Understanding the molecular and genetic mechanisms regulating senescence in wheat.

Sophie Anne Harrington

A thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

John Innes Centre

September 2019

Abstract

Monocarpic senescence in wheat is a tightly regulated process that is essential in determining grain quality levels. However, despite its importance, relatively little is known about the molecular mechanisms that underpin senescence in wheat itself. In this thesis, we have taken advantage of new genetic and genomic resources in wheat to carry out forward and reverse genetics approaches with the aim of identifying novel senescence regulators and further characterizing genes with known roles in senescence.

Initially, we carried out a forward genetic screen of the Kronos TILLING population and identified a novel dominant chlorosis phenotype. Using a combination of reduced representation sequencing techniques, we were able to map the locus underpinning this phenotype to a 4.3 Mb region containing 59 genes. We then turned out attention to further characterization of the NAM-1 genes, which are known positive regulators of wheat senescence. We identified missense mutations in highly conserved residues of the NAC domain of NAM-A1 within the Kronos TILLING population. Mutations in these residues not only disrupted protein function, leading to delayed senescence, but they specifically inhibited protein-protein binding. This suggested that the residues may be essential for NAC domain function across species.

Using overexpression and heat-shock-inducible transgenic constructs, we then demonstrated that premature expression of *NAM-B1* is not sufficient to induce premature senescence. This suggested that NAM-B1 may interact with one or more proteins to drive senescence. We carried out a yeast two-hybrid screen against NAM-B1 and identified seven candidate interactors. One of these genes, the NAC transcription factor *NAC3-B1*, is part of a triad which also positively regulates senescence and whose predicted regulatory targets are enriched for senescence-related processes. We therefore hypothesized that the NAM-1 and NAC3-1 genes may work together to regulate senescence onset and progression in wheat.

Acknowledgements

This thesis would not exist without the immense support and encouragement given to me by my supervisor, Cristobal Uauy. I cannot thank him enough for being such a brilliant supervisor and mentor over the past few years. I also can't imagine what this work would be like without the fantastic help and advice I've had from Philippa Borrill, who has supported my work from the very first day in the lab. Thank you both!

During my PhD I have been lucky enough to receive brilliant advice from Alison Smith and Mark Banfield as part of my supervisory committee, to whom I am very grateful. Thank you to the wonderful students who have worked with me—Lauren and Anna, your work was stellar and I know you will do brilliant things in your PhDs. Thank you as well to my excellent collaborators- Nicolas Cobo, Jaroslav Doležel, Miroslava Karafiátová, and all the members of the BRACT team who carried out the wheat transformations. Thank you as well to the JIC Field and Horticultural teams, and the UC Davis field trials team, and to the John Innes Foundation for funding my PhD.

I would surely have lost the will to carry on during this PhD if it wasn't for the amazing people in our lab— thank you all for always being so fun and supportive, whether squelching through mud in the Peak District or getting lost in the Rocky Mountains. Special mention must go to the residents of G09- Jemima, Nikolai, Olu, Sarah, and Aura. You have made every day in work a joy, even when the science wasn't playing ball.

To my fellow rotation students— Erin, Charlotte, Roger, and Juan-Carlos— thank you for helping me find my feet in the first year and thank you to Steph for organising a program I feel so privileged to have been a part of. To the Squad— Erin, Cat, Nicola, Lizzie, and Charlie— thank you for helping make that house a home.

This thesis would never have been possible if it weren't for the love and support of my family- thank you Mum, Dad, and John. However far away you are, you've always been there for me when I needed you. And finally, to Lachlan, my love and partner in crime, without you I would be lost.

Table of Contents

Table of Figures

Table of Tables

List of Abbreviations

3AT	3-Amino-1,2,4-Triazole
ABRE	ABA-Responsive Element
ARCs	Age-Related Change Factors
BC	Backcross
BLAST	Basic Local Alignment Search Tool
CCGs	Chlorophyll Catabolic Genes
CIMMYT	International Maize and Wheat Improvement Centre
Co-IP	Co-Immunoprecipitation
CSS	Chromosome Survey Sequence
CTD	C-Terminal Domain
DAA	Days After Anthesis
DAS	Days After Sowing
DEG	Differentially Expressed Gene
DPA	Days Post-Anthesis
DPHS	Days Post-Heat Shock
Dyn-GENIE3	Dynamic GENIE3
EMS	Ethyl Methanesulphonate
EMSA	Electrophoretic Mobility Shift Assay
ER	Estrogen Receptor
FAO	Food And Agriculture Organisation
FDR	False Discovery Rate
GO	Gene Ontology
GOI	Gene Of Interest
GPC	Grain Protein Content
JA	Jasmonic Acid
JIC	John Innes Centre
JSD	Jensen-Shannon Divergence
KASP	Kompetitive Allele Specific Pcr
Kb	Kilobase
LOESS	Locally-Estimated Scatterplot Smoothing
MAGIC	Multiparent Advanced Generation Inter-Cross
Mb	Megabase
MP	Mutant Parent
NAC	*NAM*, *ATAF1/2*, and *CUC*
NAM	No Apical Meristem
NB1	NAM-B1
PBS	Predicted Biological Score
PCD	Programmed Cell Death
qRT-PCR	Quantitative RT-PCR
QTL	Quantitative Trait Locus
RE	Relative Expression
RIL	Recombinant Inbred Line
RNA-Seq	Rna Sequencing
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
SA	Salicylic Acid
SAGs	Senescence-Associated Genes
SC	Synthetic Complete

Continued on next page

SNP	Single Nucleotide Polymorphism
TF	Transcription Factor
TGW	Thousand Grain Weight
TILLING	Targeting Induced Local Lesions In Genomes
TPM	Transcripts Per Million
TSL	The Sainsbury Lab
UC	University Of California
UTR	Untranslated Region
WGD	Whole Genome Duplication
WT	Wild Type
Y2H	Yeast Two-Hybrid
YES-1	Yellow Early Senescence-1

1 General Introduction

1.1 Wheat is an essential staple crop

In the 2019/2020 growing season, global wheat production is expected to reach over 770 million tonnes, a rise of 5.6% on the previous year (FAO 2019a). This is in contrast to other staple crops such as maize and rice which show static or decreasing production levels. This rise follows the historical trend, where over the past half-century yield has increased in close correlation with the global population (Figure 1-1). This increase in wheat yields from the latter half of the 20th century helped prevent the Malthusian prediction of the human population rapidly outgrowing its potential food sources (Malthus 1798, Shewry 2009).

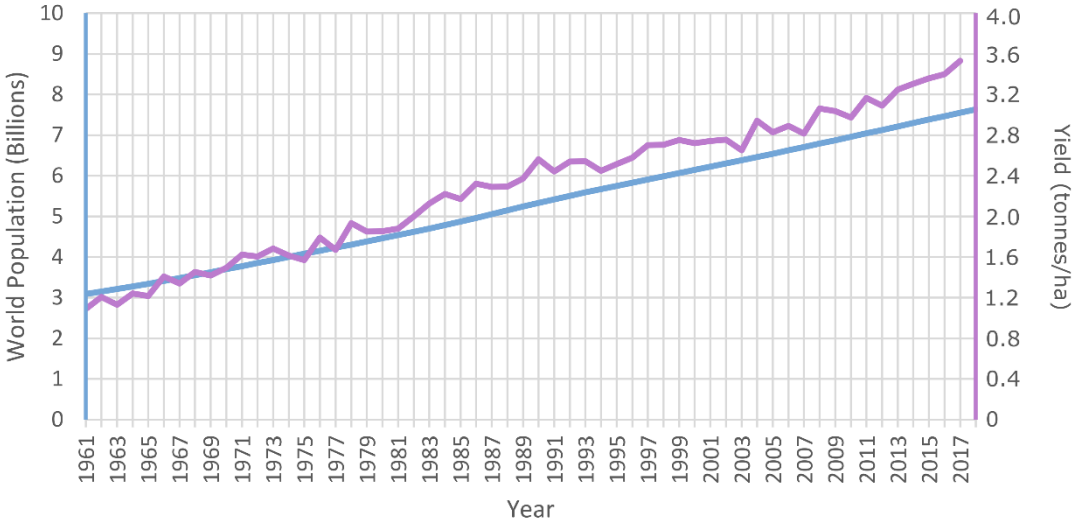

Figure 1-1: Global wheat yields closely correlate with estimated world population. The world population in billions is plotted in blue, against the left vertical axis, while global wheat yields are plotted in purple, against the right vertical axis. Data was obtained from FAOSTAT (FAO 2019b).

Since 1961, the earliest date for which the Food and Agriculture Organisation (FAO) maintains data, the number of kilocalories (kcal) provided by wheat on a per capita basis worldwide has increased from approximately 415 kcal per person per day to just under 530 kcal/day in 2013 (Figure 1-2A). The number of calories provided by wheat peaked in the late 1990s, at just over 560 kcal/day. Based on the recommended intake of approximately 2,500 calories for men and 2,000 calories for women, wheat can provide between 20% and 25% of the daily calorie intake. This level of caloric intake puts wheat and its derived food products alongside rice as the most important contributor to human calorie intake.

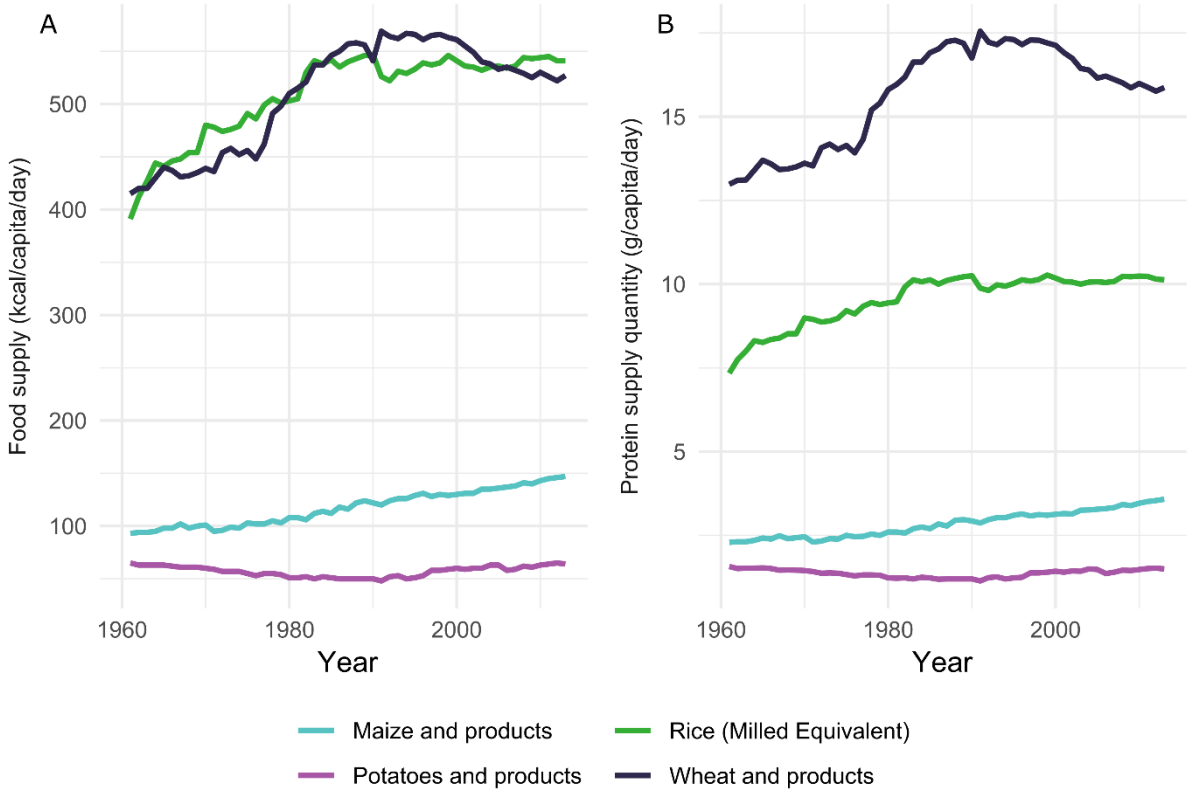

Figure 1-2: Wheat provides high levels of kilocalories and protein in the human diet. This graph depicts the relative caloric (A) and protein (B) contributions of wheat, maize, rice, and potatoes. All data was obtained from FAOSTAT (FAO 2019b).

Beyond just calories, wheat is also an essential source of protein in the human diet, today providing just under 16 g of protein per person per day (Figure 1-2B). Given a recommended protein intake of approximately 45 and 55 g/day for women and men, respectively, wheat can provide between 29 and 35% of the daily protein requirements for an adult (ONS 2014). This is substantially more than the second highest staple crop, rice, which only provides approximately 10g per person per day, despite providing a similar number of calories as wheat.

Wheat is not, of course, produced in even quantities worldwide. Its global production is highly concentrated across Europe and North America, as well as in northern India and north and north-eastern China (Figure 1-3) (Monfreda et al. 2008). Within the United Kingdom, wheat is grown across the majority of the country, with the main exception of the highland areas in Scotland, northern England, and Wales. Across the East of England and much of the Midlands, wheat makes up almost 25% of the total farming area, excluding rough grazing (Figure 1-4) (Shearer 2012). This unequal distribution of wheat production, despite the global importance of wheat in many diets, has resulted in a substantial trade in wheat products across borders. Some countries are net importers of wheat, such as Egypt which in the 2017 imported over 22 million metric tons of wheat (House 2019). In contrast, Russia exported over 40 million metric tons in the same period. As of 2016, wheat was the second most valuable crop in terms of export value globally, with a value of

approximately $35 billion USD (FAO 2019b). Wheat is therefore both a staple food crop and an economically-important agricultural commodity.

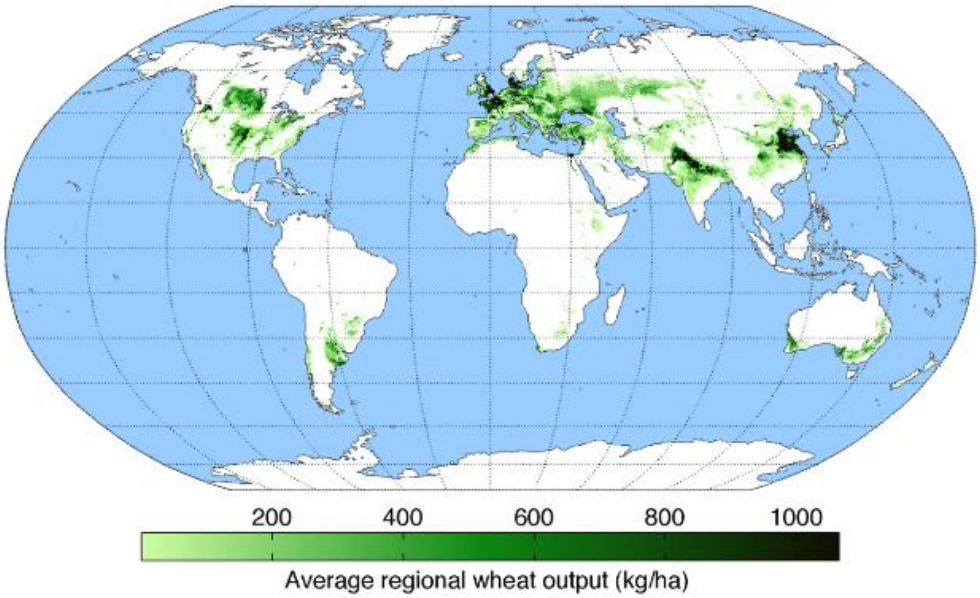

Figure 1-3: Global distribution of wheat production. Output is shown in kg of grain produced per hectare across each grid cell (size ranging from county or district to country level). Data from Monfreda et al. 2008, with map produced by AndrewMT reproduced under a CC-BY-SA 3.0 license.

