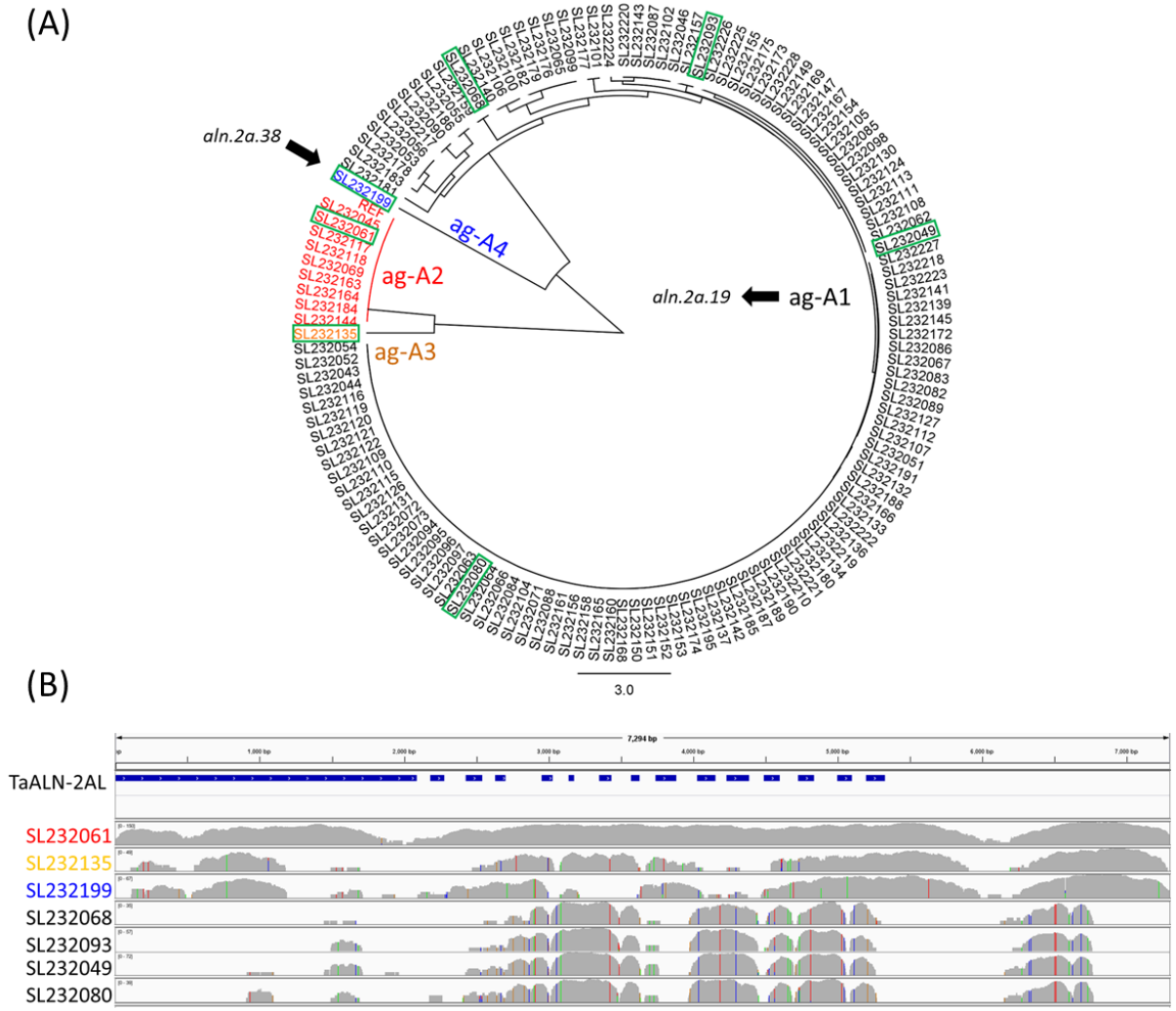


Fig. 5-2. Analysis of *TaALN-2AL* polymorphisms among 135 bread wheat accessions. (A) Polymorphism profiles were analysed for their average distance (PID; percentage of identity) and the outcome was visualised as a circular tree. Accessions were classified into major allele groups (ag): ag-A1 (black), ag-A2 (red), ag-A3 (orange) and ag-A4 (blue). Accessions presenting polymorphisms that putatively lead to a non-synonymous amino acid change of TaALN-2AL protein are indicated with an arrow and polymorphism ID is reported in italic (see Table 5-5). Note that all accessions belonging to ag-A1 present *aln.2a.19* polymorphism. Accessions highlighted with a green square are those reported in (B). Accessions details are reported in Table S5-4. The figure was prepared with the support of Geneious software. **(B)** Details of *TaALN-2AL* homeolog read mapping. *TaALN-2AL* 2kb promoter and exons are depicted in blue and correspond to reference genome of cv. Chinese Spring. Below, SL232061 corresponds to reads sequenced from Chinese Spring DNA samples, while the remaining were selected as representative of specific allele groups (ag-A) and the colour indicates the group number reported in (A). The figure was prepared with the support of IGV 2.4.1 software.

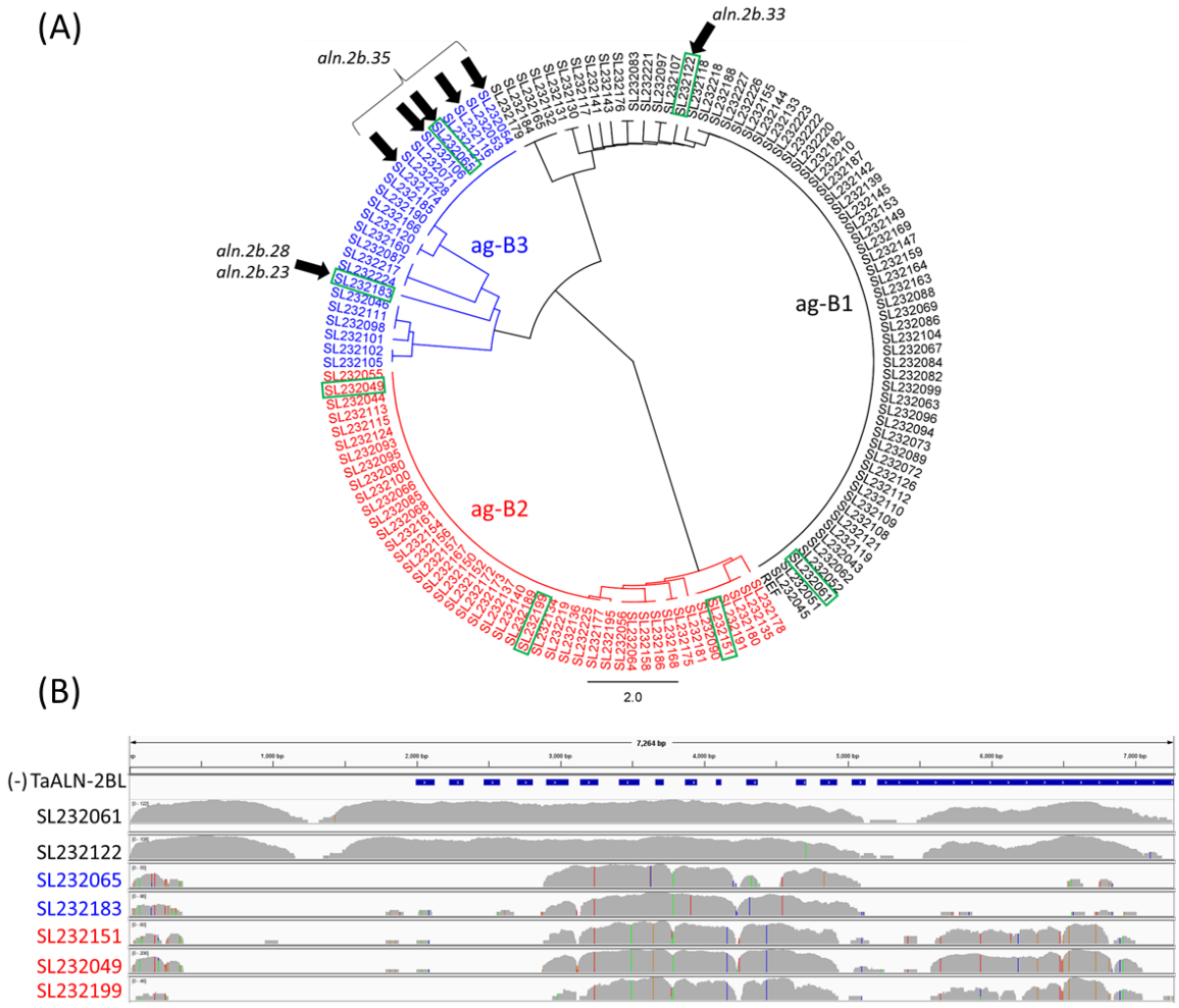


Fig. 5-3. Analysis of *TaALN-2BL* polymorphisms among 135 bread wheat accessions. See Fig. 5-4 for legend details. **(A)** Accessions were classified in major allele groups (ag): ag-B1 (black), ag-B2 (red), and ag-B3 (blue). **(B)** Note that *TaALN-2BL* coding sequence is located on the minus strand.

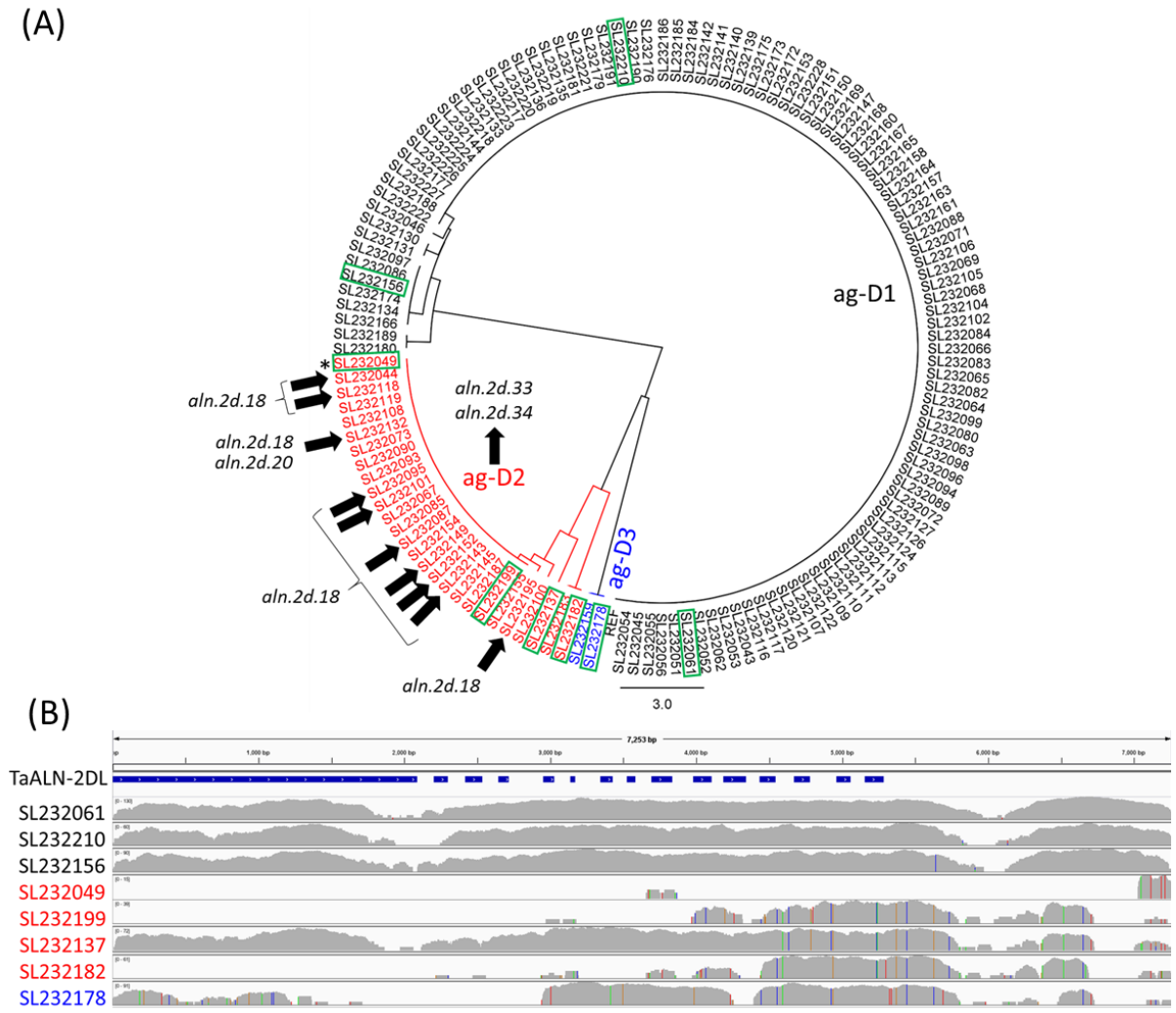


Fig. 5-4. Analysis of *TaALN-2DL* polymorphisms among 135 bread wheat accessions. See Fig. 5-4 for legend details. **(A)** Accessions were classified in major allele groups (ag): ag-D1 (black), ag-D2 (red), and ag-D3 (blue). Note that all accessions belonging to the ag-D2 have *aln.2d.33* and *aln.2d.34* polymorphism, except accession SL232049 (indicated with an asterisk).

Can the Chinese Spring reference genome capture the actual genetic diversity among TaALN homeologs?

Given the high plasticity of bread wheat genomes, events of homeolog deletion or disruption are expected to occur at a relatively fast pace (Dubcovsky and Dvorak, 2007). Interestingly, a small number (15) of accessions presented moderate to low reads coverage in all three *TaALN* homeologs (Fig. S5-4), raising the question as to whether *TaALN* and the capacity to catabolise allantoin, is completely absent in these genotypes. However, considering that purine catabolism is regarded as a housekeeping pathway involved in nutrient recycling in plants (Werner and Witte, 2011) and that the *ALN* gene is ubiquitous in most organisms (Kim *et al.*, 2009), it is expected that at least one functional *TaALN* homeolog is retained in these genotypes. For example, accession SL232049 (AMC 71), the only one sourced from Iraq (Table S5-2), had no reads mapped to the *TaALN-2DL* homeolog region (Fig. 5-4B), strongly suggesting gene absence, whilst low read coverage was observed for the *TaALN-2AL* and T*aALN-2BL* homeologs (Fig. 5-2B; Fig. 5-3B). The accession SL232199, a landrace named Mocho De Espiga Branca and the only one sourced from Portugal (Carvalho *et al.*, 2011), had a moderate coverage for the *TaALN-2AL* homeolog (Fig. 5-2B), whilst having a low coverage for *TaALN-2BL* and *TaALN-2DL* (Fig. 5-3B; Fig. 5-4B). Considering that the germplasms used in this study were sourced from different countries (Table S5-2), the results presented here suggest that the reference genome of Chinese Spring is not representative for the *TaALN* gene. Chinese Spring is thought to be a landrace originated in the Chinese province of Sichuan that is characterised by relatively low photosynthetic radiation and high humidity and temperatures during grain filling period (Liu *et al.*, 2017). Although this genotype has been widely used for cytogenetics studies since the beginning of the twentieth century, its value in breeding programs was minimal, mainly because it is not well adapted to grow in the main wheat growing areas around the world (Sears and Miller, 1985). Therefore, further studies focussed on bread wheat accessions acclimated to arid environments are required to identify more valuable *TaALN* haplotypes. Interestingly, the landrace Mocho De Espiga Branca is currently under investigation at the University of Adelaide for its tolerance to salt stress (Roy *et al*.,

unpublished). This is intriguing and indicates that assessing the *TaALN* gene (and allantoin accumulation) in relation to abiotic stress tolerance might be interesting to explore further.

Identification of putative functional polymorphisms in TaALN homeologs

In order to predict if a specific nucleotide substitution leads to an alteration of *TaALN* gene expression or enzymatic activity the identified polymorphisms were analysed in relation of their location within the *TaALN* gene model. In addition, to better assess the putative function of polymorphisms occurring in *TaALN* CDSs, ALN protein sequences from different plant species (bread wheat, Brachypodium and Arabidopsis), unicellular eukaryotes (*Chlamydomonas reinhardtii* and *Saccharomyces cerevisiae*) and unicellular prokaryotes (*Escherichia coli*) were aligned (Fig. 5-5). Overall, a large number of completely conserved residues were found in ALN protein sequences across all organisms (black colour) (Fig. 5-5). Yang and Han (2004), who first identified and characterised ALN in Arabidopsis and the leguminous tree *Robina pseudoacacia*, showed that ALN residues involved in metal binding are highly conserved across plants, animals, yeast (*S. cerevisiae*) and bacteria (*E. coli*), (Fig. 5-5; red asterisks). These putative metal binding sites are also conserved in bread wheat, Brachypodium and in the photosynthetic alga *C. reinhardtii* (Fig. 5-5; red asterisks).

Read alignment to the Chinese Spring reference genome allowed us to identify 182 putative polymorphisms across the three *TaALN* homeologs (Table S5-3). Overall, 52 polymorphisms occurred in the putative promoter region (2 kb upstream of ATG), 46 in introns, 47 in the putative 3'UTR (2 kb downstream the CDS stop codon) and 37 in exons, of which 10 putatively lead to a non-synonymous amino acid change (Table 5-3; Figure 5-5).

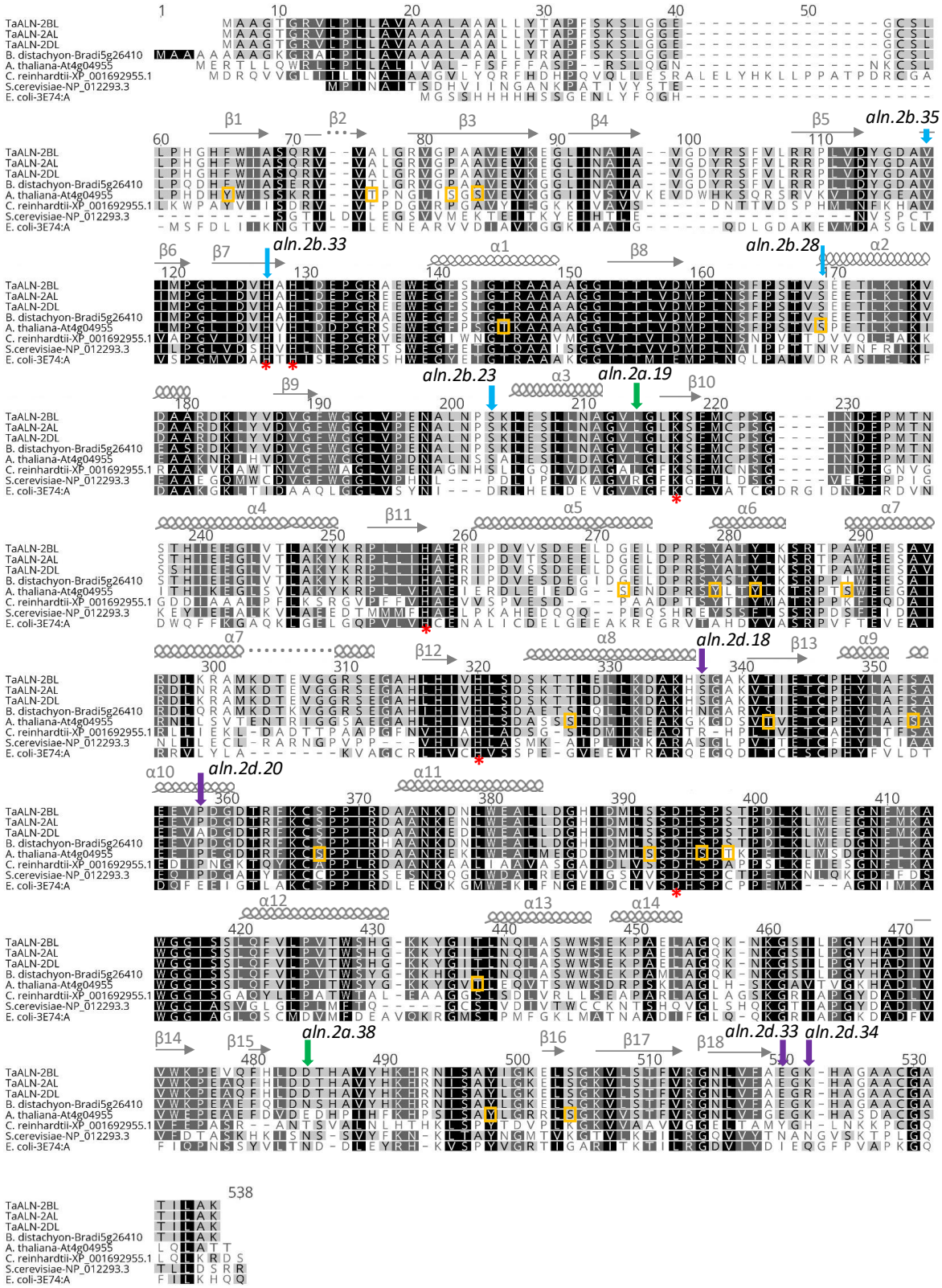

aln.2b.35
aln.2b.33
aln.2b.28
aln.2b.23
aln.2a.19
aln.2d.18
aln.2d.20
aln.2a.38
aln.2d.33
aln.2d.34
TaALN-2BL
TaALN-2AL
TaALN-2DL
B. distachyon-Bradi5g26410
A. thaliana-At4g04955
C. reinhardtii-XP_001692955.1
S.cerevisiae-NP_012293.3
E. coli-3E74:A

Fig. 5-5. Alignment of ALN amino acid sequences. ALN putative protein sequences of bread wheat (TaALN) were compared to *Brachypodium distachyon*, *Arabidopsis thaliana*, *Chlamydomonas reinhardtii*, *Saccharomyces cerev*isiae and *Escherichia coli*. Black indicates 100% similarity, dark grey 80-100% similarity, light grey 60-80% similarity and white less than 60% similarity by Blosum62 score matrix. Red asterisks denotes residues that are required for metal binding (Yang et al., 2004; Kim et al., 2009; Ho et al., 2011). Secondary protein structures, alpha-helix (α) and beta-sheet (β), are depicted in grey according to Kim et al. (2009). Orange squares indicate Arabidopsis phosphorylated sites predicted by PhosPhat 4.0 webtool (Durek et al., 2010; http://phosphat.uni-hohenheim.de/). Arrows indicate *TaALN* non-synonymous amino acid changes identified by sequence capture (see Table 5-3 for details).

a) Non-synonymous amino acid substitution.

Analysis of single nucleotide polymorphisms (SNP) putatively leading to amino acid changes in TaALN highlighted certain SNPs to be associated with specific allele groups or, conversely, to be unique (Table 5-3). For example, *aln.2a.19*, leading to substitution ALN-2AL^{Leu190Phe}, was detected in all accessions belonging to ag-A1 (Fig. 5-4A). Therefore, Phe^{190} appears as the most common residue rather than Leu^{190} as present in Chinese Spring. In *TaALN-2DL*, two SNPs in close proximity (*aln.2d.33* and *aln.2d.34*) were found leading to ALN-2DL^{Glu490Gln} and ALN-2DL^{Arg492Lys} substitutions and, interestingly, those were shared by the accessions belonging to ag-D2 (Fig. 5-3A; Table 5-3). On the other hand, five substitutions were unique (Table 5-3). A non-synonymous amino acid change (*aln.2b.33*) was identified in one of the residues involved in metal binding and was specific to an accession of ag-B1 (SL232122; Jupateco 73 from Mexico) (Fig. 5-5). This corresponded to a histidine to tyrosine substitution at position 103 within TaALN-2BL. In *E. coli*, site direct mutagenesis of EcALN^{His79} (corresponding to TaALN-2BL^{His103}) and insertion of alanine (EcALN^{Ala79}) caused the disruption of Mn^{2+}-activated EcALN specific activity (Ho *et al.*, 2011). Histidine and tyrosine are both aromatic amino acids with the former being positively charged (basic), whilst the latter is polar but uncharged. This raises the question whether the reported change is likely to affect the metal binding activity of this TaALN-2BL homeolog. *In silico* protein modelling and *in vitro* biochemical characterisation of TaALN-2DL^{His103Tyr} will be required to assess the extent of which the substitution alters the enzyme activity.

Another interesting unique non-synonymous change (*aln.2b.28*) was identified in *TaALN-2BL* at position 145 of accession SL232183 (AE.SQUARROSA 518) belonging to ag-B3 (Fig. 5-5). This SNP caused a putative serine to cysteine substitution. The residue appeared conserved among plant species and was predicted as a potential phosphorylation site by web tool PhoshPhat 4.0 (http://phosphat.uni-hohenheim.de/). Several other predicted phosphorylated residues were identified in AtALN within highly conserved positions (e.g. threonine^{145}, tyrosine^{282}, serine^{396}, with respect to the consensus sequence) and, therefore, might be interesting candidates for in-depth studies. Phosphorylation and other post-

translational modifications may play an important role in regulating TaALN activity under stress. The existence of such regulation possibly explains the inconsistencies in the timing of allantoin accumulation and *TaALN* down-regulation under stress as previously reported by Casartelli *et al.* (Chapter 4). Post-translational modification of ALN in plants have not been studied yet and information on other enzymes of the purine catabolism are scarce except on AAH. For example, Charlson *et al.* (2009), reported that soybean *GmAAH* expression was unchanged under drought and between genotypes despite the observed increase in ureide levels. Alamillo *et al.* (2010) also showed no alteration of *AAH* mRNA in common bean (*Phaseolus vulgaris*) under drought, whilst PvAAH activity was slightly reduced in nodules and roots, but not in shoots and leaves suggesting a possible post-translational regulation.

An alanine to proline substitution was observed in TaALN-2DL at position 330 (*aln.2d.20*) of accession SL232132 (LW336 from Morocco), which occurred in the putative alpha (α)-helix 10 (Fig. 5-5). Proline can alter the α-helix structure as its side chain cannot donate a hydrogen bond that is required to stabilise α-helices. Curiously, TaALN-2AL and TaALN-2BL homeologs and all other organisms, expect *E. coli*, have a conserved proline residue in that position, suggesting that the observed alanine to proline substitution is a reversion. Detailed analysis of accessions of the wheat D genome ancestor *Aegilops tauschii* may elucidate the occurrence of this substitution and determine whether this affects TaALN-2DL activity.

Wheat genome, given its hexaploidy, can handle a relatively higher rate of mutations than diploid genomes due to homeolog redundancy. Therefore, the 'rare' amino acids substitutions presented above (Table 5-3) can be expected, although they need to be investigated further. First of all, DNA amplification and sequencing of the *TaALN* homeolog from those specific accessions is required to confirm the mutation. Secondly, the screen of additional related lines (e.g. from the same country) would allow us to clarify if the muatation can be considered an allelic variation of a specific sub-group of wheat accessions.

Table 5-3. Details of SNPs identified in *TaALN* CDS that lead to a putative non-synonymous amino acid change.

Sequence	SNP name	Position (+)	Exon N	Amino acid position	Nucleotide substitution	Codon position	Substitution	Occurrence	allele group (ag) and accessions	Residue function (see Fig. 5-4)
TaALN-2AL	*aln.2a.19*	3424	Exon 7	190	C→T	I	Leu → Phe	119 (127)	ag-A1	
TaALN-2AL	*aln.2a.38*	5069	Exon 14	454	G→A	I	Asp →Asn	1 (10)	ag-A4: SL232199	
TaALN-2BL	*aln.2b.35*	2463	Exon 3	94	T→C	II	Val →Ala	5 (72)	ag-B3: SL232054, SL232116, SL232065, SL232106, SL232174	
TaALN-2BL	*aln.2b.33*	2587	Exon 4	103	C→T	I	His → Tyr	1 (109)	ag-B1: SL232122	Metal binding
TaALN-2BL	*aln.2b.28*	2980	Exon 5	145	C→G	II	Ser → Cys	1 (120)	ag-B3: SL232183	Predicted phosphorylation site
TaALN-2BL	*aln.2b.23*	3388	Exon 7	179	C→A	III	Ser → Arg	1 (134)	ag-B3: SL232183	
TaALN-2DL	*aln.2d.18*	4199	Exon 11	309	A→G	I	Ser → Gly	10 (117)	ag-D2	
TaALN-2DL	*aln.2d.20*	4262	Exon 11	330	G→C	I	Ala → Pro	1 (103)	ag-D2: SL232132	
TaALN-2DL	*aln.2d.33*	5236	Exon 15	490	G→C	I	Glu → Gln	26 (133)	ag-D2: same accessions as *aln.2d.34*	
TaALN-2DL	*aln.2d.34*	5243	Exon 15	492	G→A	II	Arg → Lys	26 (133)	ag-D2: same accessions as *aln.2d.33*	

In the 'Occurrence' column is reported the number of accessions harbouring the identified SNP, in brackets is reported the total number of accessions with reads mapped at that position.

b) Nucleotide changes in pALNphy motifs

SNPs and short INDELs occurring in *TaALN* cis-regulatory elements can potentially alter gene expression. This possibility was taken into consideration and polymorphisms were analysed in relation to the identified pALNphy motifs in the promoter regions of the *TaALN* genes (see above) as summarised in Table 5-4. Overall, five SNPs were identified, two in the upstream region of *TaALN-2AL* and three in *TaALN-2BL*. Interestingly, two of those (*aln.2a.6*, *aln.2b.36*) occurred in the pALNphy1 motif and at position 20, which corresponded to "N" of the signal sequence identified by PLACE (5'-CANNTG-3') associated with site MYCCONSENSUSAT/EBOXBNNAPA (Table 5-2). In *aln.2a.6*, guanine (G) is substituted by a cytosine (C), which is the most common nucleotide occurring at position 20 of pALNphy1 across all grass genomes analysed (Fig. S5-2). On the other hand, in *aln.2b.36*, G is substituted with adenine (A) representing a unique nucleotide for this position among grass genomes. Curiously, accession SL232140 (PBW343 from India) presented both *aln.2a.6* and *aln.2b.36* simultaneously. The remaining polymorphisms were identified in less conserved putative cis-elements (Table S5-1). In particular, *aln.2b.38* and *aln.2b.37*, were associated by PLACE to TBOXATGAPB and GT1CONSENSUS, respectively. These sequences are putatively associated with TF involved in light-regulated processes similar to other signal sequences identified as reported above (Table 5-2).

Table 5-4. Details of SNPs identified in *TaALN* promoter regions occurring in putative conserved pALNphy motifs identified with MEME web tool.

Sequence	SNP name	Position (+)	Nucleotide substitution	Putative conserved motif (MEME)	Site Name (PLACE)	Regulatory Element (PLACE)	Occurrence	Accession details
TaALN-2AL	*aln.2a.6*	1667	G→C	pALNphy1	MYCCONSENSUSAT/EBOXBNNAPA	CA**N**NTG	4 (9)	ag- A1: SL232140, SL232124, SL232220, SL232218
TaALN-2AL	*aln.2a.7*	1684	C→A	pALNphy24*	none	none	2 (7)	ag- A1: SL232124, SL232220
TaALN-2BL	*aln.2b.38*	5846	C→G	pALNphy11*	TBOXATGAPB	(-) ACTTT**G**	2 (99)	ag- B3: SL232105, SL232111
TaALN-2BL	*aln.2b.37*	5749	T→A	pALNphy9*	GT1CONSENSUS	GR**W**AAW	3 (95)	ag- B3: SL232101, SL232098, SL232111
TaALN-2BL	*aln.2b.36*	5646	C→T	pALNphy1	MYCCONSENSUSAT/EBOXBNNAPA	CA**N**NTG	4 (38)	ag- B2: SL232049, SL232191, SL232175, SL232140

Asterisk denotes the putative motif was assigned to the extended list reported in Table S5-3. Known cis-acting regulatory elements identified with PLACE are reported in column six (site name) and seven (regulatory element sequence). Red letter indicates the position where the detected SNP occurs in the signal sequence identified with PLACE database. In column eight is reported the number of accessions harbouring the identified SNP, in brackets is reported the total number of accessions with reads mapped at that position. The last column reports details of the *TaALN* haplotype groups and accession ID associated with the SNP.

Conclusions

A phylogenetic approach in combination with data mining for known plant cis-regulatory elements revealed the presence of a number of highly conserved motifs in grass *ALN* promoter regions that are possibly involved in regulating expression under abiotic stress and N restrictions. This evidence further proves that *TaALN* is induced or repressed in response to abiotic stresses. Applying this approach to genes that are co-regulated with *TaALN* might allow us to identify additional regulatory elements that are functionally related. Candidate elements may then be tested by promoter studies in *planta* under drought and N deficiency. These regulatory elements could be the object of forward and reverse genetic approaches for manipulating allantoin levels to examine the effects on wheat stress tolerance. Targeted sequence capture analysis revealed that great genetic diversity of *TaALN* homeologs among bread wheat genotypes may exist, or at least to significantly diverge from the reference cultivar Chinese Spring. In order to overcome the limitations imposed by the use of a reference sequence, *de novo* assembly of the captured sequences could be attempted and may provide more sequence information from accessions with low read coverage. Bread wheat genotypes presenting natural substitutions in predicted *TaALN* functional sites, such as metal binding sites or relevant cis-elements in the promoter, once experimentally validated, may be a valuable source for functional characterisation of *TaALN* under stress. Identification of *TaALN* haplotypes that present differential accumulation of allantoin under stress will be ultimately used to assess allantoin association to stress tolerance in wheat.

Chapter 5

Supplementary material

Table S5-1. Additional candidate motifs identified in the ALN promoter region of grass genomes with MEME web tool.

			Start position	
Name	**Hits**	**E-value**	**Average**	**SEM**
pALNphy1	10	4.60E-51	1659.80	9.44
pALNphy2	8	2.50E-41	1807.88	41.91
pALNphy3	7	9.40E-41	1875.14	11.14
pALNphy4	8	5.10E-37	1696.75	8.14
pALNphy5	9	2.20E-32	1469.67	59.87
pALNphy7	9	2.30E-29	1804.56	23.01
pALNphy8	10	9.80E-25	1919.60	4.80
pALNphy9	7	9.80E-25	1556.29	9.52
pALNphy11	6	2.20E-18	1455.33	8.83
pALNphy12	8	5.40E-17	1749.25	15.94
pALNphy14	6	3.60E-13	1624.17	8.99
pALNphy15	5	1.80E-11	1379.00	10.29
pALNphy19	9	2.20E-08	1290.22	80.97
pALNphy24	6	2.90E-04	1677.50	8.53

ALN promoter regions (2 Kb upstream of the translational start site ATG) of *Triticum aestivum*, *Aegilops tauschii*, *Triticum urartu*, *Hordeum vulgare*, *Brachypodium distachyon*, *Oryza sativa*, *Zea mays* and *Sorghum bicolor* were analysed with a phylogenetic approach using MEME v4.12.0 and conserved short regions were identified (pALNphy). Selection criteria were applied as follow: E-value ≤ 0.01, SEM of the motif start site position ≤100 and number of sequence "Hits" ≥5. Hits refers to number of orthologous promoter regions containing a specific motif (total orthologous sequence = 10).

Table S5-2. Details of bread wheat accession used for TaALN sequence capture. na, not available.