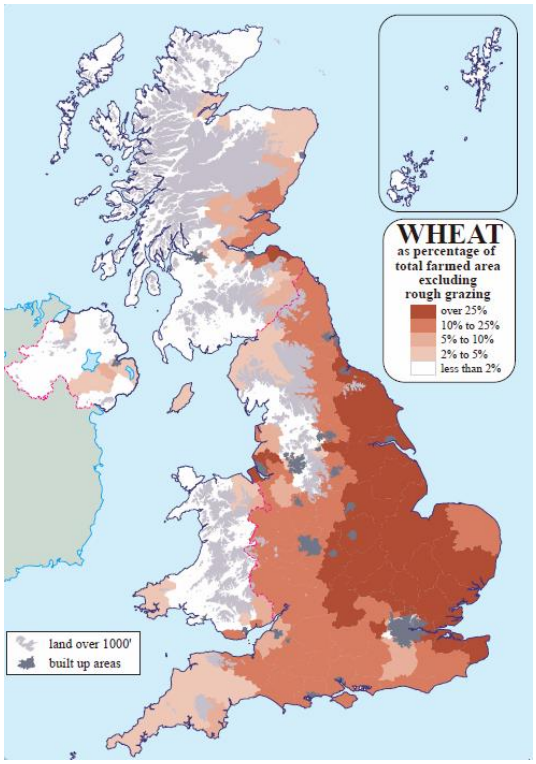

Figure 1-4: Wheat farmed area in the United Kingdom. Data is from 2012. Map was adapted from the “The United Kingdom Cereals Industry; Fifth Edition” published by AHDB (Shearer 2012).

1.2 The origin of wheat

Wheat is derived from a series of ancient hybridisations between grass species (Dubcovsky and Dvorak 2007, Borrill et al. 2019b) in and around the Fertile Crescent (Figure 1-5) (Brown et al. 2009). Tetraploid wheat, *T. turgidum*, is derived from a single hybridisation event between the wild grasses *T. urartu*, the A genome progenitor, and a species similar to *Aegilops speltoides*, the B-genome progenitor. This hybridisation occurred between 360,000 and 500,000 years ago (Huang et al. 2002, Zhang et al. 2014), forming a new tetraploid species via allopolyploidy. *T. turgidum* persisted as a wild grass (ssp. *dicoccoides*) until approximately 10 to 12,000 years ago, after which it began to be domesticated during the development of agriculture (Dubcovsky and Dvorak 2007). Domesticated Emmer wheat (most likely *T. turgidum* ssp. *dicoccum*) then underwent a second hybridisation, approximately 7000 to 10,000 years ago, with the wild goat grass *Ae. tauschii* (Feldman et al. 1995). This wild grass provided the D genome in another instance of allopolyploidy, generating hexaploid wheat, *Triticum aestivum*.

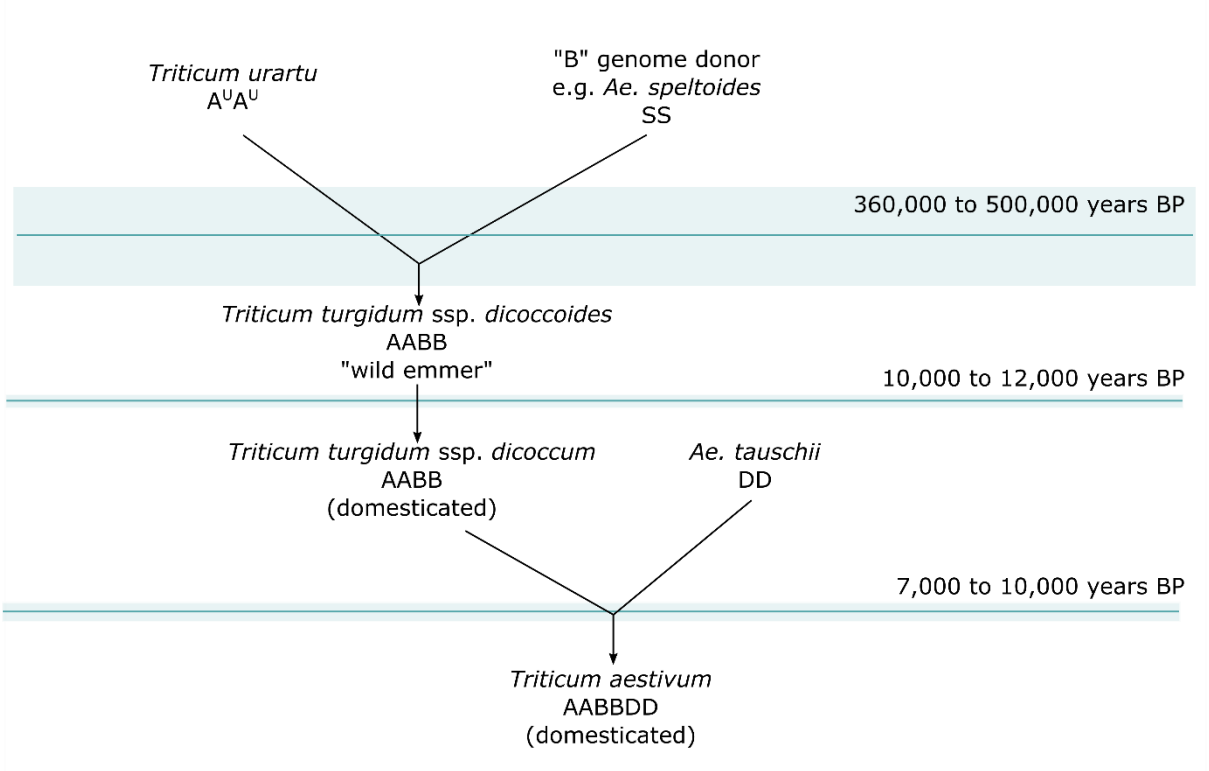

Figure 1-5: A schematic of hybridisation events leading to domesticated tetraploid and hexaploid wheat. Dates are based on those provided in Zhang et al. 2014. BP is before present (1950).

1.2.1 Selection from domestication to the Green Revolution

Following domestication of the tetraploid and hexaploid wheat, cultivation continued with consistent low-level selection for agronomically useful traits. Domestication itself likely occurred in tandem with or in proximity to the advent of free-threshing lines, containing the dominant *Q* allele that facilitates grain harvest through the loss of tough glumes, amongst other phenotypes (Simons et al. 2006). Other traits thought to have been selected for during the process of wheat

domestication include synchronous tillering and maturation, increased inflorescence number, increased grain size, and reduced seedling dormancy (Harlan et al. 1973).

The domestication of wheat heralded a step-change in human society, enabling the beginning of permanent, agrarian settlements. As the historian and philosopher Yuval Noah Harari put it in his book Sapiens: A Brief History of Humankind, "[w]ithin a couple of millennia, humans in many parts of the world were doing little from dawn to dusk other than taking care of wheat plants" (Harari 2015). Changes in domestication-related traits allowed for consistent seasons of sowing and harvest and increased the calorific content of the harvested grains. They also underpinned the delicate balance between grains being easy to remove from the mature wheat spike, but not falling off the spike prematurely (Peng et al. 2011).

Yet for the majority of wheat cultivation over the past 10,000 years, selection on agronomically useful traits occurred over slow timescales in a mostly unconscious manner. This concept was first addressed by Charles Darwin in his work The Variation of Animals and Plants under Domestication, where he highlighted the distinction between so-called "methodical selection," where systemic efforts are made to alter a particular trait, and "unconscious selection," where a natural bias occurs towards saving and propagating the seed of only the most beneficial and productive individuals (Darwin 1868). It was only by the early 19th century that concerted effort was put towards breeding in plants for beneficial traits, or what Darwin would term "methodical" selection (Lupton 1987). Much of this work was done ignorant of the effect of genetic segregation, as while Gregor Mendel first published his ideas on segregation in 1866 (Mendel 1865), they did not become part of scientific thinking until the early 1900s. In the mid-1800s, the Scottish cereal breeder Patrick Shirreff began a series of varietal intercrosses, leading to the development of new wheat varieties (Lupton 1987). Yet it wasn't until the turn of the 20th century, and the rediscovery of Mendel's law of heredity, that wheat breeding turned towards the systematic, statistically rigorous field that it is today.

During this period, distinct varieties of wheat known as landraces were being cultivated in different locations around the world. These landraces were the produce of centuries, if not millennia, of selection for their local environment, with gene flow between landraces highly limited by geography (Rufo et al. 2019). In the 1930s, A. E. Watkins collected 826 landraces which were being traded through the board of Trade in London, originating from 32 countries across Europe, the Middle East, North Africa and Asia (Wingen et al. 2014). Using population structure analysis based on a set of 14 SNP markers, the panel was found to split into ancestral sub-groups that mapped to specific geographic regions. This supports the assertion that the majority of wheat diversity at the time was dependent on geographic origin. However, the population groupings were not as strong as might have been expected if the wheat landraces had been maintained in geographical isolation, suggesting that exchange of material between farmers in different geographic regions was occurring by the 1930s. Given the expansion of global trade from the

1800s onwards (Ortiz-Ospina et al. 2014), it is possible that the presence of trade routes between specific regions led to a reduction in geographic uniqueness. Additionally, as Watkins obtained his lines through the London Board of Trade, it is possible that the cultivars obtained were geographically biased to those which had historically been part of this trade-driven germplasm exchange, as opposed to landraces which were grown in regions extant to the growing global trade network.

During the early 1900s, further advances in breeding techniques began to increase the level of selection on cultivated wheat varieties (Gayon and Zallen 1998). This included the establishment of single-seed descent breeding, where seeds from the most desirable individual plants were retained and grown again in the next generation. This contrasted with the highly heterogenous nature of most landraces, where seed was collected in bulk from multiple different plants, preventing efficient selection of desirable traits. Rowland Biffen, based at the Cambridge Plant Breeding institute, pushed for the integration of modern genetics with wheat breeding (Biffen 1905). He began using hybridisation to tackle the prevalence of low-quality wheat grown in the UK, which lacked the necessary dough strength for bread making (Kiszonas and Morris 2018). His work highlighted the importance of focussing on individual traits, rather than the whole plant, when carrying out selection on plants derived from inter-crosses. Using these methods, Biffen was able to breed a series of new varieties to increase disease resistance, such as Little Joss, and improve grain quality, such as Yeoman, which between them accounted for one quarter of the UK's planted wheat acreage by 1925 (Charnley 2009).

1.2.2 The effect of the Green Revolution on wheat breeding

Methods such as those developed by Biffen and his colleagues began to make substantial advances in breeding wheat varieties for resistance to different diseases. Yet the rate of yield increase remained low until the so-called "Green Revolution" of the 1950s. This transformation arguably began in Mexico, following the formation of the International Maize and Wheat Improvement Centre (CIMMYT) in 1943. Over the next three decades, a combination of genetic advances and the implementation of high-input agronomic practices served to substantially increase the yield potential of both wheat and other staple crops. Data obtained by the FAO on global wheat yields highlight the impact of the Green Revolution from the 1960s onwards (Figure 1-6) (Roser and Ritchie 2017, FAO 2019b). While in 1961 wheat yields were mostly at 1 to 1.5 tonnes per hectare (t/ha) outside of Europe, by 1991 yields across western Europe were reaching 6 to 7 t/ha, and between 2 and 4 t/ha in North American and China. Globally production had increased from just over 1 t/ha on average to nearly 2.5 t/ha in 1991 and over 3 t/ha in 2014, the most recent data available. Within the UK itself, wheat yields increased dramatically from a pre-Green Revolution baseline of approximately 2 to 2.5 t/ha to over 5 t/ha within three decades of the start of the Green Revolution (Figure 1-7) (Roser and Ritchie 2017).

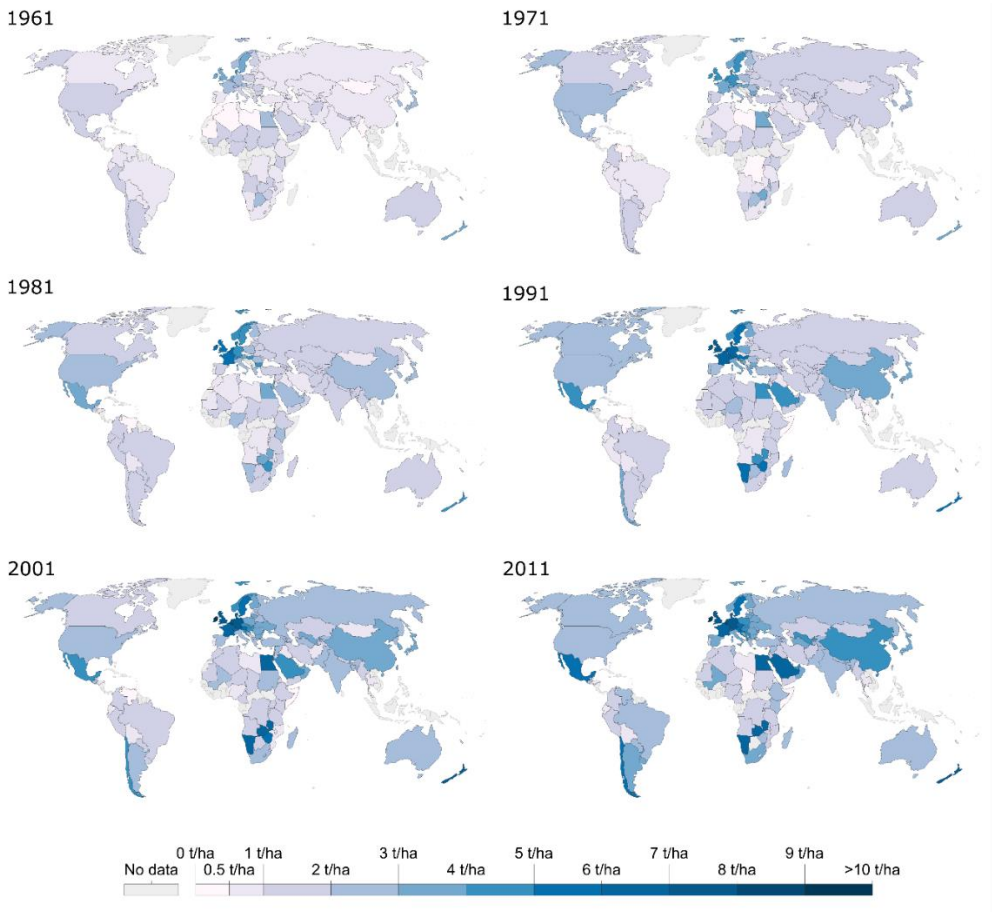

Figure 1-6: Global wheat yields (tonnes per hectare) from the early stages of the Green Revolution. Data is presented on a country-wide scale, representing the yield per hectare of wheat grown, at 10-year intervals from the start of the dataset in 1961. Data was obtained from FAOSTAT (FAO 2019b), and the maps were produced by "Our World in Data" (Roser and Ritchie 2017) and altered and reproduced here under a CC-BY license.

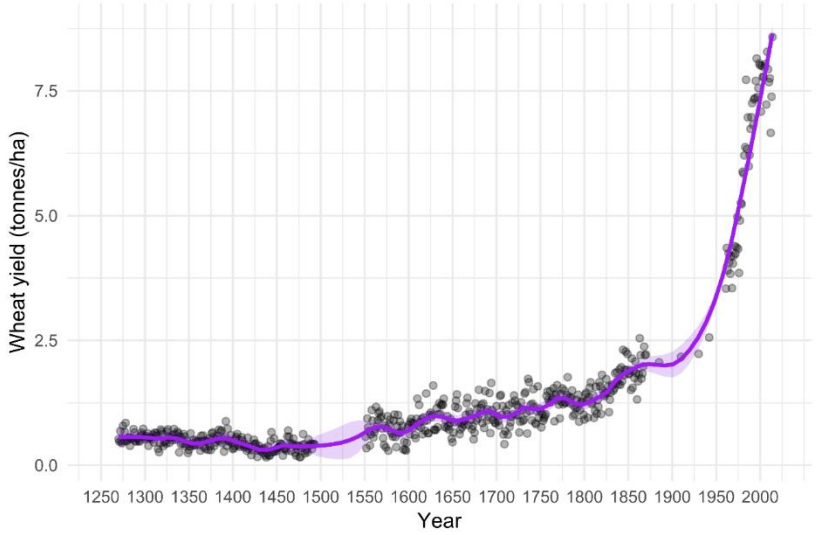

Figure 1-7: Historic wheat yields (tonnes per hectare) in the United Kingdom. A smoothed LOESS line (purple) was fit to highlight the increasing trend following the Green Revolution, with associated 95% confidence interval presented in light purple. Data was obtained from "Our World in Data" under a CC-BY license, derived from (Broadberry et al. 2015) (1270-1870), (Brassley 2000) (1885-1946), and (FAO 2019b) (1961-2014).

For wheat during the Green Revolution, increasing yield was underpinned by the development of the dwarfed and semi-dwarfed varieties by Norman Borlaug, using combinations of the *Reduced height* (*Rht*) alleles. These dwarfing genes, also essential for increasing rice yields, reduced losses caused by lodging. Lodging occurs when the stem of the plant is unable to support the spike, causing the plant to fall over. It is still observed in wheat today, particularly following strong winds and rain, but before the introduction of the dwarfing genes plant would regularly lodge as a result of the larger grains introduced through breeding for higher yield (Hedden 2003). This was compounded by the effect of nitrogen-based fertilisers, whose application led to increased biomass and correspondingly an increase in propensity for lodging (Guarda et al. 2004, Belete et al. 2018). This lodging would cause substantial crop losses, hindering the development of greater yielding varieties. Along with reducing the propensity to lodge, shorter wheat required less of the photosynthetic assimilate for stem formation and growth. This correspondingly facilitated larger yield increases by freeing up this assimilate for storage in the developing grains (Hedden 2003).

The Green Revolution also signalled the beginning of high-input agriculture, which has become the predominant form of agriculture practiced worldwide. The use of Nitrogen-based fertilisers in agriculture increased by 7-fold between 1961 and 2002, and nearly 9-fold by 2017. Increased implementation of irrigation and pesticide use also contributed to the yield increases observed from the 1960s onwards. It has been predicted that, without these efforts, the corresponding production of food in developing countries would have been 20% lower, leading to both increased prices and reduced calorie availability, and thus increased risk of famine (Pingali 2012).

1.3 Wheat is a key source of protein and nutrients

During the Green Revolution most breeding effort was focussed on increasing yield potential and stability (Davies 2003). Only more recently has substantial effort in breeding programs turned towards quality parameters that influence the end-use products of wheat crops. Historically, discussions of wheat quality in the British context referred exclusively to bread making (Kiszonas and Morris 2018). Wheat of bread-making quality is required to contain high levels of gluten (currently between 10 and 13%). Gluten comprises a complex of two proteins— gliadins and glutenin—in a 1:1 ratio which make up the vast majority (85-90%) of the protein present in the wheat grain (Biesiekierski 2017). The gluten in bread is critical for its baking properties, forming a continuous network of protein that traps the gases produced by yeast during baking and ensures the characteristic rise of leavened breads. Wheat quality parameters also include milling quality, or the ability of grains to be ground using the automated steel roller mills introduced at the turn of the 20th century (Kiszonas and Morris 2018). More recently, wheat quality has also been used to refer to the levels of macro- and micronutrients in the grain. These nutrients, including zinc, copper, and iron, are essential for human health yet are typically found at low levels in the endosperm of the grain, which is the part of the grain retained for white flour production (Figure 1-8) (Balk et al. 2019).

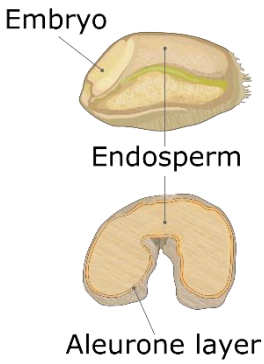

Figure 1-8: Wheat grain morphology with relevant tissues identified. Adapted from Brinton et al. 2019 under a CC-BY-4.0 license.

1.3.1 Increasing grain nutrient content could alleviate micronutrient deficiencies

Micronutrient deficiencies are widespread, particularly in developing countries, forming a so-called "silent epidemic" public health crisis (Tulchinsky 2010). The most common micronutrient deficiencies observed are in iodine, iron, zinc, and vitamin A (Tulchinsky 2010). These deficiencies can lead to a wide range of morbidities, including stunting and blindness (Black 2003, Ritchie and Roser 2017). In some countries, staple crop products are fortified with additives of these nutrients, such as iodine in salt and iron in flour. Yet many people live in areas inaccessible to such biofortification schemes, instead relying upon the nutrients present in the staple crops of their region. Unfortunately, in many cases the staple crops are themselves nutrient poor. This is often driven by cultural preferences for grains lacking the outer seed layer, often called the bran, which includes the pericarp, seed coat, and aleurone layer. White wheat flour only consists of the endosperm of grain while whole-wheat flour includes both the starchy endosperm and the bran. Whole-wheat flour typically contains, per 100 g of flour, 13 g of protein, 3.6 mg of iron and 2.6 mg of zinc. In contrast, un-enriched white flour contains only 10 g of protein, 1.2 mg of iron, and 0.7 mg of zinc (U.S. Department of Agriculture 2019).

This difference in nutrient composition stems primarily from the inclusion of the aleurone layer of the wheat grain in whole-wheat flour. The aleurone layer, which consists of a single layer of cells surrounding the wheat grain endosperm, functions as a store for lipids, minerals, and proteins in the wheat grain (Becraft and Yi 2010). Whole-wheat flour also includes the embryo, or germ, which is another key store of minerals. Previous work on iron localisation in the wheat grain, for example, has shown that in wild-type wheat, iron and zinc localise primarily to the aleurone layer and the embryo (Balk et al. 2019). Efforts to increase grain nutrient and protein levels in order to combat the effects of micronutrient deficiencies have focussed primarily in two areas— altering the distribution of minerals to increase their presence in the grain endosperm and increasing the overall levels of nutrients in the wheat grain. Here we will focus primarily on the latter approach.

1.3.2 A trade-off exists between yield and quality

Attempts to increase the levels of protein and nutrients in wheat grains have highlighted a trade-off between yield and quality. Breeding efforts which focussed on yield alone have resulted in a decrease in grain nutrient content over the past hundred years (Murphy et al. 2008). This negative correlation with yield has been observed for multiple micronutrients and for protein content (Pepe and Heiner 1975). Studies in wheat have suggested that the process underpinning this process is senescence, in which protein, starch, and nutrients are remobilised from dying leaves into the developing grain. Delayed senescence may reduce yield loss as a result of heat and drought by extending photosynthesis, thus increasing the levels of sugar reserves available during the grain filling stage (Schippers et al. 2015). In contrast, earlier senescence increases grain protein and nutrient levels often at the expense of yield (Schippers et al. 2015). This balancing act, illustrated in Figure 1-9, underlies the importance of studying senescence to increase grain quality and yield.

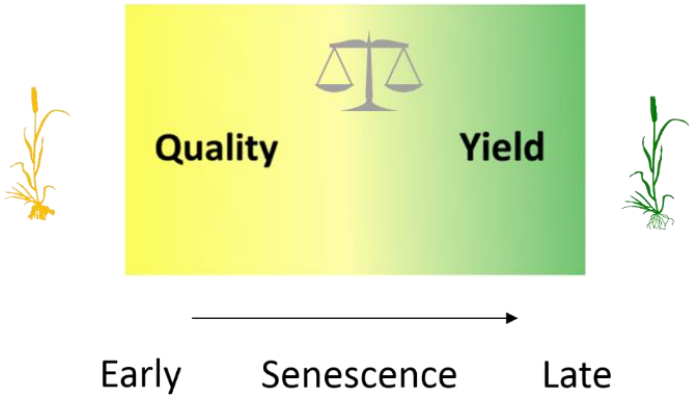

Figure 1-9: Senescence timing leads to a trade-off between grain quality and grain yield.

1.4 The many forms of senescence

Senescence begins
And middle-age ends
The day your descendants
Outnumber your friends

— Ogden Nash, *Crossing the Border*

The process of senescence is not straightforward to study. Even the definition of "senescence" differs between fields. In animals, where the concept of senescence was first introduced, it is typically defined as the developmental process that culminates in organism death (Medawar 1952, Crews 2003). Early on, three principal theories were suggested to explain the occurrence of senescence. The first, the mutation accumulation theory, was proposed by Sir Peter Medawar in 1952, arguing that aging occurs because mutations accumulate later in life as the force of selection due to reproduction declines (Medawar 1952). In contrast, the theory of antagonistic pleiotropy

proposed in 1957 by George C. Williams argued that it was the actions of genes themselves which drove senescence, as genes which were beneficial in early life, increasing fitness, lead to detrimental effects later in life (Williams 1957). Finally, the disposable soma theory was proposed in 1977 by Thomas Kirkwood, arguing that ageing occurs as a trade-off between expending resources in maintaining the reproductive germ line and the non-reproductive soma (Kirkwood 1977). These three theories were developed around the framework of animal senescence, and the extent to which they are applicable to plant senescence is still under debate (Salguero-Gómez et al. 2013).

Senescence was first discussed in the plant context by A. C. Leopold, considering it to be a "relatively gross change" that led to death of the whole plant (Leopold 1961). Yet he also recognised that senescence in plants includes both whole-organism death, such as that observed in annual plants, as well organ-specific death, such as that seen in leaves of deciduous plants during the autumn (Leopold 1980). This tension between the different scales at which senescence works is still present in the field today. In the following section, we will address the different levels at which senescence can be thought to act.

1.4.1 Monocarpic senescence

When we consider senescence in wheat breeding, it is principally in the context of *monocarpic senescence*. That is, the process by which the whole wheat plant undergoes coordinated death. During monocarpic senescence, which begins following reproduction, macromolecules in photosynthesizing tissues are broken down and the constituent parts are moved, or *remobilised*, into the developing grains (Schippers et al. 2015). This process is essential to lay down the required starch and nutrient stores in the grain, to sustain the seedling following germination (Distelfeld et al. 2014, Borrill et al. 2015b). It is this remobilisation of nutrients, protein, and starch that defines grain yield and quality parameters. Most of the remobilised nutrients stem from the flag leaf of each tiller, the topmost leaf before the wheat spike (Figure 1-10A). The remobilised nutrients move from the flag leaf through the leaf sheath, which encircles the stem, to the nearest node. At this point, the nutrients are transported into the phloem in which they are transported up the stem, through the peduncle, and into the spike and the developing grains. The breakdown of macromolecules leads to the loss of pigment in the tissues and chloroplast degradation, causing the characteristic yellowing observed in wheat plants approaching harvest (Schippers et al. 2015).

As senescence approaches its end, the cells present in the vegetative tissue lose their cellular integrity, beginning to leak electrolytes into the apoplast (Rolny et al. 2011). This stage of senescence is associated with cell death, and the complete loss of photosynthetic capacity. In wheat, monocarpic senescence occurs from the bottom of the plant to the top, with the older leaves at the base of the plant senescing first, and the flag leaf senescing last (Figure 1-10B). Whole-plant senescence is intrinsically linked with the life-history pattern of annual plants and other species

which demonstrate monocarpy, or a single reproductive episode which immediately precedes plant death (Munné-Bosch 2008). The extent to which iteroparous plants, which undergo multiple rounds of reproduction, demonstrate whole-plant senescence remains up for debate (Munné-Bosch 2008).

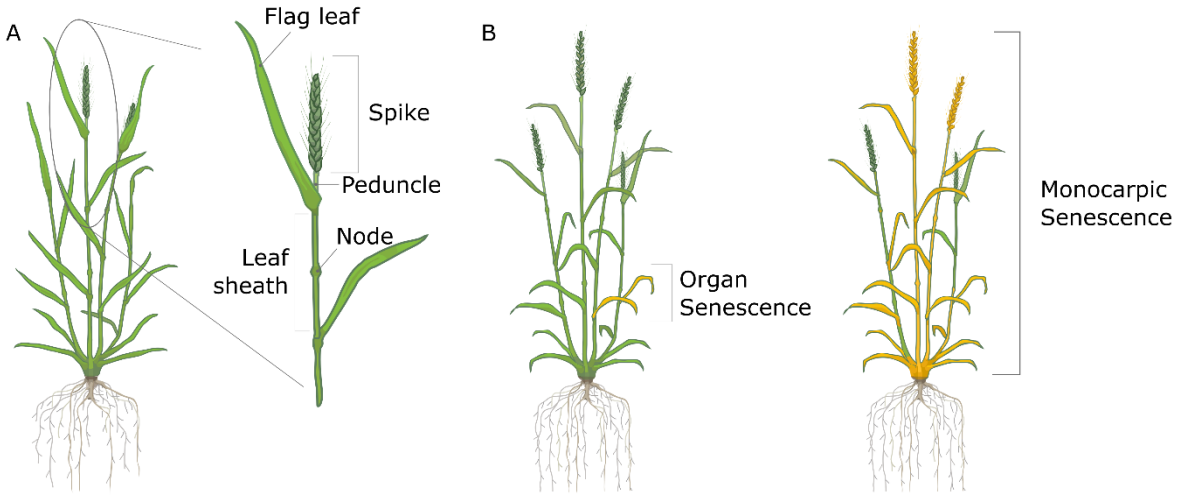

Figure 1-10: Wheat morphology and senescence. (A) Organs of wheat involved in senescence are highlighted. (B) Senescence can be separated into organ senescence of a subset of the plant, often due to environmental stresses, and monocarpic or developmental senescence, that occurs in a sequential manner across the entire plant. The plant depicted is still undergoing the process of senescence in the younger (outside) tillers, with sequential senescence of the leaves occurring from the base of the plant upwards. Illustrations adapted from Tobin Florio.

1.4.2 Organ senescence

The death of individual organs during development, whether due to environmental stresses or wounding, are also considered examples of senescence. This so-called “precocious” senescence can lead to leaf yellowing and death in a way that does not inherently lead to whole-plant mortality (Figure 1-10B). Instead, this organ-level senescence can act as a defence mechanism by which a small section of the plant is sacrificed to protect the health of the remaining plant (Schippers et al. 2015). Organ-level senescence is also observed commonly in perennial plants such as deciduous trees, which undergo a substantial annual die-back whereby individual leaves senesce and remobilise nutrients back into the trunk to aid survival during the winter season (Aerts 1996). This kind of discrete senescence is also observed in some perennial tuber plants, where the entire or near-entire above ground portion of the plant dies while the tuber persists in the soil before re-sprouting the following year (Munné-Bosch 2014). Organ senescence can also refer to the developmentally-regulated lifespan of a particular organ type, such as a flower or seed pod, which can be linked to external events such as fertilisation or seed development, respectively (Rogers 2015).

1.4.3 Programmed cell death

At an even smaller scale, the concept of senescence has also encompassed the process by which individual cells die. Highly coordinated, programmed cell death historically refers to a set

procession of steps which are required for the self-induced death of a cell. The extent to which programmed cell death (PCD) is a form of senescence is debated (van Doorn and Woltering 2004), but PCD itself forms an essential part of the final stage of organ and monocarpic senescence. Unlike PCD that occurs during earlier stages of development, senescence-associated PCD occurs across a wide range of developmentally heterogeneous cells (Rogers 2015). Autophagy-associated proteins have been found to be associated with the progression of whole-plant senescence induced by carbon-deprivation in *Arabidopsis*, highlighting the conservation between PCD at the single-cell level and the larger-scale programmes of organ and organismal senescence (van Doorn and Woltering 2004). It may be that the cellular process of programmed cell death can be considered part of the longer process of senescence itself, occurring as the final stage of a larger programme.