Accession ID	Name	Source
SL232043	117-var.12/564	PAKISTAN
SL232044	37627	EGYPT
SL232045	ACADEMIE DE PEKIN	CHINA
SL232046	Afghan 40	AFGHANISTAN
SL232049	AMC 71	IRAQ, Asia Minor expedition (1975)
SL232051	AUS29529	MEXICO, CIMMYT
SL232052	Bass	AUSTRALIA (QLD), QUEENSLAND WHEAT RESEARCH INSTITUTE
SL232053	BT2277	TUNISIA
SL232054	BT2281	TUNISIA
SL232055	BUCK ATLANTICO	ARGENTINA
SL232056	BUCKBUCK'S'	MEXICO, CIMMYT
SL232061	Chinese Spring	CHINA
SL232062	Chuanmai 18	CHINA
SL232063	cid183961sid5	MEXICO, CIMMYT
SL232064	cid394087sid30	MEXICO, CIMMYT
SL232065	cid428603sid31	MEXICO, CIMMYT
SL232066	cid473259sid69	MEXICO, CIMMYT
SL232067	COPPADRA	TURKEY
SL232068	CRUSADER	AUSTRALIA
SL232069	CS(Hope3B)	unknown
SL232071	Cul10	KAZAKHSTAN
SL232072	Cul3	KAZAKHSTAN
SL232073	Cul9	KAZAKHSTAN
SL232080	DRYSDALE	AUSTRALIA (VIC), CSIRO
SL232082	EMU'S'	MEXICO, CIMMYT
SL232083	ERADU	AUSTRALIA (WA)
SL232084	ESTANZUELA DORADO	URUGUAY
SL232085	EXCALIBUR	AUSTRALIA (SA), ROSEWORTHY AGRICULTURAL COLLEGE
SL232086	Ford	AUSTRALIA (SA), ROSEWORTHY AGRICULTURAL COLLEGE
SL232087	Free Gallipoli	AUSTRALIA (VIC), WERRIBIE
SL232088	FRONTANA 3671	BRAZIL
SL232089	G 72300	GREECE
SL232090	GAMENYA	AUSTRALIA (NSW), SYDNEY UNIVERSITY
SL232093	Gatcher	AUSTRALIA (NSW), SYDNEY UNIVERSITY
SL232094	GLADIUS	AUSTRALIA (SA)
SL232095	Glenwari	AUSTRALIA (NSW)
SL232096	Granarolo	ITALY
SL232097	H 11	AFGHANISTAN
SL232098	H 1328A	AFGHANISTAN
SL232099	H 1409	AFGHANISTAN
SL232100	H 1440	AFGHANISTAN
SL232101	H 1534	AFGHANISTAN
SL232102	H 1540	AFGHANISTAN
SL232104	H 1582	AFGHANISTAN
SL232105	H 1601	AFGHANISTAN
SL232106	H 1617	AFGHANISTAN
SL232107	H 1621	AFGHANISTAN
SL232108	H 1647	IRAN
SL232109	H 1681A	AFGHANISTAN
SL232110	H 1703	AFGHANISTAN
SL232111	H 339	AFGHANISTAN
SL232112	H 742A	ISRAEL
SL232113	HARTOG	AUSTRALIA (QLD), QUEENSLAND WHEAT RESEARCH INSTITUTE
SL232115	Heron	AUSTRALIA (NSW), WAGGA WAGGA
SL232116	HTWYT 012	MEXICO, CIMMYT
SL232117	INDIA 231	INDIA
SL232118	INDIA 322	INDIA
SL232119	INDIA 323	INDIA
SL232120	INIA 66	MEXICO, CIMMYT
SL232121	JANZ	AUSTRALIA (QLD)
SL232122	JUPATECO 73	MEXICO, CIMMYT
SL232124	KENNEDY	AUSTRALIA (QLD)
SL232126	KRICHAUFF	AUSTRALIA (SA), WAITE AGRICULTURE RESEARCH INSTITUTE
SL232127	KUKRI	AUSTRALIA (SA), ROSEWORTHY AGRICULTURAL COLLEGE
SL232130	LW334	TURKEY
SL232131	LW335	TURKEY
SL232132	LW336	MOROCCO
SL232133	LW337	ETHIOPIA
SL232134	ODESSA EXP.STA.17413	SYRIA
SL232135	ODESSA EXP.STA.19565	ETHIOPIA
SL232136	OPATA85	MEXICO, CIMMYT
SL232137	Palestinskaya	PALESTINE
SL232139	PAMUKALE	TURKEY
SL232140	PBW343	INDIA
SL232141	PERSIA 111	IRAN
SL232142	PERSIA 21	IRAN
SL232143	PERSIA 29	IRAN

SL232144	LW349	TAJIKISTAN
SL232145	PERSIA 7	IRAN
SL232147	Pitic 62	MEXICO, CIMMYT
SL232149	Purple Straw	UNITED STATES
SL232150	Quadrat	AUSTRALIA (VIC), WERRIBIE
SL232151	RAC875	AUSTRALIA (SA)
SL232152	REEVES	AUSTRALIA (WA)
SL232153	ROELFS F2007	MEXICO, CIMMYT
SL232154	S975-A4-A1	ZIMBABWE
SL232155	M708	ISRAEL
SL232156	SCOUT	AUSTRALIA (WA)
SL232157	SIETE CERROS	MEXICO, CIMMYT
SL232158	SOKOLL	MEXICO, CIMMYT
SL232159	SYNTHETIC W7984	MEXICO, CIMMYT
SL232160	Tacupeto 2001	MEXICO, CIMMYT
SL232161	TAFERSTAT	ALGERIA
SL232163	Thori 212-var.8/1	PAKISTAN
SL232164	U-Man-Syao-Mai	CHINA
SL232165	VOROBEY	MEXICO, CIMMYT
SL232166	MACE	AUSTRALIA (SA)
SL232167	WAAGAN	AUSTRALIA (VIC)
SL232168	WESTONIA	AUSTRALIA (WA)
SL232169	WILGOYNE	AUSTRALIA (WA)/CIMMYT - YAQUI VALLEY
SL232172	XIAOYAN 54	CHINA
SL232173	YECORA 70	CHILE
SL232174	YITPI	AUSTRALIA (SA), WAITE AGRICULTURE RESEARCH INSTITUTE
SL232175	YOUNG	AUSTRALIA (VIC)
SL232176	AE.SQUARROSA (309)	MEXICO, CIMMYT
SL232177	MEIRA	SPAIN
SL232178	AE.SQUARROSA (314)	MEXICO, CIMMYT
SL232179	AE.SQUARROSA (333)	MEXICO, CIMMYT
SL232180	AE.SQUARROSA (409)	MEXICO, CIMMYT
SL232181	AE.SQUARROSA (502)	MEXICO, CIMMYT
SL232182	AE.SQUARROSA (507)	MEXICO, CIMMYT
SL232183	AE.SQUARROSA (518)	MEXICO, CIMMYT
SL232184	AE.SQUARROSA (658)	MEXICO, CIMMYT
SL232185	cid460354sid132	MEXICO, CIMMYT
SL232186	W 98	PAKISTAN
SL232187	India 227	PAKISTAN
SL232188	Menflo	CHILE
SL232189	EGA Gregory	na
SL232190	Kauz	na
SL232191	RT812	AUSTRALIA
SL232195	TA-10441	na
SL232199	MOCHO DE ESPIGA	PORTUGAL
SL232210	N46	ISRAEL
SL232217	na	na
SL232218	na	na
SL232219	na	na
SL232220	na	na
SL232221	N67M2	na
SL232222	na	na
SL232223	na	na
SL232224	na	na
SL232225	na	na
SL232226	na	na
SL232227	na	na
SL232228	na	na

Table S5-3. List of polymorphisms identified in TaALN genomic regions.

TaALN-2AL						
Polymorphism ID	**Gene**	**Position**	**Reference**	**Type**	**Codon**	**Syn/Non-syn**
aln.2a.1	Promoter	776	C	C→A		
aln.2a.2	Promoter	886	T	deletion		
aln.2a.3	Promoter	1484	A	deletion		
aln.2a.4	Promoter	1543	G	G→A		
aln.2a.5	Promoter	1576	T	deletion		
aln.2a.6	Promoter	1667	G	G→C		
aln.2a.7	Promoter	1684	C	C→A		
aln.2a.8	Exon 4	2672	G	G→C	III	Syn
aln.2a.9	Intron 4	2713	C	C→A		
aln.2a.10	Intron 4	2744	C	c/+A		
aln.2a.11	Intron 4	2755	C	C→G		
aln.2a.12	Intron 4	2778	C	C→T		
aln.2a.13	Intron 4	2831	T	T→G		
aln.2a.14	Intron 4	2865	G	G→C		
aln.2a.15	Intron 4	2900	G	G→A		
aln.2a.16	Intron 4	2907	C	C→T		
aln.2a.17	Intron 5	3058	G	G→C		
aln.2a.18	Intron 5	3087	T	T→A		
aln.2a.19	Exon 7	3424	C	C→T	I	Leu→Phe
aln.2a.20	Intron 8	3699	G	G→A		
aln.2a.21	Intron 8	3732	C	C→A		
aln.2a.22	Exon 9	3811	A	A→G	III	Syn
aln.2a.23		3982	C			
aln.2a.24	Exon 10	4039	G	G→C	III	Syn
aln.2a.25		4187	C			
aln.2a.26	Exon 11	4298	T	T→C	III	Syn
aln.2a.27	Exon 12	4528	G	het (G/C)	III	Syn
aln.2a.28	Exon 12	4564	C	C→T; T/C	III	Syn
aln.2a.29	Intron 12	4657	T	T→A		
aln.2a.30	Intron 12	4676	G	G→A		
aln.2a.31	intron 12	4697	C	C-T		
aln.2a.32	Intron 12	4713	A	A→G		
aln.2a.33	Exon 13	4732	C	het C/G	III	Syn
aln.2a.34	Exon 13	4738	T	T→A; T/A	III	Syn
aln.2a.35	Exon 13	4813	C	C→T	III	Syn
aln.2a.36	Exon 14	5032	C	C→T	III	Syn
aln.2a.37	Exon 14	5047	T	T→C	III	Syn
aln.2a.38	Exon 14	5069	G	G→A	I	Asp→Asn
aln.2a.39	Intron 14	5114	A	A→C		
aln.2a.40	Intron 14	5117	A	A→C		
aln.2a.41	Exon 15	5207	A	A→G	III	Syn
aln.2a.42	3'UTR	5532	G	G/+TAC		
aln.2a.43	3'UTR	5631	C	C→T		
aln.2a.44	3'UTR	5766	G	G/+T		
aln.2a.45	3'UTR	6325	A	A→C		
aln.2a.46	3'UTR	6337	G	G→C		
aln.2a.47	3'UTR	6506	C	C→T		
aln.2a.48	3'UTR	6512	C	C-T		
aln.2a.49	3'UTR	6575	C	del		
aln.2a.50	3'UTR	6600	G	G→A		
aln.2a.51	3'UTR	6601	T	T→G		
aln.2a.52	3'UTR	6626	T	T→C		
aln.2a.53	3'UTR	6685	G	G→C		
aln.2a.54	3'UTR	6732	G	G→A		
aln.2a.55	3'UTR	7225	G	G→A		
TaALN-2BL (CDS on negative strand)						
	Gene	**Position**	**Reference**	**Type**	**Codon**	**Syn/Non-syn**
aln.2b.1	3'UTR	74	T	T→A		
aln.2b.2	3'UTR	92	T	het T/A		
aln.2b.3	3'UTR	179	C	C→T		
aln.2b.4	3'UTR	206	A	het A/C		
aln.2b.5	3'UTR	254	G	het G/T		
aln.2b.6	3'UTR	259	G	het G/T		
aln.2b.7	3'UTR	275	A	A→G		
aln.2b.8	3'UTR	299	G	G/A		
aln.2b.9	3'UTR	355	A	het G/A		
aln.2b.10	3'UTR	358	T	het A/T		
aln.2b.11	Intron 14	2161	T	T→C		
aln.2b.12	Exon 11	2982	T	T→C	III	Syn
aln.2b.13	Exon 10	3238	C	C→T	III	Syn
aln.2b.14	Exon 9	3496	G	G→A	III	Syn
aln.2b.15	Intron 8	3625	C	C→G		
aln.2b.16	Intron 8	3630	T	T→C		
aln.2b.17	Intron 8	3648	A	A→G		
aln.2b.18	Intron 7	3771	G	inser +T		
aln.2b.19	Intron 7	3774	T	Ins +TTA		
aln.2b.20	Intron 7	3777	T	het T/A		

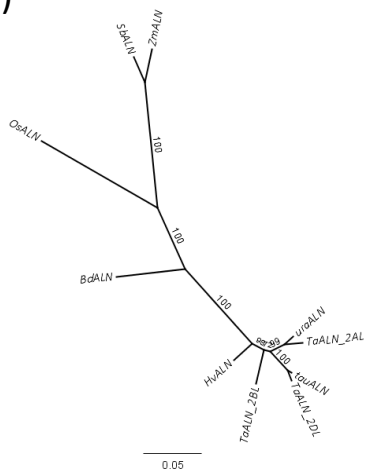
aln.2b.21	Intron 7	3780	A	ns		
aln.2b.22	Intron 7	3786	G	G→A		
aln.2b.23	Exon 7	3910	G	G→t	III	Ser→Arg
aln.2b.24	Intron 5	4160	T	T→C		
aln.2b.25	Intron 5	4161	G	G→A		
aln.2b.26	Intron 5	4203	T	del		
aln.2b.27	Intron 5	4246	C	C→T		
aln.2b.28	Exon 5	4318	G	G→C	II	Ser→Cys
aln.2b.29	Exon 5	4329	C	C→A	III	Syn
aln.2b.30	Exon 5	4362	T	T→A	III	Syn
aln.2b.31	Intron 4	4436	T	T→C		
aln.2b.32	Intron 4	4544	G	G→T		
aln.2b.33	Exon 4	4711	G	G→A	I	His→Tyr
aln.2b.34	Intron 3	4761	G	G→A		
aln.2b.35	Exon 3	4835	A	A→G	II	Val→Ala
aln.2b.36	Promoter	5646	C	C→T		
aln.2b.37	Promoter	5749	T	T→A		
aln.2b.38	Promoter	5846	G	G→C		
aln.2b.39	Promoter	5854	C	ns		
aln.2b.40	Promoter	5858	C	ns		
aln.2b.41	Promoter	5891	C	C→G		
aln.2b.42	Promoter	5903	T	T→A		
aln.2b.43	Promoter	5924	G	G→A		
aln.2b.44	Promoter	5925	C	C→T	prom	
aln.2b.45	Promoter	5979	T	del		
aln.2b.46	Promoter	5982	G	G→A		
aln.2b.47	Promoter	6034	T	T→C		
aln.2b.48	Promoter	6056	C	ns		
aln.2b.49	Promoter	6087	G	G→A		
aln.2b.50	Promoter	6186	T	T→C		
aln.2b.51	Promoter	6219	G	G→A		
aln.2b.52	Promoter	6303	G	G→A		
aln.2b.53	Promoter	6320	A	A→G		
aln.2b.54	Promoter	6369	C	C→T		
aln.2b.55	Promoter	6371	G	G→A		
aln.2b.56	Promoter	6416	C	C→T		
aln.2b.57	Promoter	6424	C	C→T		
aln.2b.58	Promoter	6476	C	C→T		
aln.2b.59	Promoter	6490	C	C→T		
aln.2b.60	Promoter	6497	A	A→G		
aln.2b.61	Promoter	6542	A	A→G		
aln.2b.62	Promoter	6572	A	A→T		
aln.2b.63	Promoter	6726	A	A→G		
aln.2b.64	Promoter	6798	A	A→G		
aln.2b.65	Promoter	6819	C	C→G		
aln.2b.66	Promoter	6842	T	del		
aln.2b.67	Promoter	6843	T	del		
aln.2b.68	Promoter	6895	T	T→C		
aln.2b.69	Promoter	6916	G	G→A		
aln.2b.70	Promoter	7107	T	T→C		
			TaALN-2DL			
	Gene	**Position**	**Reference**	**Type**	**Codon**	**Syn/Non-syn**
aln.2d.1	Promoter	83	G	del		
aln.2d.2	Promoter	84	T	del		
aln.2d.3	Promoter	187	G	G→A		
aln.2d.4	Promoter	216	A	A→G		
aln.2d.5	Promoter	314	C	/+A		
aln.2d.6	Promoter	392	T	T→C		
aln.2d.7	Promoter	939	A	A→G		
aln.2d.8	Promoter	1028	G	G→A		
aln.2d.9	Promoter	1096	T	T→C		
aln.2d.10	Promoter	1107	T	T→C		
aln.2d.11	Exon 5	3006	C	C→T	III	Syn
aln.2d.12	Exon 7	3415	G	G→A	III	Syn
aln.2d.13	Intron 7	3498	A	A→G		
aln.2d.14	Exon 10	3986	A	A→G	III	Syn
aln.2d.15	Exon 10	3995	T	T→C	III	Syn
aln.2d.16	Exon 10	4064	T	ns		
aln.2d.17	Exon 10	4067	T	T→C	III	Syn
aln.2d.18	Exon 11	4199	A	A→G	I	Ser→Gly
aln.2d.19	Exon 11	4261	C	C→T	III	Syn
aln.2d.20	Exon 11	4262	G	G→C	I	Ala→Pro
aln.2d.21	Exon 12	4446	T	T→C	III	Syn
aln.2d.22	Exon 12	4470	A	A→G	III	Syn
aln.2d.23	Exon 12	4476	A	A→G	III	Syn
aln.2d.24	Intron 12	4555	G	G→C		
aln.2d.25	Intron 12	4594	G	G→A		
aln.2d.26	Intron 12	4635	T	T→C		
aln.2d.27	Intron 12	4649	G	/+T		
aln.2d.28	Intron 13	4786	A	A→G		
aln.2d.29	Intron 13	4802	A	A→T		
aln.2d.30	Intron 13	4832	C	/+T		

aln.2d.31	Intron 13	4926	T	T→C		
aln.2d.32	Intron 13	4941	C	C→G		
aln.2d.33	Exon 15	5236	G	G→C	I	Glu→Gln
aln.2d.34	Exon 15	5243	G	G→A	II	Arg→Lys
aln.2d.35	3'UTR	5301	C	C→T		
aln.2d.36	3'UTR	5326	G	G→T		
aln.2d.37	3'UTR	5339	A	A→T		
aln.2d.38	3'UTR	5374	A	A→G		
aln.2d.39	3'UTR	5398	A	/+C		
aln.2d.40	3'UTR	5445	A	A→C		
aln.2d.41	3'UTR	5628	T	T→G/C		
aln.2d.42	3'UTR	5644	G	G→C		
aln.2d.43	3'UTR	5695	G	G→T		
aln.2d.44	3'UTR	5734	A	A→C		
aln.2d.45	3'UTR	6376	G	G→A		
aln.2d.46	3'UTR	6518	C	C→A		
aln.2d.47	3'UTR	6519	G	G→T		
aln.2d.48	3'UTR	6654	T	T→C		
aln.2d.49	3'UTR	6671	G	G→A		
aln.2d.50	3'UTR	6709	G	G→T		
aln.2d.51	3'UTR	7048	G	G→A		
aln.2d.52	3'UTR	7105	C	C→A		
aln.2d.53	3'UTR	7117	C	C→T		
aln.2d.54	3'UTR	7186	C	C/T; C→T		
aln.2d.55	3'UTR	7192	C	ns		
aln.2d.56	3'UTR	7213	C	C/T; C→T		
aln.2d.57	3'UTR	7264	C	C/T		

(A)

	TaALN-2AL	TaALN-2BL	TaALN-2DL	uraALN	tauALN	BdALN	HvALN	OsALN	ZmALN	SbALN
TaALN-2AL	100									
TaALN-2BL	95	100								
TaALN-2DL	95	94.5	100							
uraALN	97.6	94.9	95.4	100						
tauALN	94.9	95.1	98.7	95.8	100					
BdALN	87.5	87.4	87.1	87.1	87.3	100				
HvALN	94.5	94.6	93.9	94.5	94.3	87.8	100			
OsALN	81.8	81.3	81.5	81.5	81.8	84.6	82.3	100		
ZmALN	81.5	81.7	81.7	81.6	81.9	82.8	82.2	83.4	100	
SbALN	81.7	82	81.7	81.9	82	83.2	81.9	83.6	95.2	100

(B)