1.4.4 Population-wide senescence

Senescence also has implications beyond the individual plant level. At the population and species scale, the progression of senescence is studied in the context of life history trajectories. In these fields, organisms that show evidence of senescence follow a characteristic life history trajectory, whereby the proportion of individuals surviving past a given age rapidly decreases following a given point in time (Figure 1-11A, Type I) (Pearl and Miner 1935, Deevey 1947). This survivorship curve corresponds to both annual plants, which show a dramatic drop in survivorship to 0 at the end of a growing season, and species which show more variable times to mortality, such as humans. Other species, such as some long-lived trees, show high levels of mortality early in development which decreases at later stages (Figure 1-11A, Type III). In the context of populations, senescence is defined as an increase in the force of mortality following the onset of sexual maturity (Figure 1-11B). Early work on the progression of senescence in populations suggested that it was a universal characteristic (Hamilton 1966). However, this axiom has been questioned recently as more data has become available from a wider range of species, especially those with extremely long lifespans. Some species, particularly within the Plant Kingdom, can show signs of negligible or even negative senescence, defined as those for whom reproductive success does not decline, or even increases with age while death rates simultaneously decline following the start of reproduction (Figure 1-11B) (Vaupel et al. 2004). Such characteristics may be more common in species which begin reproduction before having attained their maximal size. As a result, further growth following the beginning of reproduction leads to both increased reproductive success and reduced likelihood of mortality.

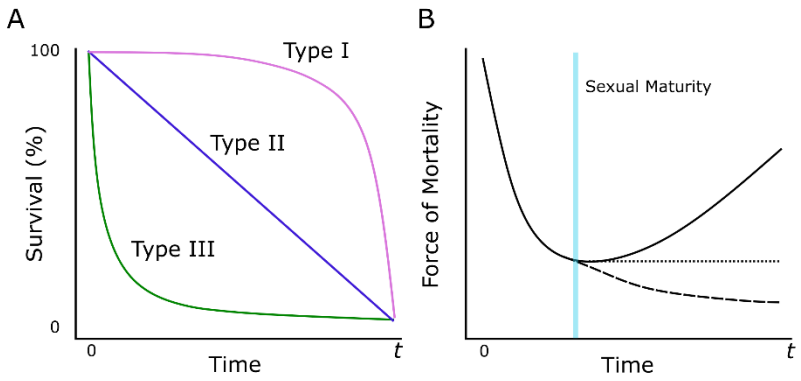

Figure 1-11: Life history patterns associate with senescence profiles. (A) Species follow characteristic life-history patterns, ranging from those species that show high levels of mortality late in life (I) to those with high levels of mortality earlier in life (III). (B) Some species show clear signatures of senescence, where the force of mortality increases following sexual maturity (solid line); others show negligible senescence (dotted line) where the force of mortality is constant following sexual maturity. Some species may exhibit negative senescence, where the force of mortality decreases following sexual maturity (dashed line). The point of sexual maturity is shown with a light blue vertical line.

In the remainder of this thesis, we will focus explicitly on *monocarpic, whole-plant senescence*. As detailed above, this is a highly regulated process that is intrinsically coupled with reproduction and is essential for the remobilisation of sugars, proteins, and nutrients into the developing grain. We will initially focus on research which has been primarily carried out in the model annual *Arabidopsis thaliana*, for which substantially more is understood regarding the regulation of monocarpic senescence. This will then be followed by a specific analysis of the known regulation of senescence in wheat.

1.5 The regulation of senescence depends on external and internal signals.

A variety of interconnected processes act to induce senescence following reproduction. The first organelles to be broken down during senescence are chloroplasts, which supply the bulk of the nutrients that are remobilised to the developing grains (Schippers et al. 2015). As senescence progresses the remaining organelles are broken down, ending with the dismantling of the nucleus and mitochondria. This process involves tightly regulated, interconnected pathways influenced by both environmental and internal signals. This tipping point can vary, but once reached the plant begins an irreversible process of senescence, irrespective of any further changes in environmental conditions. In order to ensure that senescence is triggered at precisely the correct time, these networks are precisely regulated by a range of different factors that will be discussed in the following sections.

1.5.1 Senescence is regulated by plant age.

Senescence can be understood in the context of three distinct phases of plant development (Schippers et al. 2007). The first phase consists of the early plant development, the “never

senescence" phase. Here, leaves act as sinks for nutrients and sugars and do not senescence in response to senescence-inducing factors. The second phase, "leaf maturation," involves the transition of the leaf from a net sink to a net source of sugars. At this stage, the leaf becomes competent to respond to senescence-inducing factors. Hypothetical "Age-Related Change" factors (ARCs) may accumulate in a plant as it ages, inducing senescence upon reaching a critical level even under optimal growth conditions (Jing et al. 2005, Schippers et al. 2007). However, throughout this stage, the leaves are still able to respond to senescence-delaying factors, such as cytokinin (Gan and Amasino 1995). Once the plant reaches the third stage, senescence itself, such anti-senescence signals are unable to halt or reverse the process.

Age-induced senescence is the final stage of plant development and is regulated in part by genes which act in many different processes throughout development. In Arabidopsis, two genes, the ubiquitin receptor *DA1* and the E3 ligase *Big Brother* (*BB*), regulate leaf size. The mutants act in an additive manner to increase leaf size, while also exhibiting delayed senescence (Vanhaeren et al. 2017). Correspondingly, ectopic expression of the two genes promotes the onset of senescence. Induction of *DA1* and *BB* downregulates mitosis and other cell-cycle related genes and upregulates senescence-associated genes. This suggests that *DA1* and *BB* play a role in regulating a developmental switch between cellular proliferation and cell expansion and eventual cell death, (Vanhaeren et al. 2017).

Sugars accumulate in leaves throughout plant development and may also act as age-related senescence inducing factors. In Arabidopsis, sugar accumulation in leaves has been found to promote the onset or the acceleration of senescence (Diaz et al. 2005, Wingler et al. 2009, Watanabe et al. 2013). This suggests a role for sugars as an internal ARC, inducing senescence at a critical level However, in barley, the onset and progression of senescence correlates with a decrease in glycolysis sugars in the leaves (Avila-Ospina et al. 2017). Other studies in Arabidopsis have also produced contradictory results on the impact of sugar accumulation on senescence, suggesting that this area is ripe for further study (van Doorn 2008).

1.5.2 Transcription factors interpret and integrate signals from plant hormones to induce or repress senescence.

The tight control of senescence onset and progression is governed by the action of various transcription factors which maintain the balance between senescence and non-senescence during plant development. Early in plant development, regulatory mechanisms would act to prevent the premature induction of senescence. Correspondingly, as the plant ages, transcription factors can act to promote the onset of the senescence process. These regulatory components are themselves regulated by both external factors, such as environmental stresses, and internal factors, such as plant hormones (Woo et al. 2019). Plant hormones regulate a range of different developmental and physiological processes, ranging from defence responses to leaf development (Rivas-San Vicente

and Plasencia 2011, Xie et al. 2014). Many, if not all, of the plant hormones have also been shown to have a role in regulating senescence (Figure 1-12). These plant hormones induce complex networks of gene and protein interactions, often involving crosstalk between different pathways, to regulate senescence as shown in in the simplified schematic in Figure 12.

The progression of senescence begins following the completion of flowering and reproduction, which in Arabidopsis occurs at around 21 days post-sowing (Breeze et al. 2011). Recent work has demonstrated that, before any symptoms of senescence are visible, a regulatory shift occurs whereby NAC transcription factors switch from primarily downregulating their target genes to inducing their expression (Kim et al. 2018a). Across a timecourse of gene expression taken during developmental senescence, levels of the hormone ABA increased significantly from baseline by 31 days after sowing (DAS), while jasmonic acid also peaked at 33 and 39 DAS (Breeze et al. 2011). Salicylic acid levels dropped to a minimum at 31 DAS, before rising to a peak in mature leaves (Breeze et al. 2011). The signalling pathways of these three hormones are activated during senescence, suggesting their role as positive regulators of senescence (Figure 1-12). Genes related to ABA and JA signalling were upregulated from 23 DAS in the Breeze *et al*. analysis, while those associated with ethylene signalling were upregulated from 33 DAS, as visual symptoms of senescence become visible. In contrast, the cytokinin signalling pathway was downregulated from 23 DAS, highlighting its role as a negative regulator of senescence (Figure 1-12). In the remainder of this section, we will highlight the roles of different transcription factors in regulating senescence in the context of specific plant hormones.

1.5.2.1 Responses to ABA

The Arabidopsis transcription factors *ABF2*, *ABF3*, and *ABF4* are positive regulators of leaf senescence, in particular through promoting chlorophyll degradation (Gao et al. 2016). The genes are part of the ABA-responsive element (ABRE) family of transcription factors, and correspondingly act by promoting an ABA-dependent process of chlorophyll degradation (Gao et al. 2016). In cotton, the orthologous *GhABF2* provides increased resistance to drought and salinity tolerance when overexpressed, demonstrating the potential for transferring knowledge on stress tolerance from model species into crops (Liang et al. 2016). The Myb transcription factor *MYBL* is also induced by ABA expression and is expressed in older leaves during senescence, causing the upregulation of senescence-associated genes (Zhang et al. 2010).

In a similar vein, the transcription factor *ABIG1* is suggested to respond to drought-induced increases in ABA levels to promote leaf senescence (Liu et al. 2016b). In rice, *OsNAC2*, a NAC transcription factor, acts as a positive regulator of senescence and increases ABA levels by upregulating ABA biosynthetic genes and downregulating its catabolism (Mao et al. 2017). High levels of ABA, interestingly, downregulate *OsNAC2* while lower levels upregulate its expression. This suggests that feedback repression involving *OsNAC2* may provide an upper cap on levels of ABA during senescence. These examples highlight the fact that responses to abiotic stresses such

as drought and high salinity function through pathways that overlap closely with those involved in senescence, which will be discussed in greater detail in a following section.

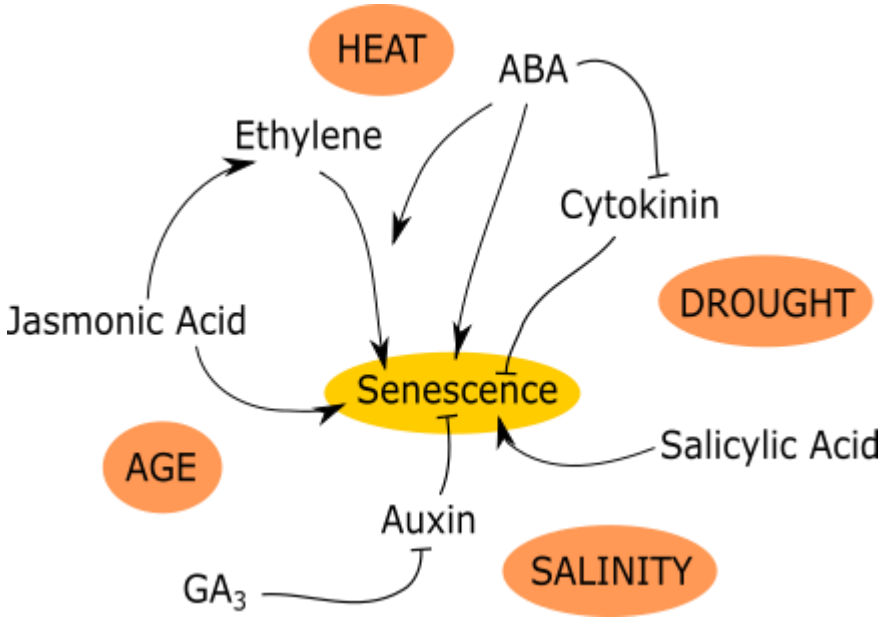

Figure 1-12: A simplified schematic of the roles of plant hormones in regulating senescence and the external factors that can influence senescence onset and progression (orange).

1.5.2.2 Responses to cytokinin

Cytokinin also plays a role in the regulation of leaf senescence, with enhanced cytokinin levels correlating with a delay in leaf senescence. This delay in leaf senescence is caused by preventing chloroplast breakdown and by increasing the sink strength of leaf tissue (Richmond and Lang 1957, Ehness and Roitsch 1997). The latter is due to cytokinin increasing the expression levels of the extracellular invertase *CIN1* and hexose transporters moving sucrose and glucose into the leaf tissue. Recently, analysis of gene expression data from a cytokinin-insensitive Arabidopsis mutant demonstrated that cytokinin inhibits leaf senescence in part by maintaining the integrity of the cellular and translational processes (Kim et al. 2018b). The transition of photosynthetic tissue, particularly the flag leaf in wheat, from a sink to a source of sugars and nutrients is thought to be critical in the initiation of senescence. This transition is highly regulated, and particularly in crop species is highly coordinated across the entire plant, if not an entire monoculture field (Schippers et al. 2015). Increased salinity stress is also thought to alter the source/sink relationship, with increased Na+ levels reducing sink strength and thus promoting premature plant senescence (Schippers et al. 2015). This importance contrasts in part with the observation that increased sugar content in Arabidopsis leaves promotes senescence, as discussed previously (Diaz et al. 2005, Wingler et al. 2009, Watanabe et al. 2013). If the latter were sufficient to drive senescence, then the increase in sugar content as a result of increased *CIN1* and hexose transporter expression should promote the onset of senescence, rather than repressing it. This, along with other contradictory results, suggests that sugar accumulation in leaves does not suffice to promote senescence (van Doorn 2008, Avila-Ospina et al. 2017).

An increase in cytokinin levels, as a result of overexpressing the isopentyl transferase (*ipt*) which controls cytokinin synthesis under the senescence-activated promoter of *Senescence Associated Gene 12* (*SAG12*), also leads to increased drought tolerance and delayed senescence in a range of species (Gan and Amasino 1995, Clark et al. 2004, Rivero et al. 2007, Merewitz et al. 2011). It was proposed that reduced leaf senescence in the *ipt* overexpression lines was a result of decreased protein degradation, particularly those involved in antioxidant activities and other major pathways (Merewitz et al. 2011).

The perception of hormones by specific receptors acts as a gateway into hormone-regulated pathways. In Arabidopsis, three histidine kinases located at the plasma membrane (*AHK2*, *AHK3* and *AHK4/CRE1*) are involved in cytokinin recognition (Hwang et al. 2012). While the action of these receptors typically filters through a cascade of downstream genes, in particular His phosphotransfer proteins (AHPs) and type-B response regulators (B-ARRs), the action of the receptors alone influence gene expression in plastids (Hwang et al. 2012, Danilova et al. 2017). Two of the receptors, *AHK2* and *AHK3*, have opposite effects on leaf senescence. *ahk2* mutants showed increased levels in chlorophyll content and decreased levels of *SAG12* expression compared to wildtype, while the opposite was true for the *ahk3* mutant (Danilova et al. 2017). Other work has also found that the overexpression of cytokinin response factors, themselves also promoted by cytokinin, promotes the premature onset of senescence (Raines et al. 2016).

The opposing influences of *AHK2* and *AHK3* appear at odds with the known role of cytokinin as an inhibitor of leaf senescence. However, the authors of the two studies in question posit that cytokinin-mediated signalling may form an incoherent Type I feedforward loop (Raines et al. 2016, Danilova et al. 2017). Such a regulatory loop leads to responses to stimuli that are proportionate to the fold change of the stimuli with respect to background noise (Goentoro et al. 2009). The fact that the response is relative to *fold-change* in a stimulus, rather than its absolute levels, could account for the apparently paradoxical induction of positive regulators of senescence by a negative regulator of senescence.

1.5.2.3 Responses to ethylene

In Arabidopsis, the ethylene response factor *EIN2* has been implicated in promoting leaf senescence through three distinct pathways, two through the induction of *EIN3*, another member of the ethylene response cascade, while also directly activating other TFs which promote senescence. In one case, *EIN3* directly promotes *AtNAP*, a NAC transcription factor which is a known positive regulator of senescence (Kim et al. 2014). *EIN3* also promotes *ORE1*, another NAC transcription factor which positively regulates senescence by inhibiting *GLK1* and *GLK2*, two G2-like transcription factors, leading to the disinhibition of Senescence Associated Genes (SAGs). In particular, *ORE1* is thought to form a heterodimer with *GLK1* and *GLK2*, possibly acting to sequester the *GLK* proteins and instead preferentially promote the activation of *ORE1* targets (Rauf et al. 2013). However, this pathway is made age-dependent through the action of miR164 which

represses *ORE1* early in development (Kim et al. 2009a). Later on, as *EIN3* increases, it is able to repress the action of miR164, both directly and indirectly promoting the activity of *ORE1* (Kim et al. 2009b, Li et al. 2013, Kim et al. 2014). *EIN3* and *ORE1* then act in concert to activate chlorophyll catabolic genes, such as *NYE1*, *NYC1*, and *PaO*, which induce chlorophyll breakdown (Qiu et al. 2015) while *ORE1* targets the nuclease *BFN1*, promoting nucleic acid degradation (Matallana-Ramirez et al. 2013).

ORE1 is also the site of crosstalk between the ethylene and ABA signalling pathways, as it is induced by both *EIN3* (Li et al. 2013) and the ABA-induced transcription factor *ATAF1* (Garapati et al. 2015). *ATAF1* concurrently activates *ORE1* expression and represses *GLK1* expression by binding to their respective promoters. This joint regulation underpins the ABA-regulated shift from a pre-senescent to actively senescing state (Garapati et al. 2015). Ethylene also acts in concert with another hormone strigolactone to promote synthesis, as while mutants in strigolactone biosynthesis exhibited delayed senescence, only application of strigolactone in concert with ethylene was sufficient to promote early-onset senescence (Ueda and Kusaba 2015). The positive regulation of senescence by ethylene is also implicated in salt-induced premature senescence in tomatoes (Ghanem et al. 2008).

1.5.2.4 Responses to jasmonates

Jasmonates (JA) are also implicated in senescence, regulating the expression of senescence related genes in multiple plant species, such as maize, rice and Arabidopsis (Qi et al. 2015). The induction of senescence by JA is also regulated by the miRNA miR319, a repressor of the TCP4 transcription factor which promotes JA biosynthesis in Arabidopsis and which also acts to upregulate senescence-associated genes such as the transcription factor *WRKY53* (Schommer et al. 2008). In the absence of JA, inhibitory JAZ proteins bind to a set of bHLH transcription factors and repress their activity. Upon synthesis of JA, however, the JA-responsive *COI1* recruits the JAZ proteins for degradation by the 26s proteasome, releasing the bHLH transcription factors. Intriguingly, these bHLH transcription factors are split between positive and negative regulators of senescence (Qi et al. 2015). Some, such as *MYC2*, activate senescence-associated genes like *SAG29*. Others, the IIId bHLH, repress this expression. As a result, the effect of JA on leaf senescence depends on a balance between the relative levels of senescence-promoting and senescence-inhibiting bHLH transcription factors (Qi et al. 2015). JA-induced senescence is also regulated by the transcription factor *WRKY57* which binds the promoters of and represses the transcription of *SENESCENCE4* and *SAG12*, inhibiting the progression of senescence (Jiang et al. 2014).

1.5.2.5 Responses to salicylic acid

The role of salicylic acid in plant stress responses is well characterised compared to the understanding of its role in plant development (Rivas-San Vicente and Plasencia 2011). However, salicylic acid (SA) also plays a role in promoting senescence. Levels of SA in Arabidopsis leaves increase up to 4X in senescing leaves, while mutants in SA signalling and biosynthesis, including

npr1 and *pad1*, exhibit delayed leaf senescence (Morris et al. 2000). SA-deficient mutants also show decreased expression levels of senescence-associated genes (SAGs) (Buchanan-Wollaston et al. 2005). More recently, the mitochondrial protease *FtSH4* in Arabidopsis was found to regulate leaf senescence via changes in SA levels (Zhang et al. 2017). The loss of *FtSH4* leads to lowered cardiolipin levels in the mitochondria, and thus promoting ROS production. The ROS then promotes expression of a cascade of WRKY TFs that themselves promote SA biosynthesis and thus senescence itself. Salicylic acid is also critical for controlled autophagy, as mutants in autophagy genes in Arabidopsis exhibit early senescence and increases in SA levels (Yoshimoto et al. 2009). Reactive oxygen species and SAGs are also upregulated, while mutations in the SA signalling pathway recovered the senescence phenotype indicating that SA alone was implicated in the premature senescence response.

1.5.2.6 Crosstalk between hormone signalling cascades is critical for senescence regulation

The discrete signalling pathways of plant hormones are connected via hub genes that facilitate crosstalk amongst hormone-response pathways. The transcription factor *AtNAP*, discussed above in the context of ethylene regulation, is also involved in feedback regulation of the ABA signalling pathway. *AtNAP* is promoted by ABA, in turn promoting the synthesis of ABA via activation of *AAO3* (Yang et al. 2014). This positive feed-forward loop may act to ensure the initiation of senescence cannot be reversed, as discussed above and in Schippers *et al*. 2015.

One of the key phenotypic characteristics of senescence onset is the yellowing of the photosynthetic tissue. This chlorosis corresponds to the breakdown of the chlorophyll pigments in the leaves as part of the nutrient remobilisation. As a result, regulating of chlorophyll catabolism is an essential part of senescence regulation. While many of the transcription factors previously discussed appear to bind and regulate further transcription factors, a large set of transcription factors have been identified in Arabidopsis which induce or repress the expression of these chlorophyll catabolic genes (CCGs). Three NAC transcription factors, ANAC046 (Oda-Yamamizo et al. 2016), ANAC016 (Kim et al. 2013, Sakuraba et al. 2016), and ANAC072 (Li et al. 2016), have all been shown to bind directly to the promoters of various subsets of CCGs, including *STAYGREEN1* (or *NYE1*) which destabilise chlorophyll binding proteins such as Lhca1 and the Lhcbs (Wu et al. 2016). The three NAC transcription factors influence senescence in response to a wide range of hormones, including JA, ABA, and ethylene. A set of MYC transcription factors, MYC2, 3, and 4, induce expression of these CCGs following treatment with Methyl Jasmonate. MYC2/3/4 are all repressed by the JA ZIM-domain repressor JAZ7 (Zhu et al. 2015), a known negative regulator of senescence, highlighting the multiple levels of transcriptional regulation involved in senescence. Indeed, MYC2 is thought to act synergistically with the NAC transcription factor ANAC019 to enhance *NYE2* expression, another CCG (Zhu et al. 2015).

The transcription factor *WRKY53* is involved in the senescence-regulatory pathways of JA, SA, and ethylene, inducing the expression of multiple downstream senescence-associated genes (Miao and

Zentgraf 2007, Rivas-San Vicente and Plasencia 2011, Xie et al. 2014). In the ethylene pathway, repression of WRKY53 by *Epithiospecifier Protein/Epithiospecifying Senescence Regulator* (*ESP/ESR*) is relieved by EIN2-mediated activation of *AtERF4* and *AtERF8* (Figure 1-13, p. 26). Meanwhile, JA promotes the activity of the gene *ESP/ESR* which physically interacts with and represses *WRKY53* (Miao and Zentgraf 2007). Levels of JA increase during senescence, controlled by TCP transcription factors (Schommer et al. 2008). This suggests that while WRKY53 levels and activity increase at the onset of senescence, as a result of induction by SA and de-repression by EIN2, as JA levels increase during the progression of senescence, WRKY53 levels decrease (Rivas-San Vicente and Plasencia 2011). This hints at a highly complex regulatory network between multiple plant hormone pathways, suggesting that there may be many more examples of highly interconnected regulation of senescence, and indeed other processes, between plant hormones.

1.5.3 Abiotic stresses promote senescence in plants

Senescence is not solely regulated by internal signals such as hormones and sugar levels; external signals such as high salinity, heat, and drought can influence the onset of senescence. These abiotic stresses are thought to induce premature senescence in order to ensure sufficient nutrients and sugars reach the seed before the plant is killed by the adverse conditions (Sade et al. 2017). The physiological characteristics of such senescence are similar to that seen as a result of plant aging, such as leaf yellowing and eventual necrosis. Many such examples of overlap between stress and senescence pathways were already discussed above in the context of hormonal control of senescence.

Network analyses of senescence-related transcription factors have highlighted substantial levels of overlap between hormone and stress regulation of transcription factors. For example, the ABA-responsive transcription factors *ABF3* and *ABF4* have been found to bind the promoter of *ANAC072*, a positive regulator of senescence, via a yeast one-hybrid assay (Hickman et al. 2013). Its promoter was also bound by Myb transcription factors known to be involved in jasmonic acid and ethylene signalling, as well as ABA responses. At the same time, a range of CBF proteins were able to bind its promoter, at a dehydration response element, suggesting that *ANAC072* is also regulated in response to a variety of stresses including cold and infection.

Drought-responsive senescence is necessary to remobilise nutrients from dying, stressed tissues to younger tissues, mitigating the detrimental effects of drought-induced stress (Munné-Bosch and Alegre 2004). Analyses of gene expression suggest that, while there is significant crosstalk between drought responses and senescence, distinct genes are upregulated during each process (Merewitz et al. 2016, Wehner et al. 2016, Zheng et al. 2016). Many of the genes involved in drought-related senescence are transcription factors. The Arabidopsis ethylene response factor *AtERF019* delays plant senescence and improves drought tolerance (Scarpeci et al. 2017). The

correspondence between delayed senescence and increased drought tolerance has been previously characterised (Borrell et al. 2000, Kumari et al. 2007). Conversely, overexpression of the rice NAC transcription factor *OsNAC2* led to early leaf senescence and increased drought stress sensitivity (Shen et al. 2017).

In salt stress, similarly, transcription factors are also involved in promoting senescence. In particular, *AtNAP* represses the expression of *AREB1*, a stress-response gene which is ABA-dependent (Seok et al. 2017). *AtNAP* was originally identified as a positive regulator of senescence; knockdown of *AtNAP* via T-DNA insertion causes a delay in senescence (Guo and Gan 2006). This illustrates that while responses to abiotic stress may use specific genetic machinery to induce senescence, in other cases genes involved in senescence in the absence of stress are also recruited by the stress response. As a result, the delineation between "internal" and "external" factors, as discussed in Figure 12, is not so clear-cut. While such factors may independently act on the senescence regulatory network, the response to these factors is often overlapping.

The level and type of light received by a plant has been suggested to be the most important external signal when regulating senescence (Woo et al. 2019). Dark-induced senescence is a well-known process that has been found to be regulated by a variety of transcription factors. *WRKY22*, a target of the positive senescence regulator *WRKY53*, is induced in darkness and its overexpression leads to accelerated dark-induced senescence (Zhou et al. 2011). The JA-responsive ZIM-domain repressor *JAZ7* also acts in dark-induced senescence (Yu et al. 2015), suppressing the induction of senescence by repressing genes such as the bHLH transcription factors *MYC2*, *3*, and *4* which promote chlorophyll breakdown (Zhu et al. 2015). Some transcription factors that regulate monocarpic senescence also regulate dark-induced senescence, such as *ATAF1* which is a positive regulator of both processes (Garapati et al. 2015). Light perception is critical in the induction of dark-induced senescence, and correspondingly the phytochrome interacting factors (PIFs) have been shown to control both age- and dark-induced senescence in Arabidopsis. *PIF3*, *PIF4* and *PIF5* act to induce multiple senescence-associated genes following a dark treatment, such as those involved in chloroplast degradation (Sakuraba et al. 2014a, Song et al. 2014, Zhang et al. 2015b). Mutations in these genes have also led to delayed developmental senescence (Song et al. 2014). The phytochrome molecules themselves have also been shown to influence senescence, with PhyA and PhyB acting antagonistically to control far red light-induced senescence (Lim et al. 2018).

Regulatory pathways responsive to both age and stress-induced senescence have been identified in different plant species. The CONSTANS-like gene *Ghd2* was found to be down-regulated under drought conditions in rice (Liu et al. 2016a). Overexpression of *Ghd2* reduced drought resistance and accelerated developmental leaf senescence. The converse was found in the *ghd2* mutant, illustrating its role in both age and stress related senescence. In barley, knockdown of the WHIRLY1 protein, which binds the promoters of senescence-associated genes, delayed senescence particularly in high-light contexts (Janack et al. 2016, Kucharewicz et al. 2017). This corresponded

with decreased expression levels of stress and senescence- associated WRKY and NAC transcription factors (Janack et al. 2016). In particular, WHIRLY1 also binds and represses the expression of *WRKY53*, a positive regulator of senescence (Miao et al. 2013). Intriguingly, the knockdown of WHIRLY1 resulted in a significant decrease in the level of H3K9 acetylation across the promoter of *HvS40*, a known senescence-associated gene. This suggests that the progression of senescence is influenced in part by chromatin marks, highlighting a potential role for chromatin-mediated regulation of senescence which remains under-explored. Perhaps such epigenetic control may provide a mechanism to explain the transition between the three stages of plant senescence discussed above, particularly between reversible and irreversible induction of senescence. Similar control mechanisms are essential in governing the switch between the vegetative and flowering state in plants, where changes in chromatin marks at *FLC* alter its expression and thus ability to flower (Hepworth and Dean 2015).

1.5.4 Chromatin changes are hallmarks of senescence

Further work has highlighted a plethora of examples where changes in chromatin modifications are associated with the regulation of senescence. Two histone methylation marks, H3K4me3 and H3K27me3 show inverse relationships with senescence associated genes (Brusslan et al. 2012). The former is a mark of open, active chromatin and is associated with upregulation of genes, while the latter marks closed chromatin and inactive genes. Correspondingly, genes containing the H3K4me3 mark in ageing leaves are more likely to be upregulated during senescence than those containing H3K27me3. The impact of other chromatin marks may be more specific to senescence itself. The Arabidopsis histone methyltransferase *SUVH2* leads to delayed senescence when over-expressed, in part due to its repression of *WRKY53* expression through inhibition of H3K4 methylation (Figure 1-13, p.26) (Ay et al. 2009). In contrast, loss-of-function of the Arabidopsis histone de-acetylase *HDA6* leads to delayed leaf senescence and downregulation of some senescence associated genes (Wu et al. 2008). However, this delay in leaf senescence co-occurs with a delay in flowering, suggesting that perhaps aspects of development in general are delayed without functioning *HDA6*, rather than senescence itself.

The sister protein of *HDA6*, *HDA9*, acts in complex with senescence-regulating transcription factors to bind the promoters of negative senescence regulators (Chen et al. 2016). It forms a complex with the POWERDRESS protein and WRKY53 and acts to inhibit negative regulators of senescence by removing the H3K9ac and H3K27ac marks and suppressing gene expression of senescence associated genes, such as the negative senescence regulator *WRKY57* (Figure 1-13, p.26) (Chen et al. 2016). Chromatin organisation can also be altered by the AT-hook DNA binding protein *ORE7* (Lim et al. 2007b). Over-expression of this gene leads to delayed leaf senescence and altered chromatin organisation during interphase. This change in chromatin organisation is associated with a shift in gene expression patterns towards that of a younger leaf, suggesting that overall chromatin architecture may also influence the progression of senescence. This is supported

by the fact that mutations in the *DRD1* and *DDM1* chromatin remodelling genes leads to delayed senescence (Cho et al. 2016).