(C)

	pTaALN-2AL	pHvALN	pTaALN-2BL	puraALN	pTaALN-2DL	ptauALN	pBdALN	pSbALN	pOsALN	pZmALN
pTaALN-2AL	100									
pHvALN	46.3	100								
pTaALN-2BL	39.8	37	100							
puraALN	46.3	40.9	56.7	100						
pTaALN-2DL	43.1	37.5	43	42.3	100					
ptauALN	44.4	40.9	47.2	47.3	84.5	100				
pBdALN	31.1	31.3	30.9	33.5	28.8	33	100			
pSbALN	33.4	30.8	29.4	33.9	30.9	34.9	32.1	100		
pOsALN	30.1	27.4	28.5	32.2	28	30	29.5	33	100	
pZmALN	29.5	28.6	27	30.2	28.5	31.2	29.4	36.7	29.5	100

(D)

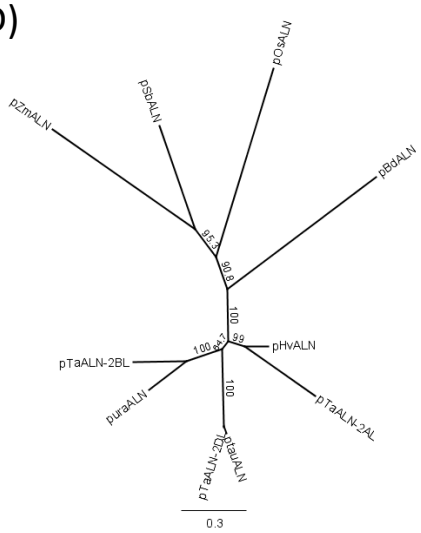


Fig. S5-1. Phylogenetic relationship of *ALN* CDS and 2kb promoter region among grasses genomes. (A) *ALN* CDS identity matrix **(B)** *ALN* CDS phylogenetic tree; **(C)** *ALN* promoter identity matrix; **(D)** *ALN* promoter phylogenetic tree. Nucleotide alignment was performed with Geneious 10.0.2 software. Identity matrix was constructed with ClustalW Alignment tool with default parameters (gap open cost:15; gap extend cost: 6.66) and trees were generated with Geneious Tree Builder tool with default parameters (genetic distance model: HKY; tree build method: Neighbor-Joining; number of replicates: 1000). Ta (*Triticum aestivum); ura (Triticum urartu*); tau (*Aegilops tauschii*); Bd (*Brachypodium distachyon*); Hv (*Hordeum vulgare*); Os (*Oryza sativa*); Zm (*Zea mays*); Sb (*Sorghum bicolor*).

pALNphy1

	Name	Strand	Start	*p*-value	Sites
6.	pHvALN	+	1675	4.84e-19	agattccaaa TCCCCGTGGGCCCCGGCCACTTGGAGGGAT ccagccagcg
5.	pura_ALN	+	1662	4.84e-19	AGATTCCAAA TCCCCGTGGGCCCCGGCCACTTGGAGGGAT CCAGCCAGCA
4.	ptau_ALN	+	1612	4.84e-19	AGATTCCAAA TCCCCGTGGGCCCCGGCCACTTGGAGGGAT CCAGCCAGCC
3.	pTaALN-2DL_NRgene_2bk	+	1647	4.84e-19	AGATTCCAAA TCCCCGTGGGCCCCGGCCACTTGGAGGGAT CCAGCCAGCC
2.	pTaALN-2BL_NRgene_2Kb	+	1633	1.38e-17	AGATTCCAAA TCCCCGTGGGCCCCGGCCAGTTGGAGGGAT CCAGCCAGCG
1.	pTaALN-2AL_Nrgene	+	1648	1.38e-17	AGATTCCAAA TCCCCGTGGGCCCCGGCCAGTTGGAGGGAT CCAGCCCGCA
9.	pSb_ALN	+	1635	1.91e-14	TCCCCAGAGA TCCCCAAAGGTTCAAGCCACTTGGATGGCG ACCCCACATG
7.	pBd_ALN	+	1702	1.36e-13	ATAAAAAGAT TCCCCGAAATTCCAGGCCACTTGGATCCAT TCGGTTCCCC
10.	pZm_ALN	+	1693	2.31e-13	CCCCAGAGAT TCCCAAAAGGTTTAAGCCACTTGGATGGCG ACCGCACACA
8.	pOs_ALN	+	1691	3.89e-13	GCAAGAGATC CCTCCAAAGGTTCCAGCCACTTGGATGCCC TCACACAACG

pALNphy5

	Name	Strand	Start	*p*-value	Sites
4.	ptau_ALN	+	1487	1.39e-18	GTCCCGGCAA GGGGCAACCAATGAAGAGCCGTCACATGAG CTCCACCCGC
3.	pTaALN-2DL_NRgene_2bk	+	1522	1.39e-18	TACCGGCAAA GGGGCAACCAATGAAGAGCCGTCACATGAG CTCCACCCGC
5.	pura_ALN	+	1528	4.09e-17	GTCCCGGCAA GGGGCAACCAATGAAGAGCCGTCACACGAC ACCCACCCGC
1.	pTaALN-2AL_Nrgene	+	1513	4.09e-17	GTCCCGGCAA GGGGCAACCAATGAAGAGCCGTCACACGAC GCCCACCCGC
6.	pHvALN	+	1549	4.41e-17	gtgccgggaa GGGGCAGCCAATGAAGGGCCGTCACATGAC ctccgcccgc
2.	pTaALN-2BL_NRgene_2Kb	+	1503	3.65e-15	TGCCGGGCAA GGGGCGGCCAATGAAGGGCCGTCACATGAG CTTTTTTTGG
9.	pSb_ALN	+	996	1.30e-12	GTTGGGAGCG GGAGCAGCCAATGGCGACCCTCCGAAATCG AAAGCCGCGT
10.	pZm_ALN	+	1565	2.27e-12	GTTGGGAGCG GGAGCAGCCAATGGAAACCCTCCGAAATCG AGAGCCGCGT
7.	pBd_ALN	+	1564	8.78e-11	CTTCCCCACG GCCCAGCCCAATCGCGCGCGCCACGTGCG CGTCACGCCG

pALNphy7

	Name	Strand	Start	*p*-value	Sites
4.	ptau_ALN	+	1771	3.80e-19	CTCGCTGGTC CAGTGCGCCGCGTCCACGGCGCGGTGCCCC AGGCGGGTGA
2.	pTaALN-2BL_NRgene_2Kb	+	1780	3.80e-19	CTCGCTGGAC CAGTGCGCCGCGTCCACGGCGCGGTGCCCC CCCCAGGCGG
3.	pTaALN-2DL_NRgene_2bk	+	1805	3.38e-17	CTCGCAGGTC CAGTGCACCGCGTGCACGGAGCGGTGCCCC TGGCGGGTGA
1.	pTaALN-2AL_Nrgene	+	1803	3.71e-16	AAACAGGGTC CAGTGCACCGAGCCCACGGCGCGGTGCCCG AGGCGGGCGA
6.	pHvALN	+	1829	8.66e-14	ctcccctgtc CAGTGCGCCGGGTTCGCGGAGCGGTGCACC gaaccccgcg
5.	pura_ALN	+	1805	1.40e-13	AATCTGCGGT GCATGCGCCGCGTCCACGGCGCGGCGCCCG AGGCGGGCGA
9.	pSb_ALN	+	1709	8.99e-12	GCCCACAGGG CCGCGCACCGCGCGCCCGGTCCAGTGGAAC ACGCCACGGT
7.	pBd_ALN	+	1774	1.49e-11	CTGCCCCAGC CACTCCCCCGAGAGCACCGTCCAGTGCGCC TTCTTCACTG
10.	pZm_ALN	-	1965	3.87e-11	GGTTTA CCCTGCCCTGCGCCGCCGCCGCCGTTCCTC CCTTGTCCTG

pALNphy8

	Name	Strand	Start	*p*-value	Sites
3.	pTaALN-2DL_NRgene_2bk	+	1927	4.84e-19	CCGAGGCGCA CGCACGCAGACGCAGACGGGGACGGGACTC TCGCCGCTGC
4.	ptau_ALN	+	1901	2.87e-18	CACGCACGCA CGCACGCAGACGCGGACGGGGACGGGACTC TCGCCGCTGC
2.	pTaALN-2BL_NRgene_2Kb	+	1906	2.87e-18	CCGAGGCGCA CGCACGCAGACGCGGACGGGGACGGGACTC TCGCCGCTGC
1.	pTaALN-2AL_Nrgene	+	1929	9.73e-14	GGCGCACGCA CGCACGCAGACGCAGACGCAGACGACAGGA CTCTCGCCAT
7.	pBd_ALN	+	1928	2.15e-13	AGCCAGTGGA GGAGGGCGCACGCAGACGGGGACGGGACTC GGACGCCAAC
10.	pZm_ALN	+	1889	5.62e-12	ACTGAGCGCG GGATCCCACGGGCAGAAAGGGACGCGACTC GTCAGGGCGA
6.	pHvALN	+	1934	1.77e-11	accgaggcga CGCACGCAGACGCGCCGGACAGGACTCTC gccatcagga
5.	pura_ALN	+	1929	4.52e-11	GAGGCGCACG CACGCACGCACGCAGACGCAGACGACAGGA CTCTCGCCAT
8.	pOs_ALN	+	1926	1.30e-10	CGAGGAAGGA AGGAAGCAGACGCAGAAAAGGACAGGAGGT CTCGTCACGA
9.	pSb_ALN	+	1927	1.35e-10	ATCCCACTTC ACTTCCCACGGGCAGAAAGGGACGCAACTC GACTGCGAGG

pALNphy19

	Name	Strand	Start	*p*-value	Sites
4.	ptau_ALN	+	1176	3.20e-16	AGTTTTTTCG TGAAGAGACGGATTAACCATTTGGGTTGG GTGAATGTGG
1.	pTaALN-2AL_Nrgene	+	1419	8.94e-15	AAATTTTTCT TGAAGAAATGGATTAAGCATTTGGGTTGC GTGAATGTGG
5.	pura_ALN	+	1435	1.07e-14	AAATTTTTCT TGAAGAAATGGATTAATCATTTGGGTTGC GTGAATGTGG
3.	pTaALN-2DL_NRgene_2bk	+	1427	1.31e-13	ATTTCGTAAA GAAATGGATTATTTAACCATTTGGGTTGG GTGAATGTGG
2.	pTaALN-2BL_NRgene_2Kb	+	1421	5.94e-12	TTTTAGCAAG GAAATGGATTATTTAGCCAGTTGGGTTGG GCCAAAGTGG
6.	pHvALN	+	1459	3.66e-11	tagcttttcc GTGAAAAATGGATTAACCATTTTCGGTCG ggtgaacgtg
9.	pSb_ALN	+	871	5.29e-10	ACAGAATGAT TGGATGGAGGGAGTAGCGTTTTGGTCTGG TGGGAAAAAG
7.	pBd_ALN	-	1490	3.31e-9	CACTCGAACG GAACGGGACGGATCGGGCGTTTTGTGGGG TAGCAGGCGC
8.	pOs_ALN	+	914	3.86e-9	TTCAAAAAAT TTAAGTAATTGTTAAATCTTTTTCTATCA TTTGATTTAT

Fig. S5-2. Details of the top motif identified using MEME web tool. ALN 2kb promoter in orthologous sequences among grass genomes was analysed with MEME v4.12.0. Conserved short regions (pALNphy) were identified using the following cut-off criteria: E-value ≤ 0.01, SEM of the motif start site position ≤100 and number of sequence "Hits" ≥9. Hits refers to number of orthologous promoter regions containing a specific motif (total orthologous sequence = 10). Ta (*Triticum aestivum); ura (Triticum urartu*); tau (*Aegilops tauschii*); Bd (*Brachypodium distachyon*); Hv (*Hordeum vulgare*); Os (*Oryza sativa*); Zm (*Zea mays*); Sb (*Sorghum bicolor*).

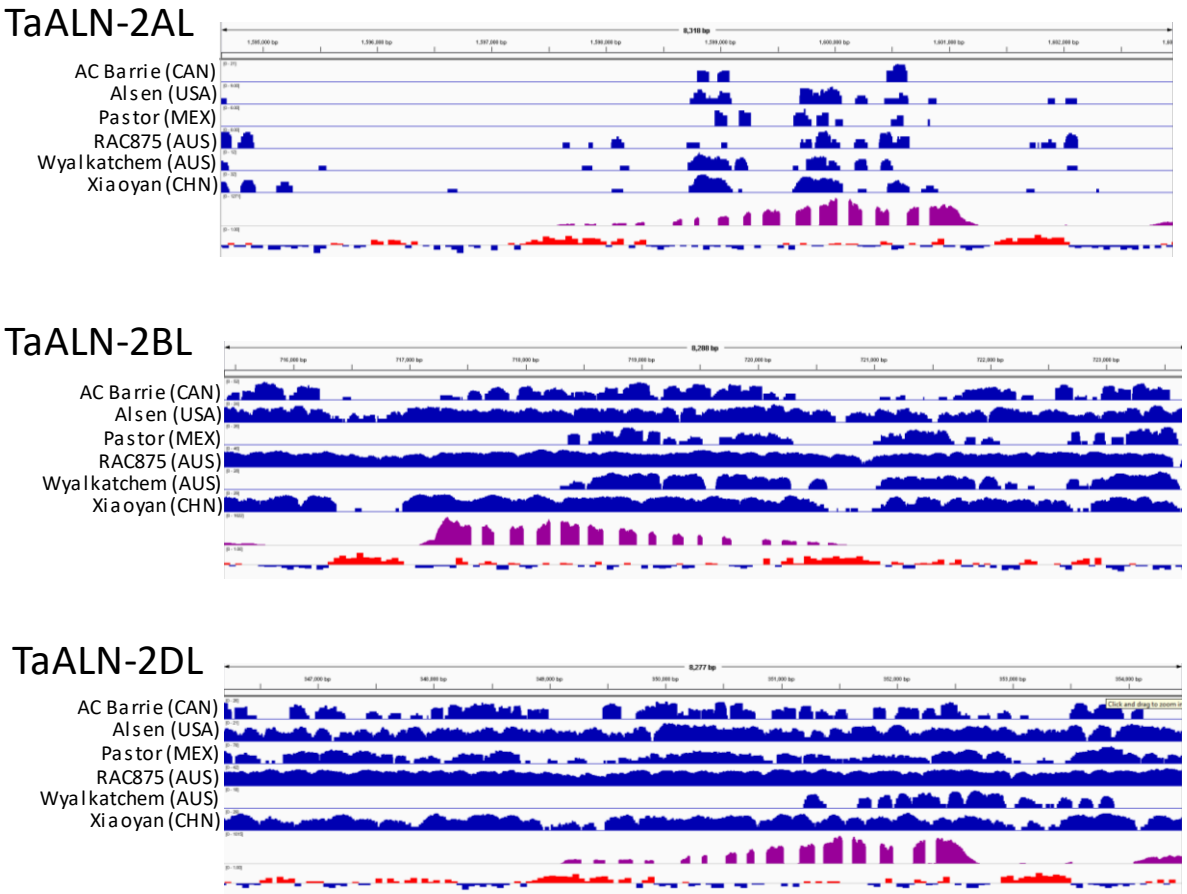


Fig. S5-3. Comparison of *TaALN* genes among different bread wheat genotypes. Whole Genome Shotgun (WGS) sequencing data of bread wheat accessions sourced from different countries (BPA, Bioplatform Australia; http://www.bioplatforms.com/wheat-sequencing/) were mapped against IWGSC Chinese Spring WGA V0.3 scaffolds. The genomic regions harbouring *TaALN* genes are reported here and the coverage of each accession in relation to the reference (Chinese Spring) is depicted in blue. IWGSC tissue series RNA-Seq dataset (https://wheat-urgi.versailles.inra.fr/Seq-Repository/Expression) mapped against IWGSC Chinese Spring WGA V0.3 scaffolds is depicted in purple. GC content (%) over 50 bp window is depicted at bottom of each panel. Note that *TaALN-2BL* CDS is located on the negative strand. The figure was drawn with the support of DAWN software (Baumann et al., unpublished, University of Adelaide).

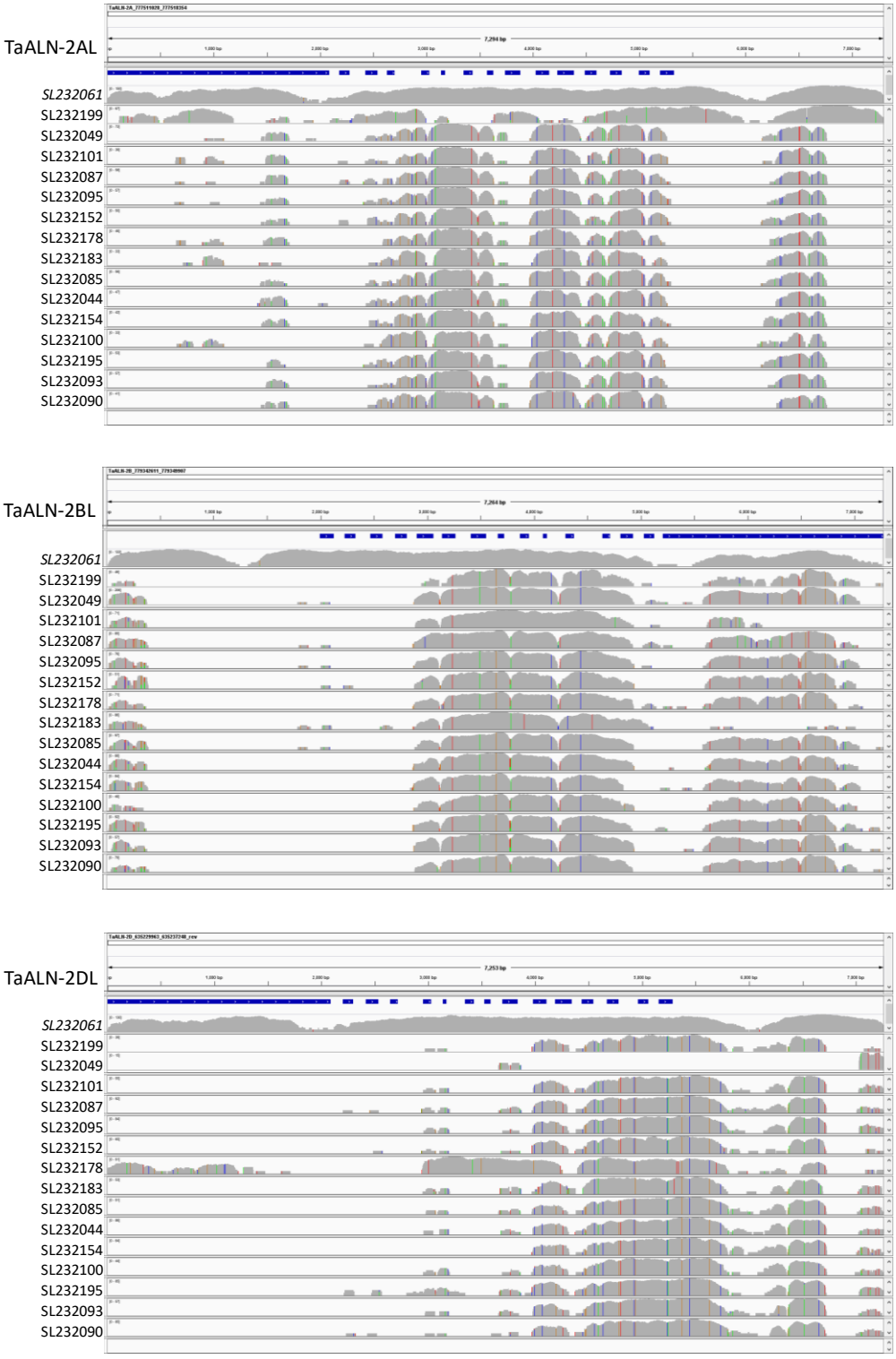


Fig. S5-4. Read mapping details of bread wheat accessions with low coverage in all TaALN all homeologs. At the top of each panel, *TaALN* 2kb promoter and exons are depicted in blue and correspond to gene model of reference genome of cv. Chinese Spring. Below, SL232061 corresponds to reads sequenced from Chinese Spring DNA sample, while the remaining are accessions with low reads coverage. Top: TaALN-2AL homeolog; middle: TaALN-2BL homeolog; bottom: TaALN-2DL homeolog. The figure was prepared with the support of the IGV 2.4.1 software.