1.5.5 Post-translational regulation of senescence

The progression of senescence can be further regulated through post-translational modifications. A characteristic example is that of *MAP Kinase 6* (*MPK6*) and *MAP Kinase Kinase 9* (*MKK9*) (Zhou et al. 2009), which are both involved in the *EIN2-EIN3-ORE9* cascade (Zhang et al. 2015a). Following perception of ethylene, which de-represses *MKK9*, *MKK9* phosphorylates and activates *MPK6*, which then phosphorylates and stabilises *EIN3* (Yoo et al. 2008). *EIN3* then goes on to induce multiple positive regulators of senescence, as discussed above, corresponding with the observed mutant phenotypes of *MKK9* and *MPK6* which lead to delayed senescence (Zhou et al. 2009). Other MAP kinase cascades regulate senescence in response to the presence of ABA (Matsuoka et al. 2015). Similarly, the MAP kinase kinase kinase *MEKK1* phosphorylates and stabilises *WRKY53* which is a critical positive regulator of senescence (Figure 1-13, p.26) (Miao et al. 2007). Intriguingly, it can also bind directly to the promoter of *WRKY53*, suggesting that it can act both directly to stabilise *WRKY53* and indirectly to induce *WRKY53* transcription during the critical transition period when *WRKY53* expression levels transition from being dependent on leaf age to being dependent on whole plant age (Miao et al. 2007).

Other kinase proteins also play a role in regulating senescence. The receptor protein kinase *RPK1* is induced by ABA and is required for timely age- and ABA-dependent senescence (Lee et al. 2011). As part of this response, Receptor protein kinase 1 (RPK1) is involved in a positive-feedback regulation loop in a complex with Calmodulin 1 (CaM1) and RbohF (Dai et al. 2018). The soybean receptor kinase *GmSARK* is also a positive regulator of leaf senescence, via auxin and ethylene-mediated pathways (Xu et al. 2011). The Arabidopsis ortholog *AtSARK* also showed early senescence when over-expressed, suggesting that the pathway is conserved between the two species (Xu et al. 2011). SARK itself is dephophorylated by the protein phosphatase SSPP, leading to its inactivation and thus negatively regulating leaf senescence (Xiao et al. 2015). Over-expression of *SSPP* is able to rescue the effects of over-expressing *SARK*.

Protein levels can also be controlled by regulated protein degradation, often in a ubiquitin-ligase mediated manner. *WRKY53* may be regulated by the HECT domain-containing E3 ubiquitin protein ligase *UPL5* (Miao and Zentgraf 2007). *UPL5* can bind directly to the WRKY53 protein and carry out poly-ubiquitination *in vitro* (Figure 1-13, p.26). As a mutant in *UPL5* phenocopies that of a *WRKY53* over-expresser, it's likely that *UPL5* targets WRKY53 for degradation via the ubiquitination pathway. In contrast, the RING-type ubiquitin ligase *ATL31* is directly activated by WRKY53 (Aoyama et al. 2014). Acting as a negative regulator of senescence, *ATL31* controls the induction of senescence in response to low CO_2 and nitrogen, preventing the rapid induction of premature senescence in response to nutrient and carbon deprivation. Protein turnover also

regulates the induction of senescence by ABA, whereby the U-box-armadillo E3 ubiquitin protein ligase *SAUL1* is thought to target the aldehyde oxygenase *AAO3* for degradation (Raab et al. 2009). *AAO3* is essential for the conversion of abscisic aldehyde to ABA, and its degradation prevents the accumulation of ABA that induces senescence. Correspondingly, mutations in *SAUL1* lead to early senescence and increased ABA biosynthesis (Raab et al. 2009) alongside upregulation of common senescence-associated genes such as *ORE1* and *WRKY53* (Vogelmann et al. 2012). The F-box protein ORE9 is also involved in mediating senescence onset through regulation of protein degradation (Woo et al. 2001). ORE9 forms part of a SCF E3 ligase complex, SCF^{ORE9}, which is thought to target negative regulators of senescence for degradation as mutants in *ORE9* lead to delayed initiation of leaf senescence. The plant U-box E3 ligases PUB12 and PUB13 are involved specifically in negatively regulating stress-responsive senescence, though the exact mechanism by which this is carried out is unknown (Zhou et al. 2015).

Ultimately, senescence is controlled through a combination of regulation at the chromatin, transcription, and post-translational levels, as demonstrated in Figure 1-13. In this example, *WRKY53* is both controlled by chromatin modifications, and controls the expression of other targets through removal of modifications. It acts as a transcription factor to promote the expression of multiple senescence-associated genes, and is itself post-translationally regulated, influencing the progression of senescence. The case of WRKY53 exemplifies the importance of integrating all three modes of regulation in our understanding of the regulation of senescence.

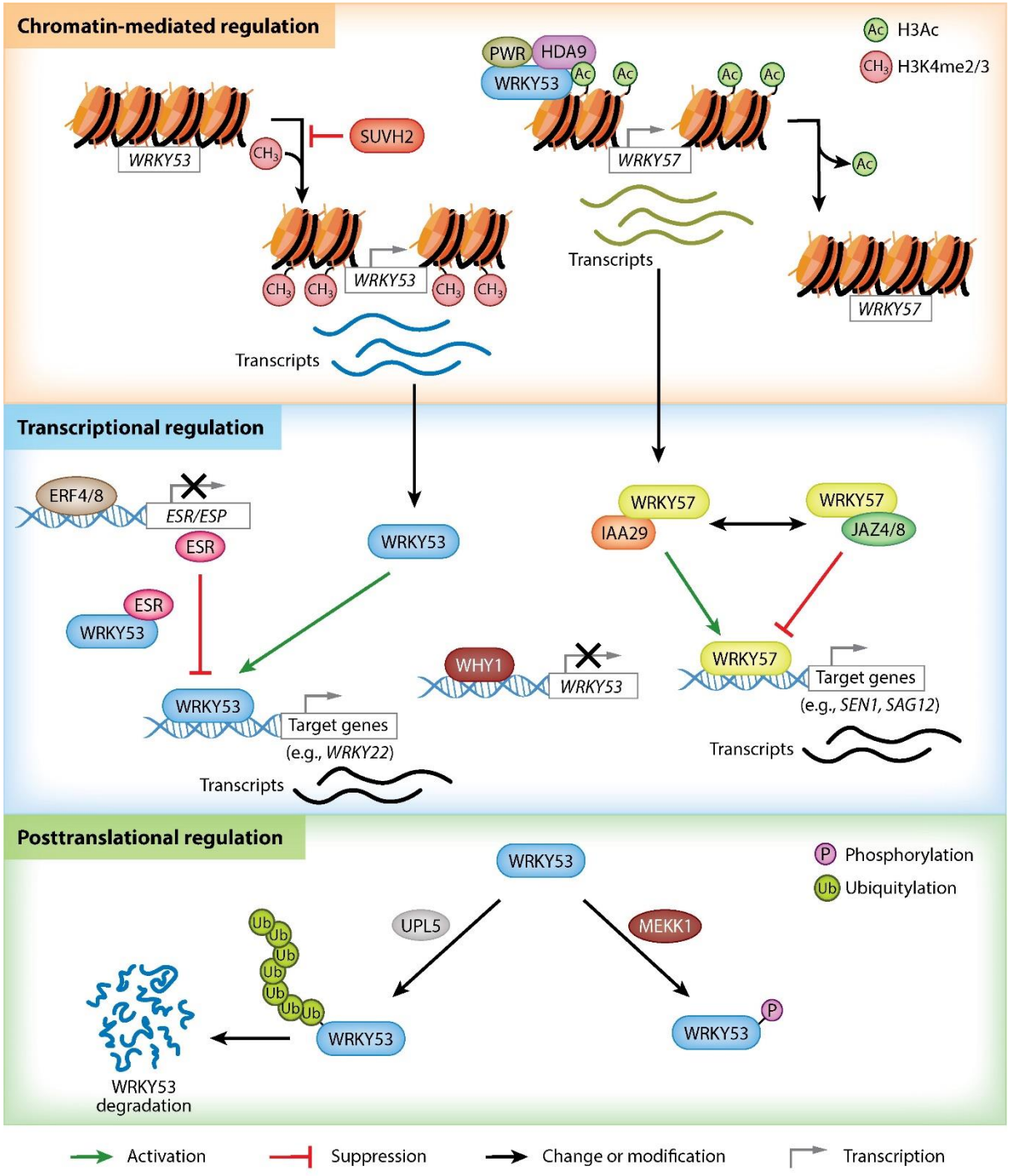

Figure 1-13: Multiple levels of regulation act on plant senescence. Depicted here in the case of the positive senescence regulator WRKY53. This figure has been reproduced from Woo et al. (2019), under license ID 4672500334507.

1.5.6 Senescence in wheat

On a large-scale level, the roles of plant hormones in promoting senescence in wheat are thought to be conserved with those in Arabidopsis and other more studied plants. ABA is upregulated in response to drought in wheat, while cytokinin levels are decreased, promoting leaf senescence (Yang et al. 2003). Additionally, application of cytokinin to wheat plants post-anthesis can increase

grain yield and heat tolerance (Yang et al. 2016). WRKY TFs in wheat with homology to *WRKY53*, discussed above as a critical regulator of senescence and cross-talk between hormone pathways, have been identified, suggesting that some may play a similar role in wheat (Niu et al. 2012). However, this principally illustrates that very little is understood regarding the particular genes involved in senescence in wheat. It is not sufficient to claim that wheat homologs to Arabidopsis genes are likely to play the same role in wheat without experimental evidence. Indeed, comparison of a timecourse of gene expression during leaf senescence in Arabidopsis (Breeze et al. 2011) and wheat (Borrill et al. 2019a) found that few wheat orthologs of the senescence-associated genes in the Arabidopsis timecourse showed equivalent expression profiles.

1.5.6.1 The role of NAC transcription factors in wheat senescence

This same timecourse from Borrill et al. found that, following 10 days after anthesis, the NAC transcription factor family, named after the first three members identified, *NAM, ATAF 1* and *2*, and *CUC2*, was significantly upregulated in flag leaf tissue (Borrill et al. 2019a). This was the only transcription factor family which was significantly upregulated in the latter half of the timecourse, suggesting that NAC transcription factors in particular may be key regulators of senescence in wheat. This hypothesis is supported by the specific NAC transcription factors with roles in senescence in wheat that have already been identified. *NAC-S* is a negative regulator of senescence, which when overexpressed delays senescence and increases Rubisco expression (Zhao et al. 2015). However, based on the expression profiles observed in Borrill et al. (2019a), it is possible that many NAC transcription factors are positive regulators of senescence and are correspondingly upregulated during senescence.

1.5.6.1.1 The NAM transcription factors are positive regulators of senescence

The wheat *NAM* genes, *NAM-1* and *NAM-2*, are NAC transcription factors that are positive regulators of senescence in wheat (Uauy et al. 2006b, Avni et al. 2014, Pearce et al. 2014, Borrill et al. 2019a). *NAM-B1* was initially identified as the gene underlying the *Grain Protein Content* (*GPC*) locus on chromosome 6B which originated from an accession of tetraploid wild emmer wheat (*T. turgidum* ssp. *dicoccoides*, accession FA15-3) (Olmos et al. 2003, Uauy et al. 2006a). *NAM-B1* functional copies are only found in wild wheat relatives, as the gene has either been lost or made non-functional by a single missense mutation in cultivated wheat (Uauy et al. 2006b). Introgression of the functional *NAM-B1* into elite tetraploid and hexaploid wheat backgrounds (cv. Langdon, Anza, RS15, and UC1041) resulted in consistently higher grain protein content and premature leaf and peduncle senescence (Uauy et al. 2006a). The *NAM-B1* gene forms part of a homoeologous triad located on chromosome 6 (*NAM-1*) and is paralogous to a second triad located on chromosome 2 (*NAM-2*).

1.5.6.1.2 Characterisation of the NAM genes

Initial characterisation found that knockdown of the NAM gene family by RNAi in hexaploid wheat (cv. Bobwhite) was sufficient to delay leaf senescence by 24 days compared to the control

(Uauy et al. 2006b). This corresponded with a decrease in GPC, Zn, and Fe levels in the grain. The RNAi knockdown targeted all NAM genes except *NAM-A2*, which had not been identified at the time of publication. Later work characterised the function of the individual NAM homoeologs and paralogs. The A and D-genome homoeologs of *NAM-B1* were individually characterised using premature stop-codon mutations in hexaploid wheat (cv. Express) (Avni et al. 2014). Single mutants in both homoeologs were sufficient to induce a delay in flag leaf senescence of approximately five days, with the *nam-a1* mutant retaining higher levels of chlorophyll in the flag leaf. Correspondingly, the single mutants led to reduced grain protein levels across different environments, though this decrease was not always significant compared to wild-type. The double mutants, *nam-a1/nam-d1*, were significantly more delayed in flag leaf senescence, by approximately 15 days compared to wild-type. There was a similarly more severe reduction in grain protein content in the double mutants, which was significant across all environments tested.

The impact of *NAM-A1* on senescence and grain protein and nutrient content was validated in tetraploid wheat, using a different premature truncation variant (Pearce et al. 2014). In this study, the paralogous *NAM-B2* gene was also characterised. The *nam-b2* mutant did lead to a slight delay in flag leaf senescence, but this was much smaller than that seen for the *nam-a1* mutant. Notably, there was no significant difference in senescence timing between the *nam-a1* single mutant and the *nam-a1/nam-b2* double mutant. Similar results were seen for grain protein content and grain zinc and iron content. This suggested that the NAM-1 genes are stronger regulators of senescence than the NAM-2 genes. More recently, mutations in the *NAM-A2* paralog were characterised using the Kronos TILLING population (Borrill et al. 2019a). In this case, no premature truncation mutants were present in *NAM-A2*. Instead, missense mutations in three highly conserved residues were selected and phenotyped for senescence. While the *nam-a2* single mutants led to no significant delay in flag leaf senescence, there was a significant delay in peduncle senescence of approximately 4 days.

Given the strong effect of the NAM genes on monocarpic senescence and grain protein and nutrient content, the role of their orthologues in rice and Arabidopsis were also investigated. The closest orthologs to the NAM genes in Arabidopsis are *AtNAM*, *AtNAC2*, and *AtNAC025*. None of these genes have been shown to have any role in regulating senescence. They instead act in shoot apical meristem development (Duval et al. 2002), responding to salt stress and lateral root development (He et al. 2005), and seed development (Kunieda et al. 2008), respectively. In rice, the closest ortholog is *OsNAC010* which is orthologous to the NAM-2 triad, located on chromosome 2. Knockdown of *OsNAC010* using RNAi had no effect on monocarpic senescence but did lead to plant sterility due to a failure of anther dehiscence (Distelfeld et al. 2012). This sterility effect was not observed in mutants in the wheat NAM genes indicating a lack of functional conservation.

1.5.6.1.3 The impact of the NAM genes on agronomic performance

Many of these mutant phenotypes recapitulate the relationship between delayed senescence and decreased grain quality that has been identified in other studies (Schippers et al. 2015). Following the cloning of *NAM-B1*, the functional copy from *T. turgidum* ssp. *dicoccoides* has been introgressed into multiple cultivars in an effort to increase grain protein content. In the majority of genotypes tested, introduction of the functional *NAM-B1* gene led to increased grain protein and nutrient content (between 90 and 100% of studies in both hexaploid and tetraploid wheat) (Tabbita et al. 2017). The effects of the *NAM-B1* gene on yield components was less clear cut. While fewer tetraploid genotypes had been tested, they were more consistent in showing no significant variation in yield or organ number, while consistently exhibiting accelerated senescence, consistent with the positive regulatory role of *NAM-B1*. In contrast, the hexaploid wheat cultivars tested showed a wide range of yield phenotypes, and indeed showed no consistent effect on senescence timing. This unexpected variation in senescence phenotypes may be due in part to variable methods of scoring senescence, a theme which will be highlighted in later chapters. The variability in phenotypes also highlighted the impact of gene by environment and gene by genotype interactions on phenotype. This is perhaps to be expected given the known impact of environmental conditions on senescence timing and progression as well as the highly complex network of regulation governing senescence, as introduced previously.

These results suggest that the presence of *NAM-B1* consistently increases grain protein and nutrient content, however the mechanism by which this occurs was unclear. Using the RNAi NAM knock-down lines introduced above, the levels of N, Fe, Zn, and other nutrients were measured across senescence in various different tissues (Waters et al. 2009). This work demonstrated that the reduction in protein and nutrient content in the RNAi lines was due to reduced translocation capacity, such that the vegetative tissues retain the nutrients rather than remobilising them into the developing grain.

Experiments with mutants in the NAM genes also consistently found no correlation between delayed senescence and increased yield or thousand kernel weight, suggesting that increasing the length of photosynthesis does not always correlate with increased grain filling (Uauy et al. 2006b, Avni et al. 2014, Pearce et al. 2014, Borrill et al. 2015b, Tabbita et al. 2017). This result has been seen in other contexts of wheat senescence. The ectopic expression of the cytokinin-synthesis gene *ipt* from *Agrobacterium tumefaciens* under the SAG12 promoter demonstrated the conservation of cytokinin's role as a negative regulator of senescence in wheat (Sykorova et al. 2008). Increases in cytokinin content as a result of the *ipt* expression corresponded with a delay in chlorophyll degradation, and thus leaf senescence. This delay in senescence also did not correspond with an increase in grain yield. The work with the NAM genes was continued in a further study with the RNAi knockdown lines with delayed senescence (Borrill et al. 2015b). The activity of starch synthase was found to be limiting grain filling capacity, with excess fructan levels accumulating in

the stem rather than being imported into the grain. This highlights that, while in theory extended photosynthesis via delayed senescence could increase yield, breeding for delayed senescence should be combined with traits to increase grain filling capacity.

It should be noted that other studies have found a negative correlation between increased grain quality and grain yield. In this context, the premature senescence associated with increasing nutrient and protein content in the grain has also been found to be associated with decreased yields (Gregersen 2011). It's unclear how premature senescence leads to increased grain quality. Two theories have been proposed, one suggesting that earlier senescence allows more complete nutrient remobilisation from leaves to the grain while the other suggests that delayed senescence reduces the relative content of nutrients and protein by increasing grain weight (Gregersen 2011). However, the second theory is not supported by research showing that delayed senescence, at least in wheat, does not increase grain size but does show decreased grain quality (Uauy et al. 2006b). Thus, it seems most likely that premature senescence allows for the more efficient and complete remobilisation of nutrients to developing grains.

1.6 How can we study complex traits like senescence in wheat?

We know far less about how senescence is controlled in wheat than we do in model species like Arabidopsis which have historically been far more tractable systems for carrying out in-depth molecular characterisations. However, we have shown that Arabidopsis is not an adequate model for all modes of senescence regulation in wheat. As a result, if we aim to untangle the complex regulatory network governing senescence in wheat, we must begin to work in wheat itself. Over the past few years, a large array of genetic and germplasm resources have been established in wheat, reviewed in detail in Adamski et al. (2018). Here we will highlight key examples of new resources that aid the study of complex traits such as senescence in wheat.

1.6.1 Sources of phenotypic and genotypic variation

1.6.1.1 Natural variation

Many traits in wheat have historically been studied through phenotypic evaluation of bi-parental populations containing novel variation and the use of SNP markers to identify quantitative trait loci (QTLs) which associate with that phenotype. Over the past few years, a large set of populations have been made available which have introduced novel sources of variation that can be screened for previously unobserved phenotypes. These populations have included sources of natural variation derived from landraces, such as the Watkins collection, previously described in section 1.2.1, and the GEDIFLUX collection of European winter wheat varieties grown between 1945 and 2000 (Wingen et al. 2014). These lines have also been crossed with the elite wheat variety Paragon to generate mapping populations (Wingen et al. 2017) and have been genotyped using the 35k SNP array (Allen et al. 2017). A subset of the Watkins collection has also been sequenced using exome capture (Gardiner et al. 2018), providing more insight into the variation observed between the

different landraces and with elite lines today. Similar collections of natural variation have been established for other regional collections. These collections have also been established for tetraploid wheat, including the Global Durum Wheat panel consisting of modern diversity (Wang et al. 2014), and the Tetraploid wheat Global Collection consisting of landraces and cultivars involved in historical cultivation and breeding (Wang et al. 2014, Maccaferri et al. 2019). In most cases, these populations have been genotyped to an extent using SNP arrays or other similar methods, facilitating their use for analyses such as GWAS.

Other sources of natural variation that are accessible for forward genetic screens are artificially generated Multiparent Advanced Generation Inter-Cross (MAGIC) populations (Huang et al. 2015a, Adamski et al. 2018, Cockram and Mackay 2018). These populations are derived from multiple generations of inter-crossing initially between a set of founder parents and later between their progeny to generate a highly recombined panel containing a defined set of allelic variation. Such populations have been generated for many different crops, and within wheat using multiple different parental sets, and have proven their utility in mapping novel alleles of agronomically relevant traits (e.g. see Boden et al. (2015), Dixon et al. (2018), Sannemann et al. (2018)). As with the panels of natural variation, many of these MAGIC populations have been supplemented with genotyping data from SNP chips which facilitates genetic mapping of phenotypes.

Variation found within landraces and wheat cultivars can also be supplemented with the substantial variation retained in the wild grasses, both wheat progenitors and related species. Increasing work has been carried out to characterise the grass accessions available in gene banks, such as the Seeds of Discovery initiative at CIMMYT which is working to genotype tens of thousands of wild wheat accessions using DArTseq (Adamski et al. 2018). On a smaller scale, introgression of genome sequence from wild grasses into modern wheat varieties can allow the testing of these novel alleles within a domesticated background (Adamski et al. 2018). New advances in genotyping have facilitated the tracking and characterisation of these introgressed regions (Grewal et al. 2019).

1.6.1.2 Induced variation

Alongside the allelic variation available through wheat landraces, cultivars, and wild relatives, completely novel variation can be generated through induced mutagenesis. This mutagenesis can be carried out using chemical means, such as the mutagen ethyl methanesulphonate (EMS) which typically causes single base-pair mutations, or through irradiation of a population by, for instance, γ radiation. A γ-irradiated population has been developed in the winter wheat cultivar Paragon, which is currently undergoing skim sequencing to characterise the presence of large deletions or insertions in the lines (Uauy, personal communication). The wheat TILLING population, developed in both hexaploid (cv. Cadenza) and tetraploid (cv. Kronos) wheat, utilised EMS mutagenesis to develop a set of 1,200 and 1,535 mutated lines, respectively (Uauy et al. 2009, Krasileva et al. 2017). Uniquely, these lines were exome-captured and sequenced to identify the induced SNPs present in the coding regions of the genomes. SNP calling confirmed a uniform SNP

density across the coding regions in both cultivars, with 23 mutations per Mb present in Kronos individuals and 33 mutations per Mb present in Cadenza individuals. Analysis of the hexaploid and tetraploid lines found that, of the genes sequenced using the exome capture, 96% and 94%, respectively, contained at least one missense mutation in the population while 59% and 61% contained at least one premature truncation variant. The TILLING lines can now be screened for novel phenotypes caused by the induced mutations, and the causal mutation can then be mapped using markers designed against the SNPs present in the line of interest.

1.6.2 The wheat (pan)genome

Yet while the vast array of genotypic diversity available as a result of natural and induced variation has facilitated the identification of novel alleles and traits in wheat, any forward genetic screen in wheat is still hindered by its polyploid genome. As discussed earlier in this chapter, wheat is derived from two or three hybridisation events, for tetraploid and hexaploid wheat respectively, which has resulted in most genes consisting of two or three copies in the genome (Borrill et al. 2015a). These homoeologs often act in a functionally redundant manner, such that mutations in all two or three copies are required for a clear mutant phenotype to be observed (Borrill et al. 2019b). As a result, any forward genetic screen will only identify novel phenotypes caused by either dominant mutations in genes, or by mutations in genes where the homoeologs are not completely functionally redundant. Instead, this so-called "hidden variation" can only be identified through targeted studies of specific genes—that is, through reverse genetics approaches (Uauy et al. 2017).

Until recently, attempts to carry out reverse genetics studies in wheat were hindered by the lack of a high-quality genomic reference sequence for wheat. The development of such a resource was made more difficult by the polyploid genetic structure of wheat. At approximately 17 Gb (Uauy 2017), the sheer size of the wheat genome, 35 X that of rice, hindered efforts to develop a full reference sequence. The large percentage of repetitive DNA sequences in wheat, approaching 80%, and the high degree of conservation between homoeologous chromosomes, 95-99% in coding regions, further increased the sequencing difficulty (Borrill et al. 2015a). Despite these difficulties, since the mid-2000s efforts began to develop increasingly complete wheat genome sequences.

1.6.2.1 Developing the Chinese Spring reference genome

Early sequencing attempts were hampered by small scaffolds and significant regions of poor or no quality sequencing (Uauy 2017). These initial efforts used Bacterial Artificial Chromosome libraries to develop physical maps of individual chromosomes (Paux et al. 2008), which, while a significant step forward, was nevertheless time-consuming and difficult to carry out for the complete 21 chromosome complement of bread wheat. Next-generation sequencing techniques were also used to provide increasingly complete bread wheat sequences, including the Chromosome Survey Sequence (CSS) genome developed by the International Wheat Genome Sequencing Consortium (IWGSC) which functioned as the wheat reference genome until 2017

(Table 1-1) (IWGSC 2014). Nevertheless, the wheat genome remained incomplete and fragmented, hindering genome-wide analyses.

In 2017, two more complete assemblies were released, which each made substantial gains in the physical contiguity of the genome. The first, released by a team working at TGAC, what is now the Earlham Institute, utilized the open-source w2rap package (Clavijo et al. 2017a) alongside Illumina sequencing to assemble the wheat genome with a mean scaffold size of 88.7 Kb, an over 10-fold increase on the original CSS assembly (Table 1-1) (Clavijo et al. 2017b). This assembly also substantially improved the quality of the gene annotations through its use of strand-specific RNA-Seq data and long PacBio reads of cDNAs. In the same year, a second assembly of the wheat genome was produced which generated the most complete assembly to date (Zimin et al. 2017). This approach combined short-read Illumina sequencing with long PacBio reads of genomic DNA using two assembly methods, MaSuRCA and Falcon. Following iteration to maximize contig size and merging of the two assemblies, the final assembly accounted for 15.3 Gb of sequence, with a contig N50 of approximately 233 Kb (Table 1-1). However, while this assembly was highly complete in terms of genome sequence, it lacked the accompanying annotation data that would facilitate its use in genetics studies and, with the exception of the D-genome, the contigs were not associated with a physical position on each chromosome.

Most recently, a chromosome-scale assembly was released by the IWGSC (IWGSC et al. 2018). This genome, referred to as RefSeqv1.0, was not as complete as the assembly produced by Zimin *et al*., with a contig N50 of 52 Kb and covering only 14.5 Gb of the wheat genome (Table 1-1). However, this assembly has become the standard used for wheat research due to its high-quality gene annotation and physical anchoring. Using a combination of multiple gene prediction pipelines, including that previously used for the TGAC assembly, over 100,000 high confidence and 160,000 low confidence gene models were predicted and anchored in their physical order along the 21 chromosomes.

Table 1-1: Quality statistics for wheat genome assemblies. Adapted from Adamski et al. 2019 and updated with additional information.

Genome	**CSS**	**TGACv1**	**Triticum 3.0**	**RefSeqv1.0**	**Durum wheat**
Publication	IWGSC, 2014	Clavijo *et al.*, 2017	Zimin *et al.*, 2017	IWGSC, 2018	Maccaferri *et al.*, 2019
Contigs/ Chromosomes	> 1 million	735,943	N/A	21 chromosomes + ChrU	14 chromosomes + ChrU
Mean contig size	2.1 kb	16.7 kb	233 kb	52 kb	56 kb
Mean scaffold size	7.7 kb	88.7 kb	N/A	Chromosomes	Chromosomes

Assembly Size	10.2 Gb	13.4 Gb	15.3 Gb	14.6 Gb	10.5 Gb
Order	Crude order	Large Bins	Physical order*	Physical order	Physical order
Coding genes	133,090 HC 88,998 LC	104,091 HC 103,660 LC	N/A	107,891 HC 161,537 LC	66,559 HC 303,404 LC

* D genome only

1.6.2.2 Pangenome assemblies

Following the release of RefSeqv1.0, effort is now being put into expanding the range of wheat cultivars and grass species which also have their genome sequence available. The tetraploid durum wheat cultivar Svevo was recently sequenced to the same standard as RefSeqv1.0, including its own gene annotation (Table 1-1) (Maccaferri et al. 2019). Efforts to sequence the so-called "wheat pangenome" are ongoing, with high quality assemblies of cultivars from around the world currently being developed (10+ Wheat Genomes Project 2016). Alongside elite wheat varieties, wild grass species are also being sequenced, particularly those which are the progenitors of tetraploid and hexaploid wheat. A high-quality assembly of wild emmer wheat, which was domesticated to form durum wheat, was released in 2017 which was also assembled into physically contiguous chromosomes with its own gene annotation. The D-genome progenitor *Aegilops tauschii* was assembled to a final contig N50 of 93 kb (Luo et al. 2017, Zhao et al. 2017), while the A-genome progenitor *Triticum urartu* was assembled to chromosome-level pseudomolecules (Ling et al. 2018). The putative B-genome donor *Aegilops speltoides* was identified through sequencing of the total RNA of a series of species which are members of the Sitopsis section of *Aegilops* (Miki et al. 2019).

1.6.2.3 Advances in genetic resources facilitate gene functional characterisation

With these new, gold-standard reference sequences available, the genetic resources available in wheat have also expanded. The exome sequencing of the TILLING mutant populations has facilitated the shift towards reverse genetics approaches in wheat. Using the *in-silico* TILLING database on Ensembl Plants (Krasileva et al. 2017), these SNPs can now be rapidly identified for any gene of interest and crossed to generate double- or triple-mutant lines that completely eliminate the function of all two or three homoeologs (Figure 1-14). Combining these inducible variation methods with accelerated plant growth conditions, as described by "Speed Breeding" (Ghosh et al. 2018, Watson et al. 2018), provides the potential to rapidly produce the knock-out lines needed to carry out phenotypic characterisation. These induced mutations can now be increasingly complemented with transgenic approaches, as the efficiencies of *Agrobacterium*-mediated transformation are slowly improving, with substantial potential for increased efficiencies upon application of the *Baby boom* and *Wuschel2*-mediated transformation technique to wheat. (Lowe et al. 2016) Not only can transgenic approaches be used to produce over-expression lines of a gene or genes of interest, but the CRISPR-Cas9 system has been shown to work successfully in wheat (Shan et al. 2013, Shan et al. 2014), highlighting its potential to generate bespoke loss-of-function mutants, amongst other possible outcomes.