References

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2003. Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell* **15**, 63-78.

Alamillo JM, Diaz-Leal JL, Sanchez-Moran MV, Pineda M. 2010. Molecular analysis of ureide accumulation under drought stress in *Phaseolus vulgaris* L. *Plant, Cell & Environment* **33**, 1828-1837.

Bailey TL, Williams N, Misleh C, Li WW. 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research* **34**, W369-W373.

Bi Y-M, Wang R-L, Zhu T, Rothstein SJ. 2007. Global transcription profiling reveals differential responses to chronic nitrogen stress and putative nitrogen regulatory components in Arabidopsis. *BMC Genomics* **8**, 281.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120.

Bowne JB, Erwin TA, Juttner J, Schnurbusch T, Langridge P, Bacic A, Roessner U. 2011. Drought responses of leaf tissues from wheat cultivars of differing drought tolerance at the metabolite level. *Molecular Plant* **5**, 418-429.

Bucher P. 1990. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *Journal of Molecular Biology* **212**, 563-578.

Carvalho A, Guedes-Pinto H, Lima-Brito J. 2011. Intergenic spacer length variants in Old Portuguese bread wheat cultivars. *Journal of genetics* **90**, 203-208.

Casartelli A, Riewe D, Hubberten HM, Altmann T, Hoefgen R, Heuer S. 2018. Exploring traditional aus-type rice for metabolites conferring drought tolerance. *Rice* **11**, 9.

Charlson DV, Korth KL, Purcell LC. 2009. Allantoate amidohydrolase transcript expression is independent of drought tolerance in soybean. *Journal of Experimental Botany* **60**, 847-851.

Chinnusamy V, Ohta M, Kanrar S, Lee BH, Hong X, Agarwal M, Zhu JK. 2003. ICE1: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. *Genes & Development* **17**, 1043-1054.

Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E. 2014. Structural and functional partitioning of bread wheat chromosome 3B. *Science* **345**, 1249721.

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.

Dubcovsky J, Dvorak J. 2007. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* **316**, 1862-1866.

Durek P, Schmidt R, Heazlewood JL, Jones A, MacLean D, Nagel A, Kersten B, Schulze WX. 2010. PhosPhAt: the *Arabidopsis thaliana* phosphorylation site database. An update. *Nucleic Acids Research* **38**, D828-834.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797.

Hanks JF, Tolbert NE, Schubert KR. 1981. Localization of Enzymes of Ureide Biosynthesis in Peroxisomes and Microsomes of Nodules. *Plant Physiology* **68**, 65-69.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* **27**, 297-300.

Ho Y-Y, Hsieh H-C, Huang C-Y. 2011. Biochemical Characterization of Allantoinase from *Escherichia coli* BL21. *The Protein Journal* **30**, 384-394.

Irani S, Todd CD. 2016. Ureide metabolism under abiotic stress in *Arabidopsis thaliana*. *Journal of Plant Physiology* **199**, 87-95.

Kim GJ, Lee DE, Kim H-S. 2000. Functional expression and characterization of the two cyclic amidohydrolase enzymes, allantoinase and a novel phenylhydantoinase, from *Escherichia coli*. *Journal of Bacteriology* **182**, 7021-7028.

Kim K, Kim M-I, Chung J, Ahn J-H, Rhee S. 2009. Crystal Structure of Metal-Dependent Allantoinase from *Escherichia coli*. *Journal of Molecular Biology* **387**, 1067-1074.

Laloum T, De Mita S, Gamas P, Baudin M, Niebel A. 2013. CCAAT-box binding transcription factors in plants: Y so many? *Trends in Plant Science* **18**, 157-166.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**, 357-359.

Lescano C, Martini C, González C, Desimone M. 2016. Allantoin accumulation mediated by allantoinase downregulation and transport by Ureide Permease 5 confers salt stress tolerance to *Arabidopsis* plants. *Plant Molecular Biology* **91**, 581-595.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079.

Li J, Qin R-Y, Li H, Xu R-F, Yang Y-C, Ni D-H, Ma H, Li L, Wei P-C, Yang J-B. 2015. Low-Temperature-Induced Expression of Rice Ureidoglycolate Amidohydrolase is Mediated by a C-Repeat/Dehydration-Responsive Element that Specifically Interacts with Rice C-Repeat-Binding Factor 3. *Frontiers in plant science* **6**.

Liu D, Zhang L, Hao M*, et al.* 2017. Wheat breeding in the hometown of Chinese Spring. *The Crop Journal*.

Lopes MS, El-Basyoni I, Baenziger PS, Singh S, Royo C, Ozbek K, Aktas H, Ozer E, Ozdemir F, Manickavelu A. 2015. Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *Journal of Experimental Botany* **66**, 3477-3486.

Mulrooney SB, Hausinger RP. 2003. Metal ion dependence of recombinant *Escherichia coli* allantoinase. *Journal of Bacteriology* **185**, 126-134.
Nam M, Bang E, Kwon T, Kim Y, Kim E, Cho K, Park W, Kim B, Yoon I. 2015. Metabolite Profiling of Diverse Rice Germplasm and Identification of Conserved Metabolic Markers of Rice Roots in Response to Long-Term Mild Salinity Stress. *International Journal of Molecular Sciences* **16**, 21959-21974.
Nourimand M, Todd CD. 2016. Allantoin increases cadmium tolerance in Arabidopsis via activation of antioxidant mechanisms. *Plant & Cell Physiology* **57**, 2485-2496.
Oliver MJ, Guo L, Alexander DC, Ryals JA, Wone BW, Cushman JC. 2011. A sister group contrast using untargeted global metabolomic analysis delineates the biochemical regulation underlying desiccation tolerance in *Sporobolus stapfianus*. *The Plant Cell* **23**, 1231-1248.
Pate JS, Atkins CA, White ST, Rainbird RM, Woo KC. 1980. Nitrogen Nutrition and Xylem Transport of Nitrogen in Ureide-producing Grain Legumes. *Plant Physiology* **65**, 961-965.
Ramazzina I, Folli C, Secchi A, Berni R, Percudani R. 2006. Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nature Chemical Biology* **2**, 144.
Rushton PJ, Somssich IE, Ringler P, Shen QJ. 2010. WRKY transcription factors. *Trends in Plant Science* **15**, 247-258.
Scaraffia PY, Tan G, Isoe J, Wysocki VH, Wells MA, Miesfeld RL. 2008. Discovery of an alternate metabolic pathway for urea synthesis in adult *Aedes aegypti* mosquitoes. *Proceedings of the National Academy of Sciences* **105**, 518-523.
Schubert KR. 1986. Products of biological nitrogen fixation in higher plants: synthesis, transport, and metabolism. *Annual Review of Plant Physiology*, 539-574.
Sears E, Miller T. 1985. The history of Chinese Spring wheat. *Cereal Research Communication*, 261-263.
Stalberg K, Ellerstom M, Ezcurra I, Ablov S, Rask L. 1996. Disruption of an overlapping E-box/ABRE motif abolished high transcription of the napA storage-protein promoter in transgenic *Brassica napus* seeds. *Planta* **199**, 515-519.
Takagi H, Ishiga Y, Watanabe S*, et al.* 2016. Allantoin, a stress-related purine metabolite, can activate jasmonate signaling in a MYC2-regulated and abscisic acid-dependent manner. *Journal of Experimental Botany* **67**, 2519.
Vogels G, Van der Drift C. 1966. Allantoinases from bacterial, plant and animal sources II. Effect of bivalent cations and reducing substances on the enzymic activity. *Biochimica et Biophysica Acta (BBA)-Enzymology and Biological Oxidation* **122**, 497-509.
Vogels GD, Van der Drift C. 1976. Degradation of purines and pyrimidines by microorganisms. *Bacteriological Reviews* **40**, 403-468.
Watanabe S, Matsumoto M, Hakomori 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.
Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-1191.
Werner AK, Sparkes IA, Romeis T, Witte CP. 2008. Identification, biochemical characterization, and subcellular localization of allantoate amidohydrolases from Arabidopsis and soybean. *Plant Physiology* **146**, 418-430.
Werner AK, Witte CP. 2011. The biochemistry of nitrogen mobilization: purine ring catabolism. *Trends in Plant Science* **16**, 381-387.
Yang J, Han KH. 2004. Functional characterization of allantoinase genes from Arabidopsis and a nonureide-type legume black locust. *Plant Physiology* **134**, 1039-1049.
Yobi A, Wone BW, Xu W, Alexander DC, Guo L, Ryals JA, Oliver MJ, Cushman JC. 2013. Metabolomic profiling in *Selaginella lepidophylla* at various hydration states provides new insights into the mechanistic basis of desiccation tolerance. *Molecular Plant* **6**, 369-385.
Zhao M, Ding H, Zhu JK, Zhang F, Li WX. 2011. Involvement of miR169 in the nitrogen-starvation responses in Arabidopsis. *New Phytologist* **190**, 906-915.
Zrenner R, Stitt M, Sonnewald U, Boldt R. 2006. Pyrimidine and purine biosynthesis and degradation in plants. *Annual Review of Plant Biology* **57**, 805.

Chapter 6: General discussion & future directions

Purine catabolites are used for N transport in wheat

Little was known regarding purine catabolism contribution to N homeostasis in plants that do not use allantoin and allantoate for long-distance transport of N, such as the dicotyledon model plant *Arabidopsis thaliana* or monocotyledonous plants, such as bread wheat. The research undertaken here provides solid evidence that allantoin is used for nitrogen remobilisation during grain filling in bread wheat. In particular, allantoin progressively accumulated in developing grains and its concentration at maturity was comparable with those of free amino acids (Chapter 4). These free N pools can represent a readily available source during early stages of germination. My study further demonstrated that the growth of bread wheat can be supported by purine catabolites (xanthine and allantoin) as the sole N source. This suggests that wheat is well equipped with systems mediating the uptake of purine catabolic intermediates from external media and for their catabolism internally to recycle the released nitrogen. Similar experiments conducted previously in Arabidopsis revealed contrasting results. In fact, Col-0 wild type (WT) lines were able to grow on agar media supplemented with allantoin or xanthine as a sole source of N (Desimone *et al.*, 2002; Brychkova *et al.*, 2008), however, they showed significant growth delays when compared to seedlings grown on media supplemented with inorganic N. Further work is therefore needed to dissect this in more details and assess whether the difference lies in the media used (liquid or agar) or whether it is species-specific (e.g. wheat is more efficient in purine bases uptake and/or utilisation).To test this, we could grow Arabidopsis seedlings in a hydroponic system with a similar setup described in Chapter 4. Intriguingly, the results of the hydroponic study may have important implications in a broader context, as this may indicate that organic sources of N, such as products of nucleic acid break down, may be readily taken up by wheat. Purines are the most widespread heterocycle molecules in nature (Rosemeyer, 2004) and together with pyrimidines were estimated to account for 35% of the total soil organic N (Schulten and Schnitzer, 1997). Perhaps, these sources are taken up into the root when inorganic N becomes limiting. They additionally represent an energy effective source of reduced N and C. UPS transporters, able to translocate allantoin and other nucleotide derivates, have been identified and

characterised in common bean (*Phaseolus vulgaris*) and in Arabidopsis (Desimone *et al.*, 2002; Pélissier *et al.*, 2004). However, this transporter family has not yet been characterised in cereals. Identification of *UPS* orthologous in bread wheat and analysis of their substrate specificity and expression pattern may allow us to further understand whether allantoin and xanthine are taken up from growth media and if ureides are translocated across plant compartments. Schmidt *et al.* (2006) reported that AtUPS1 and AtUPS5l had higher affinities for xanthine than for allantoin, whilst AtUPS2 had similar affinities for both. Intriguingly, these observations may support our results showing that wheat seedlings supplied with 0.25 mM xanthine displayed a higher shoot to root ratio than with 0.25 mM allantoin supplementation (Chapter 4). Further, a comparison of tissue specific expression of bread wheat *UPS* genes during grain filling may also provide important evidence of the contribution of allantoin to grain filling. Ultimately, allantoin isotope labelling studies are required to confirm the hypothesis that allantoin is actively used as an N transport form in bread wheat.

Ureides degradation provides nutrient supply under N deficiency

Disruption of *ALN* the allantoin-degrading gene in Arabidopsis had no apparent effects on plant growth under control conditions (sufficient N supplementation) (Yang and Han, 2004) prompting researchers to question whether the recycling of purines was important for N economy. My study of purine catabolism in wheat suggested that the catabolic activity of the pathway is enhanced in plants grown under N deficiency (Chapter 4). This is supported by the concomitant reduction of allantoin and up-regulation of the key gene *TaALN*. Together with the evidence that uptake and catabolism of xanthine and allantoin by wheat plants can support their growth, we can speculate that enhanced degradation of purine-containing compounds under nitrogen restriction represents an alternative supply of N, C and P that plants can access to maintain growth, at least in the short term. In order to ultimately prove this we could supply [^{15}N]-allantoin and other purine/pyrimidine bases to plants and assess the recovery of [^{15}N] in other N pools, such as amino acids. We could further dissect the link between the liberated ammonium and N assimilation machinery by employing mutants in GS/GOGAT that would allow us to determine if other pathways may be involved in this process. Additionally, analysis of TaALN enzymatic activity under N deficiency would support the transcriptional and metabolic data reported here. To further prove the importance of N remobilisation through the purine catabolic pathway during N deficiency we could undertake a transgenic approach. Interestingly, Werner *et*

al. (2013) reported that *Ataln* KO lines overexpressing *AtALN* (driven by the 35S promoter) showed higher dry weight than Col-0 lines when grown on 10 mM allantoin as a sole N source for six weeks. This suggests that the recycling of N through purine catabolic intermediates can be controlled by fine tuning *ALN* expression. Arabidopsis transgenic lines (Col-0 and *Ataln* KO) transformed with *AtALN* under the control of an inducible promoter (*XVE*, activated by estradiol) or constitutive promoter (CaMV 35S) were produced in collaboration with Dr. Rainer Höfgen at MPI-Golm (Appendix 2). These lines are an important resource to test our hypothesis that purine recycling can supply nutrients for growth. In particular, a question that needs to address is to what extent ureide catabolism contributes to growth maintenance under N restrictions. Comparing the growth of *Ataln* KO, Col-0 and *AtALN* overexpression lines under prolonged N deficiency would help address this question. In particular, the ability to control *AtALN* expression, would allow us to fine tuning the rates of allantoin degradation, by increasing or decreasing the levels of *XVE*::*AtALN* construct induction. These could be correlated to specific growth traits, such as biomass or chlorophyll, and will ultimately provide a quantitative measure of ureides contribution toward plant nutrition. Of great interest would be to dissect this in bread wheat, however, undertaking a similar approach in wheat is constrained by the complexity of its hexaploid genome. Diploid cereals, such as rice, for which mutant lines are available, represent an ideal alternative. For this purpose, rice insertional mutants in the *XDH*, *ALN* and *AAH* genes were sourced from different international collections (Appendix 3) and are available for further studies.

Has the TaXDH2-6DS homeolog acquired a novel function?

Purine catabolism has recently gained attention because of its involvement in plant abiotic stress responses, however, reports showed tobacco and wheat XDH to be involved in the biotrophic growth of certain fungi (Montalbini, 1992; Marte and Montalbini, 2001). Recently, Ma *et al.* (2016) elegantly showed that AtXDH1 in Arabidopsis accession MS-0 is involved in the mechanism of resistance against powdery mildew by producing reactive oxygen species (ROS) in the fungal haustorium. The authors proposed that AtXDH1 may function as an oxidase producing superoxide anions using NADH as substrate instead of xanthine: $NADH + O_2 \rightarrow NAD^+ + O_2^{\cdot-}$. Identification of *XDH* orthologous in bread wheat revealed that a duplication of the *XDH* gene had occurred after speciation from barley. In fact, *TaXDH* homeologs were located on chromosome group 1 and group 6 (Chapter 3). Protein sequence and structure analyses revealed a likely loss of function of the xanthine binding site (moco-binding

domain) in TaXDH2-6DS, which is the only homeolog on chromosome group 6 that is expressed. Whilst no alterations were observed in the other two main protein domains (2Fe-2S and FAD-binding) that can bind and oxidise NADH to produce superoxide in an uncoupled reaction from moco-binding domain (Zarepour *et al.*, 2010). The possibility that TaXDH2-6DS may have acquired a specific function as an oxidase is intriguing. Further experimental studies are required to confirm the biological function of the *TaXDH2-6DS* homeolog. As a first step, *in vitro* biochemical characterisation of purified TaXDH2-6DS protein is required to test our hypothesis that TaXDH2-6DS can use NADH but not xanthine as a substrate. In addition, detailed transcriptional analysis should be performed to assess a possible cell-type specific expression, which would explain the relatively low *TaXDH2-6D* expression when compared to *TaXDH1-1AL/-1BL/-1DL* expression in whole tissue analyses (Chapter 3).

Purine metabolism under drought stress: catabolism or salvage?

The data reported in this thesis in relation to the regulation of purine catabolism under drought are in accordance with the evidence reported in the literature showing allantoin accumulation to be coordinated with transcriptional changes of purine catabolic genes (Chapter 4). Moreover, up-regulation of genes encoding for enzymes involved in the conversion of xanthine to allantoin (i.e. *TaXDH1*, *TaUOX* and *TaAS*) suggests that an increased amount of xanthine precursors are directed towards the catabolic pathway. However, xanthine precursors (e.g. hypoxanthine, guanine and xanthosine) can also take part in purine salvage pathway that mediate the inter-conversion of purine bases, nucleosides and nucleotides that are derived from several metabolic processes (Zrenner *et al.*, 2006). This pathway allows the regeneration of plant nucleotide pools and requires less energy than *de novo* purine synthesis (Zrenner *et al.*, 2006). It would be of interest to assess whether purine bases are preferentially salvaged or degraded under drought stress. For example, adenosine kinase (ADK) that mediates the salvage of adenosine to adenosine monophosphate (AMP) was reported to increase under drought in a recent proteomics study conducted in maize (Kim *et al.*, 2014). Purine derivatives are important substrates for the synthesis of functional molecules and secondary metabolites. In fact, ATP is used for S-adenosyl methionine (SAM) synthesis that, in turn, is required for the production of polyamines and ethylene, both known to be involved in drought-related responses (for review see Gupta *et al.*, 2013; Müller and Munné-Bosch, 2015). As purine catabolic intermediates cannot be salvaged (Parks and Agarwal, 1972; Nguyen, 1978), diverting purine pools towards degradation would reduce substrate availability

for other reactions. The phytohormone ethylene negatively regulates ABA metabolism (Ghassemian *et al.*, 2000; Harrison, 2012), which in turn, is activated by allantoin accumulation (Watanabe *et al.*, 2014; Takagi *et al.*, 2016). Although merely speculative, this observation may indicate that the coordination between purine salvage and degradation acts in synergy with the components belonging to ABA and ethylene signalling. Comprehensive studies, integrating purine metabolomics analysis with ABA and ethylene pathway analyses under drought stress would be required to address this in detail.

Allantoin accumulates under drought to improve plant N economy

The physiological advantage of accumulating allantoin under stress in relation to N metabolism still remains elusive in the available literature. We hypothesized that accumulating allantoin under drought stress prevents the formation of ammonium, which under drought, may not be efficiently re-assimilated by the GS/GOGAT cycle (Chapter 2). Allantoin degradation would otherwise lead to ammonia (NH_3) emission from leaves and consequent N loss from plant shoots as has been reported in the literature (Mattsson *et al.*, 1997; Kumagai *et al.*, 2011). Quantifying NH_3 emissions under drought would therefore address our question. Ammonia emissions have been correlated with the activity of the GS2 enzyme and experimentally quantified in rice cultivars with contrasting GS activities (Kumagai *et al.*, 2011), as well as in barley carrying a detrimental mutation in GS2, the isoform involved in reassimilating ammonium generated from photorespiration (Mattsson *et al.*, 1997). Early report in perennial C_4 grasses showed that NH_3 volatilisation accounted for 2-10% of the N shoot reduction under drought (Heckathorn and DeLucia, 1995). However, the authors reported that NH_3 emissions generally decreased under drought as a consequence of stomatal closure and that the majority of N was translocated to the root system (up to 70% of the total N). However, to the best of our knowledge, there are no reports on NH_3 emissions under drought in annual C_3 cereals. Studies on rice and wheat previously suggested that drought-tolerant genotypes have higher GS activity than sensitive genotypes under drought (Nagy et al., 2013; Singh and Ghosh, 2013), therefore, it would be expected that NH_3 emissions under drought are lower in tolerant genotypes. The fact that genotypes with superior tolerance were shown to accumulate allantoin with a higher magnitude is in support of this hypothesis. We could determine NH_3 emissions from plant shoots by collecting NH_3 in an acid trap and measuring by a colorimetric method (Kumagai *et al.*, 2011). Alternatively, Mattsson and Schjoerring (1996) reported a method for real-time measurements employing an NH_3 monitor and detection by conductometry. Bread wheat genotypes with contrasting drought

tolerance and/or GS activity may be tested to assess if genotypes with more efficient N assimilation under drought are also emitting lower amounts of volatile NH_3. However, in order to show that allantoin accumulation contributes to reduced NH_3 emission requires further analyses. This could be tested by feeding labelled purine precursors of allantoin to the plant shoot and then measure allantoin isotope enrichment in tissues and volatile labelled NH_3 under drought by GC-MS. Alternatively, allantoin contribution to NH_3 emission may be tested in the aforementioned transgenic Arabidopsis lines (Appendix 2). Inducing *AtALN* overexpression upon water stress would prevent allantoin to accumulate, thereby an increment of emitted NH_3 would be expected. To confirm the relevance of allantoin accumulation under stress in cereals, the rice mutant collection (Appendix 3) should be tested with a similar approach. Although, complementation with an inducible *OsALN* construct would be required first, as null controls (wild type) would accumulate allantoin when subjected to drought stress.

Allantoin and stress tolerance in wheat

Allantoin association with stress tolerance has also been suggested by recent metabolomics studies (Bowne *et al.*, 2011; Oliver *et al.*, 2011; Degenkolbe *et al.*, 2013; Yobi *et al.*, 2013; Nam *et al.*, 2015; Casartelli *et al.*, 2018). In this thesis, the Australian bread wheat genotypes RAC875 and Mace were characterised under drought (Chapter 4). The former has been an object of previous studies for its high tolerance to cyclic drought (Izanloo *et al.*, 2008; Bowne *et al.*, 2011). However, in this research and other unpublished work (Pers. Comm. V. Melino 16/09/2017), the recently released Australian variety Mace showed slightly better performances under short-term terminal drought stress, as Mace yield loss was lower than that of RAC875. Interestingly, greater differences in allantoin levels between RAC875 and Mace were observed under control conditions (optimal water and N supply) than under stress (Chapter 4). In fact, analysis of mature grain revealed that Mace accumulated 66% more allantoin than RAC875. Previous work suggested that grain allantoin is rapidly used during seed germination in wheat grains (Montalbini 1992) and, in rice, grain allantoin concentration positively correlated with survival of seedlings subjected to water deficit and cold stress (Wang *et al.*, 2012). Therefore, screening a large wheat diversity panel for grain allantoin could allow us to associate this metabolite to relevant traits that can increase plant performances under stress at later stages of development. Generally, it can be assumed that breeding wheat in the dry environments of southern Australia produce genotypes that can withstand drought better than genotypes bred under favourable

conditions. This may explain why there was no discernible difference in allantoin accumulation in these two wheat genotypes under drought. Increasing the genetic diversity of wheat accessions tested for allantoin accumulation under drought is thus required for assessing the relationship between purine catabolism and drought tolerance in wheat. For example, the large difference in allantoin accumulation reported in the rice drought study by Casartelli *et al.* (2018) (Chatper 2 and Appendix 1) was observed between significantly contrasting genotypes: the aus-type rice varieties Dular and N22 that are landraces compared to IR64 and IR74 that are modern irrigated varieties. Evidence of genetic diversity of *TaALN* homeologs among bread wheat genotypes was discussed in Chapter 5. Interestingly, several wheat accessions with *TaALN* sequences that were highly divergent from the *TaALN* gene of the reference cultivar Chinese Spring were identified. Mutations that could putatively alter TaALN enzyme activity were also identified. These wheat accessions represent a valuable resource for testing allantoin accumulation in relation to drought tolerance. The high-throughput allantoin GC-QqQ-MS that was developed as part of this research (Chapter 4) would allow us to screen a large number of wheat flag leaves and grains from different accessions grown under both well-water and drought conditions. Our ultimate goal is to establish a link between allantoin in flag leaves under drought with reduced yield losses and/or allantoin in the grain as a predictor of growth performance at a later stage of development.

In summary, the research conducted and presented in this dissertation is the first comprehensive molecular characterisation of purine catabolism in a monocot grass. In particular, the data reported here highlights the nutritional role of purine catabolic intermediates, which so far have been a neglected area of research in cereals because of their presumed preferential use of amino acids for long-distance transport of N. The high N to C ratio of ureides means that they are very energy-efficient forms of N that can be readily released or remobilised to actively growing plant tissues. The data reported here provide the basis for further fundamental research exploring the dual role of ureides in wheat, both in N recycling and abiotic stress tolerance. In addition, bread wheat genotypes were identified as ideal candidates for further analysis of allantoin under stress and perhaps allantoin could be ultimately used as a trait in breeding program.

References (Chapter 1, 2 and 6)

Alamillo JM, Diaz-Leal JL, Sanchez-Moran MV, Pineda M. 2010. Molecular analysis of ureide accumulation under drought stress in *Phaseolus vulgaris* L. *Plant, Cell & Environment* **33**, 1828-1837.

Amiour N, Imbaud S, Clément G*, et al.* 2012. The use of metabolomics integrated with transcriptomic and proteomic studies for identifying key steps involved in the control of nitrogen metabolism in crops such as maize. *Journal of Experimental Botany* **63**, 5017-5033.

Arbona V, Manzi M, Ollas C, Gómez-cadenas A. 2013. Metabolomics as a Tool to Investigate Abiotic Stress Tolerance in Plants. *International Journal of Molecular Sciences* **14**, 4885-4911.

Asfaw S, Lipper L. 2012. Economics of plant genetic resource management for adaptation to climate change. ESA Working paper.

Atkins CA, Storer P. 1997. Reexamination of the intracellular localization of de novo purine synthesis in cowpea nodules. *Plant Physiology* **113**, 127-135.

Barbottin A, Lecomte C, Bouchard C, Jeuffroy M-H. 2005. Nitrogen remobilization during grain filling in wheat: genotypic and environmental effects. *Crop Science* **45**, 1141.

Barraclough PB, Howarth JR, Jones J, Lopez-Bellido R, Parmar S, Shepherd CE, Hawkesford MJ. 2010. Nitrogen efficiency of wheat: Genotypic and environmental variation and prospects for improvement. *European Journal of Agronomy* **33**, 1-11.

Bassett CL, Baldo AM, Moore JT, Jenkins RM, Soffe DS, Wisniewski ME, Norelli JL, Farrell RE. 2014. Genes responding to water deficit in apple (*Malus* × *domestica* Borkh.) roots. *BMC Plant Biology* **14**, 182.

Beckles DM, Roessner U. 2012. 5 - Plant metabolomics: Applications and opportunities for agricultural biotechnology. In: Altman A, Hasegawa PM, eds. *Plant Biotechnology and Agriculture*. San Diego: Elsevier Inc.

Bernard S, Møller A, Dionisio G*, et al.* 2008. Gene expression, cellular localisation and function of glutamine synthetase isozymes in wheat (*Triticum aestivum* L.). *An International Journal on Molecular Biology, Molecular Genetics and Biochemistry* **67**, 89-105.

Bernard SM, Habash DZ. 2009. The importance of cytosolic glutamine synthetase in nitrogen assimilation and recycling. *New Phytologist* **182**, 608.

Bowne JB, Erwin TA, Juttner J, Schnurbusch T, Langridge P, Bacic A, Roessner U. 2011. Drought responses of leaf tissues from wheat cultivars of differing drought tolerance at the metabolite level. *Molecular Plant* **5**, 418-429.

Bowsher CG, Lacey AE, Hanke GT, Clarkson DT, Saker LR, Stulen I, Emes MJ. 2007. The effect of Glc6P uptake and its subsequent oxidation within pea root plastids on nitrite reduction and glutamate synthesis. *Journal of Experimental Botany* **58**, 1109-1118.

Braun H-J, Atlin G, Payne T. 2010. Multi-location testing as a tool to identify plant response to global climate change. *Climate Change and Crop Production* **1**, 115-138.

Britto DT, Kronzucker HJ. 2002. NH_4^+ toxicity in higher plants: a critical review. *Journal of Plant Physiology* **159**, 567-584.

Britto DT, Siddiqi MY, Glass AD, Kronzucker HJ. 2001. Futile transmembrane NH_4^+ cycling: a cellular hypothesis to explain ammonium toxicity in plants. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4255-4258.

Brychkova G, Alikulov Z, Fluhr R, Sagi M. 2008a. A critical role for ureides in dark and senescence-induced purine remobilization is unmasked in the *Atxdh1* Arabidopsis mutant. *The Plant Journal* **54**, 496-509.

Brychkova G, Fluhr R, Sagi M. 2008b. Formation of xanthine and the use of purine metabolites as a nitrogen source in Arabidopsis plants. *Plant Signaling & Behavior* **3**, 999-1001.

Butterbach-Bahl K, Dannenmann M. 2011. Denitrification and associated soil N 2O emissions due to agricultural activities in a changing climate. *Current Opinion in Environmental Sustainability* **3**, 389-395.

Casartelli A, Riewe D, Hubberten HM, Altmann T, Hoefgen R, Heuer S. 2018. Exploring traditional aus-type rice for metabolites conferring drought tolerance. *Rice* **11**, 9.

Cheng Z, Dong K, Ge P, Bian Y, Dong L, Deng X, Li X, Yan Y. 2015. Identification of leaf proteins differentially accumulated between wheat cultivars distinct in their levels of drought tolerance. *PloS one* **10**, e0125302.

Clark W, Kelly E. 2004. Energy Efficiency in Fertilizer Production and Use. In: Clark W, Kornelis B, eds. *Efficient Use and Conservation of Energy*. Oxford, UK: Eolss.

Coleto I, Pineda M, Rodino AP, De Ron AM, Alamillo JM. 2014. Comparison of inhibition of N_2 fixation and ureide accumulation under water deficit in four common bean genotypes of contrasting drought tolerance. *Annals of Botany* **113**, 1071-1082.

Coneva V, Simopoulos C, Casaretto JA*, et al.* 2014. Metabolic and co-expression network-based analyses associated with nitrate response in rice. *BMC Genomics* **15**, 1056.

Cossani CM, Slafer GA, Savin R. 2010. Co-limitation of nitrogen and water, and yield and resource-use efficiencies of wheat and barley. *Crop and Pasture Science* **61**, 844-851.

Crawford NM, Arst HN, Jr. 1993. The molecular genetics of nitrate assimilation in fungi and plants. *Annual Review of Genetics* **27**, 115.

Daryanto S, Wang L, Jacinthe P-A. 2016. Global synthesis of drought effects on maize and wheat production. *PloS one* **11**, e0156362.
Day D, Poole P, Tyerman S, Rosendahl L. 2001. Ammonia and amino acid transport across symbiotic membranes in nitrogen-fixing legume nodules. *Cellular and Molecular Life Sciences* **58**, 61-71.
De Datta S, Tauro A, Balaoing S. 1968. Effect of Plant Type and Nitrogen Level on the Growth Characteristics and Grain Yield of Indica Rice in the Tropics. *Agronomy Journal* **60**, 643.
de Ruiter H, Kollöffel C. 1983. Arginine catabolism in the cotyledons of developing and germinating pea seeds. *Plant Physiology* **73**, 525-528.
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.
deSilva M, Purcell LC, King CA. 1996. Soybean petiole ureide response to water deficits and decreased transpiration. *Crop Science* **36**, 611-616.
Desimone M, Catoni E, Ludewig U, Hilpert M, A S, Kunze R, Tegeder M, Frommer WB, Schumacher K. 2002. A Novel Superfamily of Transporters for Allantoin and Other Oxo Derivatives of Nitrogen Heterocyclic Compounds in Arabidopsis. *The Plant Cell* **14**, 847-856.
Dhanda S, Sethi G. 2002. Tolerance to drought stress among selected Indian wheat cultivars. *The Journal of Agricultural Science* **139**, 319-326.
Duan J, Tian H, Gao Y. 2016. Expression of nitrogen transporter genes in roots of winter wheat (*Triticum aestivum* L.) in response to soil drought with contrasting nitrogen supplies. *Crop and Pasture Science* **67**, 128-136.
Ehdaie B, Merhaut DJ, Ahmadian S, Hoops AC, Khuong T, Layne AP, Waines JG. 2010. Root System Size Influences Water-Nutrient Uptake and Nitrate Leaching Potential in Wheat. *Journal of Agronomy and Crop Science* **196**, 455-466.
Farmer BH. 2008. Perspectives on the 'Green Revolution' in South Asia. *Modern Asian Studies* **20**, 175-199.
Forde BG. 2000. Nitrate transporters in plants: structure, function and regulation. *Biochimica et Biophysica Acta* **1465**, 219-235.
Forde BG, Lea PJ. 2007. Glutamate in plants: metabolism, regulation, and signalling. *Journal of Experimental Botany* **58**, 2339-2358.
Fresneau C, Ghashghaie J, Cornic G. 2007. Drought effect on nitrate reductase and sucrose-phosphate synthase activities in wheat (*Triticum durum* L.): role of leaf internal CO_2. *Journal of Experimental Botany* **58**, 2983-2992.
Galili G, Amir R, Hoefgen R, Hesse H. 2005. Improving the levels of essential amino acids and sulfur metabolites in plants. *Biological Chemistry* **386**, 817-831.
Gamuyao R, Chin JH, Pariasca-Tanaka J, Pesaresi P, Catausan S, Dalid C, Slamet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S. 2012. The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature* **488**, 535.
Garnett T, Conn V, Plett D*, et al.* 2013. The response of the maize nitrate transport system to nitrogen demand and supply across the lifecycle. *New Phytologist* **198**, 82-94.
Ghassemian M, Nambara E, Cutler S, Kawaide H, Kamiya Y, McCourt P. 2000. Regulation of abscisic acid signaling by the ethylene response pathway in Arabidopsis. *The Plant Cell* **12**, 1117-1126.
Gil-Quintana E, Larrainzar E, Seminario A, Diaz-Leal JL, Alamillo JM, Pineda M, Arrese-Igor C, Wienkoop S, Gonzalez EM. 2013. Local inhibition of nitrogen fixation and nodule metabolism in drought-stressed soybean. *Journal of Experimental Botany* **64**, 2171-2182.
González-Schain N, Dreni L, Lawas LMF, Galbiati M, Colombo L, Heuer S, Jagadish KSV, Kater MM. 2016. Genome-Wide Transcriptome Analysis During Anthesis Reveals New Insights into the Molecular Basis of Heat Stress Responses in Tolerant and Sensitive Rice Varieties. *Plant and Cell Physiology* **57**, 57-68.
Gravenmade EJ, Vogels GD, Van der Drift C. 1970. Hydrolysis, racemization and absolute configuration of ureidoglycolate, a substrate of allantoicase. *Biochimica et Biophysica Acta* **198**, 569-582.
Gupta K, Dey A, Gupta B. 2013. Plant polyamines in abiotic stress responses. *Acta Physiologiae Plantarum* **35**, 2015-2036.
Habash D, Bernard S, Schondelmaier J, Weyen J, Quarrie S. 2007. The genetics of nitrogen use in hexaploid wheat: N utilisation, development and yield. *Theoretical and Applied Genetics* **114**, 403-419.
Han M, Okamoto M, Beatty PH, Rothstein SJ, Good AG. 2015. The genetics of nitrogen use efficiency in crop plants. *Annual Review of Genetics* **49**, 269-289.
Hanks JF, Tolbert N, Schubert KR. 1981a. Localization of enzymes of ureide biosynthesis in peroxisomes and microsomes of nodules. *Plant Physiology* **68**, 65-69.
Hanks JF, Tolbert NE, Schubert KR. 1981b. Localization of Enzymes of Ureide Biosynthesis in Peroxisomes and Microsomes of Nodules. *Plant Physiology* **68**, 65-69.
Hare P, Cress W. 1997. Metabolic implications of stress-induced proline accumulation in plants. *Plant Growth Regulation* **21**, 79-102.
Harrison MA. 2012. Cross-talk between phytohormone signaling pathways under both optimal and stressful environmental conditions. *Phytohormones and abiotic stress tolerance in plants*: Springer, 49-76.
Häusler RE, Lea PJ, Leegood RC. 1994. Control of photosynthesis in barley leaves with reduced activities of glutamine synthetase or glutamate synthase. *Planta* **194**, 418-435.

Hawkesford MJ. 2014. Reducing the reliance on nitrogen fertilizer for wheat production. *Journal of Cereal Science* **59**, 276-283.
Hawkesford MJ, Araus J-L, Park R, Calderini D, Miralles D, Shen T, Zhang J, Parry MAJ. 2013. Prospects of doubling global wheat yields. *Food and Energy Security* **2**, 34-48.
Heckathorn S, DeLucia E. 1995. Ammonia volatilization during drought in perennial C_4 grasses of tallgrass prairie. *Oecologia* **101**, 361-365.
Henry A, Gowda VRP, Torres RO, McNally KL, Serraj R. 2011. Variation in root system architecture and drought response in rice (*Oryza sativa*): Phenotyping of the OryzaSNP panel in rainfed lowland fields. *Field Crops Research* **120**, 205-214.
Herridge DF, Atkins CA, Pate JS, Rainbird RM. 1978. Allantoin and Allantoic Acid in the Nitrogen Economy of the Cowpea (*Vigna unguiculata* [L.] Walp.). *Plant Physiology* **62**, 495-498.
Hewitt MM, Carr JM, Williamson CL, Slocum RD. 2005. Effects of phosphate limitation on expression of genes involved in pyrimidine synthesis and salvaging in Arabidopsis. *Plant Physiology and Biochemistry* **43**, 91-99.
Hill CB, Taylor JD, Edwards J, Mather D, Langridge P, Bacic A, Roessner U. 2015. Detection of QTL for metabolic and agronomic traits in wheat with adjustments for variation at genetic loci that affect plant phenology. *Plant Science* **233**, 143-154.
Hirel B, Le Gouis J, Ney B, Gallais A. 2007. The challenge of improving nitrogen use efficiency in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. *Journal of Experimental Botany* **58**, 2369-2387.
Hochman Z, Gobbett D, Horan H. 2017. Changing climate has stalled Australian wheat yields: study. *The Conversation*.
Hodson D, White J. 2007. Use of spatial analyses for global characterization of wheat-based production systems. *Journal of Agricultural Science* **145**, 115-125.
Hope P, Grose MR, Timbal B, Dowdy AJ, Bhend J, Katzfey JJ, Bedin T, Wilson L, Whetton PH. 2015. Seasonal and regional signature of the projected southern Australian rainfall reduction. *Australian Meteorological and Oceanographic Journal* **65**, 54-71.
Horridge M, Madden J, Wittwer G. 2005. The impact of the 2002–2003 drought on Australia. *Journal of Policy Modeling* **27**, 285-308.
Irani S, Todd CD. 2016. Ureide metabolism under abiotic stress in *Arabidopsis thaliana*. *Journal of Plant Physiology* **199**, 87-95.
Izanloo A, Condon AG, Langridge P, Tester M, Schnurbusch T. 2008. Different mechanisms of adaptation to cyclic water stress in two South Australian bread wheat cultivars. *Journal of Experimental Botany* **59**, 3327-3346.
James H, Brad R, Tony S*, et al.* 2016. Early sowing in South Australia results from 2015 and a summary of two years of trials. *GRDC Update Papers*.
Kaiser WM, Spill D. 1991. Rapid Modulation of Spinach Leaf Nitrate Reductase by Photosynthesis: II. *In vitro* Modulation by ATP and AMP. *Plant Physiology* **96**, 368-375.
Kamachi K, Yamaya T, Mae T, Ojima K. 1991. A Role for Glutamine Synthetase in the Remobilization of Leaf Nitrogen during Natural Senescence in Rice Leaves. *Plant Physiology* **96**, 411-417.
Kanani H, Dutta B, Klapa MI. 2010. Individual vs. combinatorial effect of elevated CO_2 conditions and salinity stress on *Arabidopsis thaliana* liquid cultures: Comparing the early molecular response using time-series transcriptomic and metabolomic analyses. *BMC Systems Biology* **4**, 177-177.
Kant S, Bi Y-M, Rothstein SJ. 2011. Understanding plant response to nitrogen limitation for the improvement of crop nitrogen use efficiency. *Journal of Experimental Botany* **62**, 1499-1509.
Kaplan F, Kopka J, Haskell DW, Zhao W, Schiller KC, Gatzke N, Sung DY, Guy CL. 2004. Exploring the Temperature-Stress Metabolome of Arabidopsis. *Plant Physiology* **136**, 4159-4168.
Kawakami N, Watanabe A. 1988. Senescence-Specific Increase in Cytosolic Glutamine Synthetase and Its mRNA in Radish Cotyledons. *Plant Physiology* **88**, 1430-1434.
Kharin VV, Zwiers FW, Zhang X, Wehner M. 2013. Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change* **119**, 345-357.
Khedr AHA, Abbas MA, Wahid AAA, Quick WP, Abogadallah GM. 2003. Proline induces the expression of salt-stress-responsive proteins and may improve the adaptation of *Pancratium maritimum* L. to salt-stress. *Journal of Experimental Botany* **54**, 2553-2562.
Kichey T, Hirel B, Heumez E, Dubois F, Le Gouis J. 2007. In winter wheat (*Triticum aestivum* L.), post-anthesis nitrogen uptake and remobilisation to the grain correlates with agronomic traits and nitrogen physiological markers. *Field Crops Research* **102**, 22-32.
Kim SG, Bae HH, Jung HJ, Lee J-S, Kim J-T, Go TH, Son B-Y, Baek S-B, Kwon Y-U, Woo M-O. 2014. Physiological and protein profiling response to drought stress in KS141, a Korean maize inbred line. *Journal of Crop Science and Biotechnology* **17**, 273-280.
Kim K, Park J, Rhee S. 2007. Structural and Functional Basis for (S)-Allantoin Formation in the Ureide Pathway. *Journal of Biological Chemistry* **282**, 23457-23464.
King CA, Purcell LC. 2005. Inhibition of N_2 fixation in soybean is associated with elevated ureides and amino acids. *Plant Physiology* **137**, 1389.

Kong L, Wang F, López-bellido L, Garcia-mina JM, Si J. 2013. Agronomic improvements through the genetic and physiological regulation of nitrogen uptake in wheat (*Triticum aestivum* L.). *Plant Biotechnology Reports* **7**, 129-139.
Kronzucker HJ, Siddiqi MY, Glass ADM. 1996. Kinetics of NH_4^+ influx in spruce. *Plant Physiology* **110**, 773–779.
Kumagai E, Araki T, Hamaoka N, Ueno O. 2011. Ammonia emission from rice leaves in relation to photorespiration and genotypic differences in glutamine synthetase activity. *Annals of Botany* **108**, 1381-1386.
Lambers H, Chapin FSI, Pons T. 2008. *Plant physiological ecology*. New York: Springer.
Lamberto I, Percudani R, Gatti R, Folli C, Petrucco S. 2010. Conserved Alternative Splicing of Arabidopsis Transthyretin-Like Determines Protein Localization and S-Allantoin Synthesis in Peroxisomes. *The Plant Cell* **22**, 1564-1574.
Leegood RC, Lea PJ, Adock MD, Hausler RE. 1995. The regulation and control of photorespiration. *Journal of Experimental Botany* **46**, 1379-1414.
Lemaire G, Oosterom Ev, Sheehy J, Jeuffroy MH, Massignam A, Rossato L. 2007. Is crop N demand more closely related to dry matter accumulation or leaf area expansion during vegetative growth? *Field Crops Research* **100**, 91-106.
Léran S, Varala K, Boyer J-C*, et al.* 2014. A unified nomenclature of NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER family members in plants. *Trends in Plant Science* **19**, 5-9.
Lescano C, Martini C, González C, Desimone M. 2016. Allantoin accumulation mediated by allantoinase downregulation and transport by Ureide Permease 5 confers salt stress tolerance to *Arabidopsis* plants. *Plant Molecular Biology* **91**, 581-595.
Li X, Lawas LMF, Malo R*, et al.* 2015. Metabolic and transcriptomic signatures of rice floral organs reveal sugar starvation as a factor in reproductive failure under heat and drought stress. *Plant, Cell & Environment* **38**, 2171-2192.
Li YC, Meng FR, Zhang CY, Zhang N, Sun MS, Ren JP, Niu HB, Wang X, Yin J. 2012. Comparative analysis of water stress-responsive transcriptomes in drought-susceptible and-tolerant wheat (*Triticum aestivum* L.). *Journal of Plant Biology*, 1-12.
Lillo C. 2008. Signalling cascades integrating light-enhanced nitrate metabolism. *Biochemical Journal* **415**, 11.
Liu J, Chen F, Olokhnuud C, Glass A, Tong Y, Zhang F, Mi G. 2009. Root size and nitrogen-uptake activity in two maize (*Zea mays*) inbred lines differing in nitrogen-use efficiency. *Journal of Plant Nutrition and Soil Science* **172**, 230-236.
Lobell DB, Schlenker W, Costa-Roberts J. 2011. Climate trends and global crop production since 1980. *Science* **333**, 616.
Lodwig EM, Hosie AH, Bourdes A, Findlay K. 2003. Amino-acid cycling drives nitrogen fixation in the legume-Rhizobium symbiosis. *Nature* **422**, 722.
Londo JP, Yu-Chung C, Kuo-Hsiang H, Tzen-Yuh C, Barbara AS. 2006. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 9578.
Loqué D, von Wirén N. 2004. Regulatory levels for the transport of ammonium in plant roots. *Journal of Experimental Botany* **55**, 1293.
Ma X, Wang W, Bittner F*, et al.* 2016. Dual and opposing roles of xanthine dehydrogenase in defense-associated reactive oxygen species metabolism in Arabidopsis. *The Plant Cell* **28**, 1108-1126.
Maksimović Z, Malenović A, Jančić B, Kovačević N. 2004. Quantification of allantoin in various *Zea mays* L. hybrids by RP–HPLC with UV detection. Die Pharmazie. *Pharmazie* **59**, 524-527.
Manivannan P, Jaleel CA, Somasundaram R, Panneerselvam R. 2008. Osmoregulation and antioxidant metabolism in drought-stressed *Helianthus annuus* under triadimefon drenching. *Comptes Rendus Biologies* **331**, 418-425.
Marschner H, Marschner P. 2012. *Mineral nutrition of higher plants*. London Waltham, MA: Elsevier Academic Press.
Marte M, Montalbini P. 2001. Effect of allopurinol treatment on biotrophic growth of some powdery mildew fungi in their specific hosts. *Physiological and Molecular Plant Pathology* **59**, 201-211.
Martin A, Lee J, Kichey T, *et al.* 2006. Two Cytosolic Glutamine Synthetase Isoforms of Maize Are Specifically Involved in the Control of Grain Production. *The Plant Cell* **18**, 3252-3274.
Masclaux-Daubresse C, Reisdorf-Cren M, Pageau K, Lelandais M, Grandjean O, Kronenberger J, Valadier M-H, Feraud M, Jouglet T, Suzuki A. 2006. Glutamine synthetase-glutamate synthase pathway and glutamate dehydrogenase play distinct roles in the sink-source nitrogen cycle in tobacco. *Plant Physiology* **140**, 444.
Matson PA, Naylor R, Ortiz-Monasterio I. 1998. Integration of Environmental, Agronomic, and Economic Aspects of Fertilizer Management. *Science* **280**, 112-115.
Mattsson M, Häusler RE, Leegood RC, Lea PJ, Schjoerring JK. 1997. Leaf-Atmosphere NH_3 Exchange in Barley Mutants with Reduced Activities of Glutamine Synthetase. *Plant Physiology* **114**, 1307-1312.
Mattsson M, Schjoerring JK. 1996. Ammonia emission from young barley plants: influence of N source, light/dark cycles and inhibition of glutamine synthetase. *Journal of Experimental Botany* **47**, 477-484.
McAllister CH, Beatty PH, Good AG. 2012. Engineering nitrogen use efficient crop plants: the current status. *Plant Biotechnology Journal* **10**, 1011-1025.
McDonald G, Hooper P. 2013. Nitrogen decision - Guidelines and rules of thumb. Grain Research and Development Corporation.
Melino VJ, Fiene G, Enju A, Cai J, Buchner P, Heuer S. 2015. Genetic diversity for root plasticity and nitrogen uptake in wheat seedlings. *Functional Plant Biology* **42**, 942.
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 of the United States of America* **104**, 4759-4764.
Micallef BJ, Shelp BJ. 1989. Arginine metabolism in developing soybean cotyledons. *Plant Physiology* **90**, 624-630.

Millenium Ecosystem Assesment. 2005. Ecosystem and Human Well-being: Desertification Synthesis. World Resources Institute, Washington DC. .
Miller A, Cramer M. 2005. Root Nitrogen Acquisition and Assimilation. *Plant and Soil* **274**, 1-36.
Miller AJ, Cookson SJ, Smith SJ, Wells DM. 2001. The use of microelectrodes to investigate compartmentation and the transport of metabolized inorganic ions in plants. *Journal of Experimental Botany* **52**, 541.
Mohanty P, Matysik J. 2001. Effect of proline on the production of singlet oxygen. *Amino Acids* **21**, 195-200.
Moll RH, Kamprath EJ, Jackson WA. 1982. Analysis and Interpretation of Factors Which Contribute to Efficiency of Nitrogen Utilization. *Agronomy Journal* **74**, 562-564.
Montalbini P. 1992. Ureides and enzymes of ureide synthesis in wheat seeds and leaves and effect of allopurinol on *Puccinia recondita* f. sp. tritici infection. *Plant Science* **87**, 225-231.
Mott IW, Wang RR-C. 2007. Comparative transcriptome analysis of salt-tolerant wheat germplasm lines using wheat genome arrays. *Plant Science* **173**, 327-339.
Müller M, Munné-Bosch S. 2015. Ethylene response factors: a key regulatory hub in hormone and stress signaling. *Plant Physiology* **169**, 32-41.
Munné-Bosch S, Alegre L. 2004. Die and let live: leaf senescence contributes to plant survival under drought stress. *Functional Plant Biology* **31**, 203-216.
Nagy Z, Németh E, Guóth A, Bona L, Wodala B, Pécsváradi A. 2013. Metabolic indicators of drought stress tolerance in wheat: Glutamine synthetase isoenzymes and Rubisco. *Plant Physiology and Biochemistry* **67**, 48-54.
Nakagawa A, Sakamoto S, Takahashi M, Morikawa H, Sakamoto A. 2007. The RNAi-mediated silencing of xanthine dehydrogenase impairs growth and fertility and accelerates leaf senescence in transgenic Arabidopsis plants. *Plant & Cell Physiology* **48**, 1484-1495.
Nam M, Bang E, Kwon T, Kim Y, Kim E, Cho K, Park W, Kim B, Yoon I. 2015. Metabolite Profiling of Diverse Rice Germplasm and Identification of Conserved Metabolic Markers of Rice Roots in Response to Long-Term Mild Salinity Stress. *International Journal of Molecular Sciences* **16**, 21959-21974.
Nguyen J. 1978. Some properties and subcellular localization of xanthine dehydrogenase in pea leaves. *Plant Science Letters* **13**, 125-132.
Nigro D, Blanco A, Anderson OD, Gadaleta A. 2014. Characterization of Ferredoxin-Dependent Glutamine-Oxoglutarate Amidotransferase (Fd-GOGAT) Genes and Their Relationship with Grain Protein Content QTL in Wheat. *PloS ONE* **9**, e103869.
Nigro D, Gu YQ, Huo N, Marcotuli I, Blanco A, Gadaleta A, Anderson OD. 2013. Structural analysis of the wheat genes encoding NADH-dependent glutamine-2-oxoglutarate amidotransferases and correlation with grain protein content. *PloS ONE* **8**, e73751.
Nikiforova VJ, Kopka J, Tolstikov V, Fiehn O, Hopkins L, Hawkesford MJ, Hesse H, Hoefgen R. 2005. Systems rebalancing of metabolism in response to sulfur deprivation, as revealed by metabolome analysis of Arabidopsis plants. *Plant Physiology* **138**, 304.
Nourimand M, Todd CD. 2016. Allantoin increases cadmium tolerance in Arabidopsis via activation of antioxidant mechanisms. *Plant & Cell Physiology* **57**, 2485-2496.
Okamoto M, Kumar A, Li W, Wang Y, Siddiqi MY, Crawford NM, Glass ADM. 2006. High-affinity nitrate transport in roots of Arabidopsis depends on expression of the NAR2-like gene *AtNRT3.1*. *Plant Physiology* **140**, 1036-1046.
Okamoto M, Vidmar JJ, Glass AD. 2003. Regulation of *NRT1* and *NRT2* gene families of *Arabidopsis thaliana*: responses to nitrate provision. *Plant & Cell Physiology* **44**, 304-317.
Oliver MJ, Guo L, Alexander DC, Ryals JA, Wone BW, Cushman JC. 2011. A sister group contrast using untargeted global metabolomic analysis delineates the biochemical regulation underlying desiccation tolerance in *Sporobolus stapfianus*. *The Plant Cell* **23**, 1231-1248.
Orsel M, Chopin F, Leleu O, Smith SJ, Krapp A, Daniel-Vedele F, Miller AJ. 2006. Characterization of a two-component high-affinity nitrate uptake system in Arabidopsis. Physiology and protein-protein interaction. *Plant Physiology* **142**, 1304-1317.
Owen AG, Jones DL. 2001. Competition for amino acids between wheat roots and rhizosphere microorganisms and the role of amino acids in plant N acquisition. *Soil Biology & Biochemistry* **33**, 651-657.
Pariasca-Tanaka J, Satoh K, Rose T, Mauleon R, Wissuwa M. 2009. Stress Response Versus Stress Tolerance: A Transcriptome Analysis of Two Rice Lines Contrasting in Tolerance to Phosphorus Deficiency. *Rice* **2**, 167-185.
Parks R, Agarwal R. 1972. 16 Purine Nucleoside Phosphorylase. *The enzymes* **7**, 483-514.
Passioura J. 1976. Physiology of grain yield in wheat growing on stored water. *Functional Plant Biology* **3**, 559-565.
Pate JS, Atkins CA, White ST, Rainbird RM, Woo KC. 1980. Nitrogen Nutrition and Xylem Transport of Nitrogen in Ureide-producing Grain Legumes. *Plant Physiology* **65**, 961-965.
Pélissier HC, Frerich A, Desimone M, Schumacher K, Tegeder M. 2004. PvUPS1, an Allantoin Transporter in Nodulated Roots of French Bean. *Plant Physiology* **134**, 664-675.
Pèlissier HC, Tegeder M. 2007. PvUPS1 plays a role in source–sink transport of allantoin in French bean (*Phaseolus vulgaris*). *Functional Plant Biology* **34**, 282.
Pessoa J, Sárkány Z, Ferreira-da-Silva F, Martins S, Almeida MR, Li J, Damas AM. 2010. Functional characterization of *Arabidopsis thaliana* transthyretin-like protein. *BMC Plant Biology* **10**, 30.

Peterhanse C, Maurino VG. 2011. Photorespiration redesigned. *Plant Physiology* **155**, 49-55.
Planchet E, Limami AM. 2015. Amino Acid Synthesis under Abiotic Stress. In: D'Mello JPF, ed. *Amino Acids in Higher Plants*. Oxfordshire and Boston: CABI Publishers, 262-276.
Qin D, Wu H, Peng H, Yao Y, Ni Z, Li Z, Zhou C, Sun Q. 2008. Heat stress-responsive transcriptome analysis in heat susceptible and tolerant wheat (*Triticum aestivum* L.) by using Wheat Genome Array. *BMC Genomics* **9**, 432.
Rai VK. 2002. Role of Amino Acids in Plant Responses to Stresses. *Biologia Plantarum* **45**, 481-487.
Ramazzina I, Folli C, Secchi A, Berni R, Percudani R. 2006. Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nature Chemical Biology* **2**, 144-148.
Rastogi R, Bate N, Sivasankar S, Rothstein S. 1997. Footprinting of the spinach nitrite reductase gene promoter reveals the preservation of nitrate regulatory elements between fungi and higher plants. *Plant Molecular Biology* **34**, 465-476.
Raun WR, Johnson GV. 1999. Improving Nitrogen Use Efficiency for Cereal Production. *Agronomy Journal* **91**, 357.
Rawat SR, Silim SN, Kronzucker HJ, Siddiqi MY, Glass AD. 1999. *AtAMT1* gene expression and NH_4^+ uptake in roots of *Arabidopsis thaliana*: evidence for regulation by root glutamine levels. *The Plant Journal* **19**, 143.
Redinbaugh MG, Huber SC, Campbell WH. 1994. Regulation of Maize Leaf Nitrate Reductase Activity Involves Both Gene Expression and Protein Phosphorylation. *Plant Physiology* **106**, 1667-1674.
Reinbothe H, Mothes K. 1962. Urea, Ureides, and Guanidines in Plants. *Annual Review of Plant Physiology* **13**, 129-149.
Riedelsheimer C, Grieder C, Melchinger AE, Lisec J, Willmitzer L, Czedik-Eysenberg A, Sulpice R, Flis A, Stitt M, Altmann T. 2012. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8872-8877.
Rose MT, Rose TJ, Pariasca-Tanaka J, Yoshihashi T, Neuweger H, Goesmann A, Frei M, Wissuwa M. 2012. Root metabolic response of rice (*Oryza sativa* L.) genotypes with contrasting tolerance to zinc deficiency and bicarbonate excess. *Planta* **236**, 959-973.
Rosemeyer H. 2004. The Chemodiversity of Purine as a Constituent of Natural Products. *Chemistry & Biodiversity* **1**, 361-401.
Roychoudhury A, Datta K, Datta SK. 2011. Abiotic stress in plants: From Genomics to Metabolomics In: Tuteja N, Gill SS, Tureja R, eds. *Omics and Plant Abiotic Stress Tolerance*. India: Bentham e books.
Sadras V, Lawson C. 2013. Nitrogen and water-use efficiency of Australian wheat varieties released between 1958 and 2007. *European Journal of Agronomy* **46**, 34-41.
Sadras VO. 2005. A quantitative top-down view of interactions between stresses: theory and analysis of nitrogen–water co-limitation in Mediterranean agro-ecosystems. *Australian Journal of Agricultural Research* **56**, 1151.
Sadras VO, Hayman PT, Rodriguez D, Monjardino M, Bielich M, Unkovich M, Mudge B, Wang E. 2016. Interactions between water and nitrogen in Australian cropping systems: physiological, agronomic, economic, breeding and modelling perspectives. *Crop and Pasture Science* **67**, 1019-1053.
Saito K, Matsuda F. 2010. Metabolomics for functional genomics, systems biology, and biotechnology. *Annual Review of Plant Biology* **61**, 463-490.
Schauer N, Kopka J, Willmitzer L, *et al.* 2006. Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nature Biotechnology* **24**, 447-454.
Schlüter U, Colmsee C, Scholz U, Bräutigam A, Weber APM, Zellerhoff N, Bucher M, Fahnenstich H, Sonnewald U. 2013. Adaptation of maize source leaf metabolism to stress related disturbances in carbon, nitrogen and phosphorus balance. *BMC Genomics* **14**, 442.
Schmidt A, Baumann N, Schwarzkopf A, Frommer W, Desimone M. 2006. Comparative studies on Ureide Permeases in *Arabidopsis thaliana* and analysis of two alternative splice variants of *AtUPS5*. *Planta* **224**, 1329-1340.
Schmidt A, Su YH, Kunze R, Warner S, Hewitt M, Slocum RD, Ludewig U, Frommer WB, Desimone M. 2004. UPS1 and UPS2 from Arabidopsis mediate high affinity transport of uracil and 5-fluorouracil. *Journal of Biological Chemistry* **279**, 44817-44824.
Schubert KR. 1981. Enzymes of Purine Biosynthesis and Catabolism in *Glycine max*: I. Comparison of Activities with N_2 Fixation and Composition of Xylem Exudate during Nodule Development. *Plant Physiology* **68**, 1115-1122.
Schubert KR. 1986. Products of biological nitrogen fixation in higher plants: synthesis, transport, and metabolism. *Annual Review of Plant Physiology*, 539-574.
Schulten H-R, Schnitzer M. 1997. The chemistry of soil organic nitrogen: a review. *Biology and Fertility of Soils* **26**, 1-15.
Sears E, Miller T. 1985. The history of Chinese Spring wheat. *Cereal Research Communication*, 261-263.
Serventi F, Ramazzina I, Lamberto I, Puggioni V, Gatti R, Percudani R. 2010. Chemical basis of nitrogen recovery through the ureide pathway: formation and hydrolysis of S-ureidoglycine in plants and bacteria. *ACS Chemical Biology* **5**, 203.
Shelp B, Atkins C, Storer P, Canvin D. 1983. Cellular and subcellular organization of pathways of ammonia assimilation and ureide synthesis in nodules of cowpea (*Vigna unguiculata* L. Walp). *Archives of biochemistry and biophysics* **224**, 429-441.
Silvente S, Sobolev AP, Lara M. 2012. Metabolite Adjustments in Drought Tolerant and Sensitive Soybean Genotypes in Response to Water Stress (Metabolite Adjustments of Soybean to Drought). *PLoS ONE* **7**, e38554.
Sinclair TR, Serraj R. 1995. Legume nitrogen fixation and drought. *Nature* **378**, 344-344.
Singh K, Ghosh S. 2013. Regulation of glutamine synthetase isoforms in two differentially drought-tolerant rice (*Oryza sativa* L.) cultivars under water deficit conditions. *Plant Cell Reports* **32**, 183-193.

Smith MT, Emerich D. 1993. Alanine dehydrogenase from soybean nodule bacteroids. Kinetic mechanism and pH studies. *Journal of Biological Chemistry* **268**, 10746-10753.
Smith PMC, Atkins CA. 2002. Purine Biosynthesis. Big in Cell Division, Even Bigger in Nitrogen Assimilation. *Plant Physiology* **128**, 793-802.
Suzuki A, Knaff D. 2005. Glutamate synthase: structural, mechanistic and regulatory properties, and role in the amino acid metabolism. *Photosynthesis Research* **83**, 191-217.
Szabados L, Savoure A. 2010. Proline: a multifunctional amino acid. *Trends in Plant Science* **15**, 89-97.
Tabuchi M, Abiko T, Yamaya T. 2007. Assimilation of ammonium ions and reutilization of nitrogen in rice (*Oryza sativa* L.). *Journal of Experimental Botany* **58**, 2319-2327.
Tabuchi M, Sugiyama K, Ishiyama K, Inoue E, Sato T, Takahashi H, Yamaya T. 2005. Severe reduction in growth rate and grain filling of rice mutants lacking OsGS1;1, a cytosolic glutamine synthetase1;1. *Plant Journal* **42**, 641-651.
Takagi H, Ishiga Y, Watanabe S*, et al.* 2016. Allantoin, a stress-related purine metabolite, can activate jasmonate signaling in a MYC2-regulated and abscisic acid-dependent manner. *Journal of Experimental Botany* **67**, 2519.
Todd CD, Polacco JC. 2006. *AtAAH* encodes a protein with allantoate amidohydrolase activity from *Arabidopsis thaliana*. *Planta* **223**, 1108-1113.
Tricker PJ, Haefele SM, Okamoto M. 2016. The interaction of drought and nutrient stress in wheat. *Water Stress and Crop Plants*: John Wiley & Sons, Ltd, 695-710.
Triplett EW, Blevins DG, Randall DD. 1982. Purification and properties of soybean nodule xanthine dehydrogenase. *Archives of Biochemistry and Biophysics* **219**, 39-46.
Tsay Y-F, Chiu C-C, Tsai C-B, Ho C-H, Hsu P-K. 2007. Nitrate transporters and peptide transporters. *FEBS Letters* **581**, 2290-2300.
Vanrensburg L, Kruger GHJ, Kruger H. 1993. Proline accumulation as drought-tolerance selection criterion - its relationship to membrane integrity and chloroplast ultrasrtucture in *Nicotiana tabacum* L. *Journal of Plant Physiology* **141**, 188-194.
Venu R, Sreerekha M, Sheshu Madhav M, Nobuta K, Mohan K, Chen S, Jia Y, Meyers B, Wang G-L. 2013. Deep transcriptome sequencing reveals the expression of key functional and regulatory genes involved in the abiotic stress signaling pathways in rice. *Journal of Plant Biology* **56**, 216-231.
Verbruggen N, Hermans C. 2008. Proline accumulation in plants: a review. *Amino Acids* **35**, 753-759.
Villas-Bôas SG, Roessner U, Hansen M. 2007. *Metabolome analysis: An introduction*. New Jersey, NJ, USA.
Wang MY, Siddiqi MY, Ruth TJ, Glass ADM. 1993. Ammonium uptake by rice roots. II. Kinetics of $^{13}NH_4^+$ influx across the plasmalemma. *Plant Physiology* **103**.
Wang P, Kong CH, Sun B, Xu XH. 2012. Distribution and function of allantoin (5-ureidohydantoin) in rice grains. *Journal of Agricultural and Food Chemistry* **60**, 2793-2798.
Wang W-S, Zhao X-Q, Li M, Huang L-Y, Xu J-L, Zhang F, Cui Y-R, Fu B-Y, Li Z-K. 2016. Complex molecular mechanisms underlying seedling salt tolerance in rice revealed by comparative transcriptome and metabolomic profiling. *Journal of Experimental Botany* **67**, 405-419.
Wasternack C, Hause B. 2013. Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany. *Annals of Botany* **111**, 1021-1058.
Watanabe S, Matsumoto M, Hakomori 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.
Watanabe S, Nakagawa A, Izumi S, Shimada H, Sakamoto A. 2010. RNA interference-mediated suppression of xanthine dehydrogenase reveals the role of purine metabolism in drought tolerance in Arabidopsis. *FEBS Letters* **584**, 1181-1186.
Werner AK, Medina-Escobar N, Zulawski M, Sparkes IA, Cao FQ, Witte CP. 2013. The ureide-degrading reactions of purine ring catabolism employ three amidohydrolases and one aminohydrolase in Arabidopsis, soybean, and rice. *Plant Physiology* **163**, 672-681.
Werner AK, Romeis T, Witte CP. 2010. Ureide catabolism in *Arabidopsis thaliana* and *Escherichia coli*. *Nature Chemical Biology* **6**, 19-21.
Werner AK, Sparkes IA, Romeis T, Witte CP. 2008. Identification, biochemical characterization, and subcellular localization of allantoate amidohydrolases from Arabidopsis and soybean. *Plant Physiology* **146**, 418-430.
Werner AK, Witte CP. 2011. The biochemistry of nitrogen mobilization: purine ring catabolism. *Trends in Plant Science* **16**, 381-387.
Winter G, Todd CD, Trovato M, Forlani G, Funck D. 2015. Physiological implications of arginine metabolism in plants. *Frontiers in plant science* **6**.
Wissuwa M, Ae N. 2001. Genotypic variation for tolerance to phosphorus deficiency in rice and the potential for its exploitation in rice improvement. *Plant Breeding* **120**, 43-48.
Witte C-P. 2011. Urea metabolism in plants. *Plant Science* **180**, 431-438.
Wu H, Liu X, You L, Zhang L, Zhou D, Feng J, Zhao J, Yu J. 2012. Effects of Salinity on Metabolic Profiles, Gene Expressions, and Antioxidant Enzymes in Halophyte *Journal of Plant Growth Regulation* **31**, 332-341.
Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ. 2006. *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature* **442**, 705-708.

Yang J, Han KH. 2004. Functional characterization of allantoinase genes from Arabidopsis and a nonureide-type legume black locust. *Plant Physiology* **134**, 1039-1049.
Yobi A, Wone BW, Xu W, Alexander DC, Guo L, Ryals JA, Oliver MJ, Cushman JC. 2013. Metabolomic profiling in *Selaginella lepidophylla* at various hydration states provides new insights into the mechanistic basis of desiccation tolerance. *Molecular Plant* **6**, 369-385.
Yoshiba Y, Kiyosue T, Katagiri T, Ueda H, Mizoguchi T, Yamaguchi-Shinozaki K, Wada K, Harada Y, Shinozaki K. 1995. Correlation between the induction of a gene for δ1-pyrroline-5-carboxylate synthetase and the accumulation of proline in *Arabidopsis thaliana* under osmotic stress. *The Plant Journal* **7**, 751-760.
Yoshida T, Mogami J, Yamaguchi-Shinozaki K. 2014. ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. *Current Opinion in Plant Biology* **21**, 133-139.
Zarepour M, Kaspari K, Stagge S, Rethmeier R, Mendel RR, Bittner F. 2010. Xanthine dehydrogenase AtXDH1 from Arabidopsis thaliana is a potent producer of superoxide anions via its NADH oxidase activity. *Plant Molecular Biology* **72**, 301-310.
Zonia LE, Stebbins NE, Polacco JC. 1995. Essential role of urease in germination of nitrogen-limited *Arabidopsis thaliana* seeds. *Plant Physiology* **107**, 1097-1103.
Zozaya-Hinchliffe M, Potenza C, Ortega JL, Sengupta-Gopalan C. 2005. Nitrogen and metabolic regulation of the expression of plastidic glutamine synthetase in alfalfa (*Medicago sativa*). *Plant Science* **168**, 1041-1052.
Zrenner R, Stitt M, Sonnewald U, Boldt R. 2006. Pyrimidine and purine biosynthesis and degradation in plants. *Annual Review of Plant Biology* **57**, 805.

Appendix 1: Exploring traditional aus-type rice for metabolites conferring drought tolerance

Library note: Appendix 1 : blank pages supplied

Appendix 2: Arabidopsis constitutive and inducible *ALN* transgenic lines

Appendix 2. Arabidopsis constitutive and inducible *ALN* transgenic lines

Target gene	Background	Construct	Eco-type
AtALN	WT	35S::ALN	Col-0
AtALN	WT	XVE::ALN	Col-0
AtALN	*aln-1*	35S::ALN	Col-0
AtALN	*aln-1*	XVE::ALN	Col-0
AtALN	*aln-2*	35S::ALN	Col-0
AtALN	*aln-2*	XVE::ALN	Col-0
AtALN	WT	35S::empty	Col-0
AtALN	WT	XVE::empty	Col-0
AtALN	*aln-1*	35S::empty	Col-0
AtALN	*aln-1*	XVE::empty	Col-0
AtALN	*aln-2*	35S::empty	Col-0
AtALN	*aln-2*	XVE::empty	Col-0

WT, wild type; SALK T-DNA insertion lines (Watanabe *et al.*, 2014): *aln-1*, SALK-000325; aln-2, SALK_146783; 35S::ALN; *AtALN* constitutive expression driven by a cauliflower mosaic virus promoter (CaMV 35S), plasmid pK7WG2; XVE::ALN, *AtALN* indiciuble expression driven by an estrogen-based receptor, plasmid pMDC7; empty, control vector.

Appendix 3: Rice insertional mutants in purine catabolic genes

Appendix 3. Rice insertional mutants in purine catabolic genes

Target gene	Insertion ID	Line name	Background	Insertion type	Region	Sense	Source
OsXDH	1B-05035L	APNA04	Nipponbare	T-DNA	Intron 13	plus	OTL
OsXDH	FL058720	NE9382	Nipponbare	Tos 17	Exon 12	plus	NIAS
OsALN	NC2763_0_504_1A	NC2763	Nipponbare	Tos 17	Exon 8	minus	NIAS
OsALN	NF0046_0_101_1A	NF0046	Nipponbare	Tos 17	Exon 8	plus	NIAS
OsALN	NF_1040_0_501_1A	NF1040	Nipponbare	Tos 17	Exon 9	minus	NIAS
OsAAH	CU325184	AHWA11	Nipponbare	Tos17	Intron 2	Minus	OTL
OsAAH	CU314235	AVTA04	Nipponbare	Tos17	Intron 5	Minus	OTL
OsAAH	FL056879	NE7420	Nipponbare	Tos 17	Exon 3	plus	NIAS

T-DNA, Agrobacterium-mediated insertion; Tos 17; transposon-mediated insertion, OTL, Oryza Tag Line, France (oryzatagline.cirad.fr); NIAS; National Institute of Agrobiological Sciences, Japan, (tos.nias.affrc.go.jp)