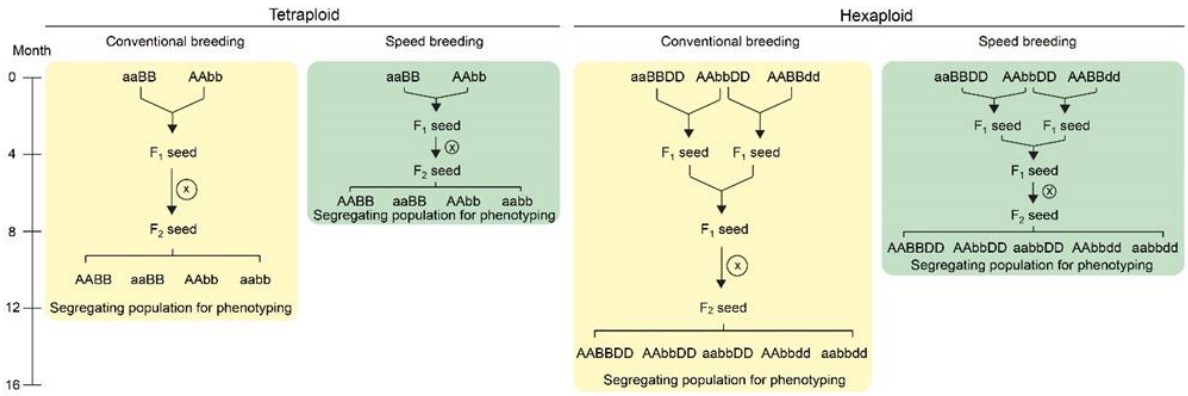

Figure 1-14: Crossing schemes to combine TILLING mutations in tetraploid and hexaploid wheat. The required two or three mutations, respectively, can be crossed to generate the double and triple knock-out mutants. Under conventional breeding conditions (yellow), this takes between 12 and 16 months for tetraploid and hexaploid wheat, respectively. Under "Speed Breeding" conditions (green), this takes 7.5 and 10 months, respectively. Figure reproduced from Adamski et al. (2018), under a CC-BY-4.0 license.

1.6.2.4 Gene expression databases are key resources for hypothesis generation

Yet while these reverse genetics approaches have the potential to substantially improve our understanding of genetic processes in wheat, they are limited by our ability to predict gene function in wheat. In most cases, we can only predict gene function based on orthology with distantly related model species, such as the dicot *Arabidopsis thaliana* (Adamski et al. 2018). In some cases, wheat genes share orthology with genes from rice (*Oryza sativa*) which have been characterised, and which are more likely to share similar roles. However, as we have shown with the example of the *NAM* genes involved in senescence in wheat, some wheat genes do not share roles with their mostly closely related orthologs in *Arabidopsis* (Uauy et al. 2006b) nor rice (Distelfeld et al. 2012). As a result, the selection of genes for characterisation through a reverse genetics approach based on their function in a different species may not lead to the identification of genes which act in that process in wheat. Instead, reverse genetics approaches based on orthology with other species should be complemented with wheat-specific data to identify high quality candidate genes with roles in the process of interest.

Gene expression datasets obtained from wheat itself can begin to account for this discrepancy. Using expVIP, a platform for visualisation and manipulation of gene expression data, publicly-available RNA-Seq data has been assigned metadata categories that allow the independent datasets to be compared amongst themselves (Borrill et al. 2016, Ramírez-González et al. 2018). These expression datasets can be explored to provide insight into the expression patterns of a gene of interest, and to suggest possible functional roles for such a gene. At a higher level, these RNA-Seq datasets have been utilised to produce a series of gene regulatory networks (Ramírez-González et al. 2018). One set is based on gene co-expression, grouping genes into clusters based on similar expression patterns across a set of samples. In this case, clusters were created based on tissue-specific and stress-specific expression patterns. Given knowledge of the functions of a few key

genes within a cluster, it can be possible to predict general functions of other genes in the cluster which can then be tested through targeted phenotyping of mutants.

The RNA-Seq data has also been used to create a GENIE3 network (Huynh-Thu et al. 2010) that uses correlations in expression patterns to predict the target genes for the set of annotated transcription factors from RefSeqv1.0 (Ramírez-González et al. 2018, Harrington et al. 2019a). In this case, predicted targets of a transcription factor with a known function, such as *NAM-A1*, could be considered as candidate genes also involved in that process, here senescence. However, in most cases no biological role is known for a given transcription factor. Here, enrichment analyses of the gene ontology (GO) terms (Ashburner et al. 2000, The Gene Ontology Consortium 2018) associated with target genes of the transcription factor can be used to suggest broad phenotypes that this transcription factor may be involved in regulating, though these analyses are limited by the lack of experimental evidence that supports GO terms annotations (Rhee and Mutwil 2014).

1.7 Statement of purpose

In this thesis we will address the question of how senescence is regulated in wheat. As we have demonstrated in this introduction, a substantial portion of senescence regulation is carried out by transcription factors. Yet in wheat our knowledge of the transcription factors regulating senescence, and the interactions between any such genes, is minimal. In this work, we aim to both identify novel regulators of senescence in wheat and to further characterise the role of the NAM-1 genes, known positive regulators of senescence.

We will first take advantage of the TILLING population to carry out a phenotypic screen for novel alleles affecting the onset and progression of senescence. From this screen, we phenotypically characterised and mapped a dominant chlorosis and early senescence mutant to a 4.3 Mb locus on chromosome 3A. Having also identified as part of this screen a pair of mutants in *NAM-A1*, we will then carry out an analysis of the effect of missense mutations on the functioning of *NAM-A1*. We will also expand this analysis to the effect of the mutations on the functioning of NAC transcription factors in general, particularly considering their effect on protein-protein interactions.

We will then discuss the development and characterisation of a heat-shock inducible transgenic system to characterise the impact of ectopic expression of *NAM-B1* on senescence onset in wheat. Having used this system to show that premature expression of *NAM-B1* is not sufficient to induce early senescence, we will then address the question of whether NAM-B1 and its homoeologs physically interact with other proteins to induce senescence. Using a yeast two-hybrid screen, we identify a set of NAC transcription factors which bind NAM-B1, one of which we show to be a positive regulator of senescence. We then characterise the regulatory targets of this transcription factor using a wide range of gene expression datasets and network analyses.

2 Identification of a dominant chlorosis phenotype through a forward screen of the *Triticum turgidum* cv. Kronos TILLING population.

The work included in this chapter was previously published in *Frontiers in Plant Science* in July 2019 (Harrington et al. 2019b). The author was supported in this work by Dr. Philippa Borrill, who carried out the initial TILLING mutant screen at the JIC in 2015, and by Dr. Nicolas Cobo who managed the Davis field trials and assisted with phenotyping in the 2016-2017 field season. The chromosome flow sorting was carried out by Prof. Jaroslav Doležel and Dr. Miroslava Karafiátová

2.1 Introduction

2.1.1 Polyploidy in wheat hides genetic variation

Domesticated wheat is the product of one or two allopolyploidsation events, in the cases of tetraploid durum wheat and hexaploid bread wheat, respectively. As explained in more detail in the main introduction (Figure 1-5, p.4), tetraploid wheat is derived from a single hybridisation event between two wild grasses approximately 400,000 years ago (Dubcovsky and Dvorak 2007, Borrill et al. 2019b). This hybridisation produced the allotetraploid wild grass, known as wild emmer wheat or *Triticum turgidum* ssp. *dicoccoides*. This grass was itself domesticated to form durum wheat, *Triticum turgidum* ssp. *durum*. More recently, a second hybridisation took place to produce hexaploid wheat. Approximately 10,000 years ago, domesticated emmer wheat hybridised with the diploid grass *Aegilops speltoides*, producing a hexaploid grass which would be domesticated as bread wheat (*Triticum aestivum*).

This comparatively recent set of hybridisations, occurring on a time scale of thousands of years rather than millions, has resulted in a highly redundant complement of genes. Unlike in other staple crop species such as rice (*Oryza sativa*) and maize (*Zea mays*), most genes are present as two or three copies, in tetraploid and hexaploid wheat, respectively. These gene copies, referred to as homoeologs, often have very similar expression patterns and similar functions (Ramírez-González et al. 2018). It appears that, in the majority of cases, the duos and triads of homoeologous genes have not yet begun to undergo the pseudogenisation and neofunctionalization that is characteristic following whole genome duplication (WGD) events (Dodsworth et al. 2016, Clark and Donoghue 2018). WGD events are common in the history of plants, particularly angiosperms (Soltis and Soltis 2016), and have been thought to help drive innovation in part through neofunctionalization of duplicated genes (Dodsworth et al. 2016, Clark and Donoghue 2018). Yet these processes have occurred over timescales of millions of years—the monocot specific duplication, τ, is thought to have occurred at least 20 million years before the divergence of the *Poaceae* (Paterson et al. 2012). All monocot species, including crops such as rice and maize, are therefore ancestral polyploids, or paleopolyploids (Levy and Feldman 2002). Since their ancestral polyploidisation events, the

functional redundancy between the homoeologous genes has been lost through diploidisation processes, such as neo-, sub-, and non-functionalisation (Levy and Feldman 2002). Wheat allopolyploidy, on the other hand, is too recent for the functional redundancy between homoeologs to have been lost.

A direct result of this homoeolog redundancy in wheat is that many traits are controlled by all three of the homoeologous genes. This genetic redundancy can appear in many forms, as detailed in Figure 2-1. Some genes in wheat do act in a dominant manner, such as *VERNALISATION1* (*VRN1*) (Yan et al. 2003) and *PHOTOPERIOD* (*PPD*) (Wilhelm et al. 2009), where a mutation in a single homoeolog is sufficient to yield the full mutant phenotype. Often these dominant mutations are driven by variation in regulatory regions of the genome, such as promoters. However, most cases currently described involve at least a level of homoeolog redundancy. This can be additive, as in the case of *GPC* (*NAM*) (Avni et al. 2014) or *GW2* (Wang et al. 2018b), where the mutation of each homoeolog leads to a stepwise increase towards the mutant phenotype. This dosage effect can also be non-additive, such as in *Waxy* (Kim et al. 2003), where mutations in one or two of the homoeologs leads to a small phenotypic effect, but the triple mutant is required to show the complete mutant phenotype. Finally, some wheat homoeologs exhibit functional redundancy, such as *SBE-II* (Slade et al. 2012) and *MLO* (Acevedo-Garcia et al. 2017). In these cases, mutations in only one or two of the homoeologs causes no mutant phenotype; only the complete triple mutant leads to any effect at all.

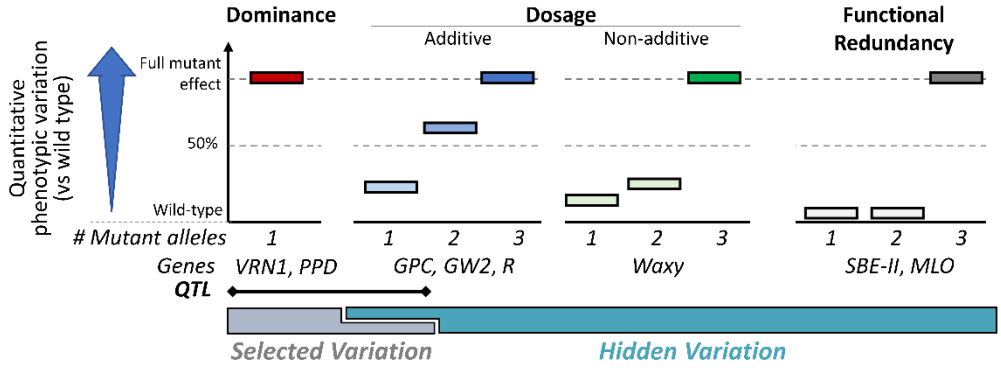

Figure 2-1: The effects of homoeolog redundancy on accessible genetic variation in wheat. Mutations in single homoeologs can lead to dominant effects which underpin many quantitative trait loci. Other mutations, leading to dosage effects or functional redundancy, are more difficult to select as they cause little to no phenotypic variation. This figure is reproduced from Borrill et al. (2019b), under the terms of a CC-BY-4.0 license.

Efforts to carry out forward genetic screens in wheat are, as a result, substantially hindered by the redundancy present between homoeologous genes. Not only are quantitative traits influenced by the actions of many independent genes, but any mutant phenotype is often masked by the redundancy provided by the two remaining homoeologs present in bread wheat (Borrill et al.

2019b). In response to this, tetraploid durum wheat can be a powerful tool to facilitate forward genetic screens in wheat. Having only undergone one round of hybridisation (Dubcovsky and Dvorak 2007), durum wheat typically contains two copies of each gene, reducing the redundancy present in the genome.

2.1.2 New resources in wheat aid the characterisation of gene function

Recently, large advances have been made not only in wheat resources but in tetraploid wheat resources specifically. Shortly after the release of the hexaploid wheat genome, RefSeqv1.0 (IWGSC et al. 2018), the tetraploid wheat genome, derived from the commercial cultivar Svevo (Maccaferri et al. 2019), was released to the same quality standard. As part of the on-going wheat pangenome project, the Californian durum wheat variety Kronos is also being sequenced. This variety was used to create a TILLING population, containing over 4 million EMS-induced point mutations, that was then exome-captured and sequenced (Uauy et al. 2009, Krasileva et al. 2017). These mutations have been made available through Ensembl Plants (Kersey et al. 2017), providing an *in-silico* TILLING platform through which mutations in specific genes of interest can be rapidly identified. This TILLING population can now be used for forward genetic screens, with the reduced redundancy burden of tetraploid wheat.

Yet even if mutants of interest could be identified in the Kronos TILLING population, identifying the causal mutations underpinning the phenotype is not trivial. This is particularly the case because of the large size of the wheat genome, spanning 16 Gb in hexaploid wheat (IWGSC et al. 2018) and 11 Gb in tetraploid wheat (Maccaferri et al. 2019). Sequencing strategies to identify loci associated with a mutant phenotype must, as a result, attempt to reach a balance between sufficient coverage over a sufficient area of the genome and reasonable costs. Many different approaches have been suggested for this, typically involving a form of reduced representation where only a particular part of the genome is sequenced. This include exome capture sequencing, where baits are used to enrich only the coding regions of the genome before sequencing (Mamanova et al. 2010, Krasileva et al. 2017). Even particular gene families can be enriched and sequenced, as in the case of R-gene enrichment sequencing (Jupe et al. 2013, Steuernagel et al. 2016), further reducing the required sequencing space. However, these methods suffer from bias towards pre-existing knowledge; that is, exome sequencing or R-gene enrichment sequencing require baits to be designed to enrich the desired region of the genome. If part of the genome isn't represented in the bait library, it won't be enriched and sequenced. Additionally, these methods are typically biased towards enriching coding regions of the genome, though this is beginning to change as new captures are developed which include some promoter and intron sequence (Gardiner et al. 2019). This results in variation outside of the coding regions being lost, which can be of particular importance when dealing with dominant variation which is often, though not exclusively (Simons et al. 2006, Greenwood et al. 2017), due to variation outside of the coding sequence in regulatory regions (Yan et al. 2004, Fu et al. 2005, Borrill et al. 2015a).

Some methods do exist for comparatively unbiased reduced representation sequencing. One such method is chromosome flow-sorting, which reduces the size of the genome to only that of the desired chromosome which is isolated, amplified, and then sequenced (Doležel et al. 2012). Alternative methods include skim sequencing, previously implemented in rice and currently being implemented in wheat, which carries out low coverage sequencing across the entire genome. This can allow the cost-effective sequencing of large panels at the whole-genome scale, and dependent on the sequencing depth, facilitate the identification of features such as deletions, duplications, and SNPs across the genome (Huang et al. 2009).

2.1.3 Aims and hypotheses

Given the wealth of novel variation present within the TILLING populations, we hypothesized that we could use a forward screen of the TILLING population to identify novel senescence mutants. In this chapter we take advantage of these new advances in sequencing technologies to identify and fine-map a novel locus controlling chlorosis and early senescence. We screen the tetraploid Kronos TILLING population for early and late senescence phenotypes, identifying a single line that segregates for a dominant chlorosis, early senescence mutation. We find that this phenotype is consistent across multiple years of field trials in the UK but is not present when grown in California and under glasshouse conditions. Using bulked segregant analysis and exome capture sequencing, we define the causal locus to a single interval on chromosome 3A, which we call *Yellow Early Senescence 1* (*YES-1*). We then identify novel SNPs in the TILLING line by flow-sorting and sequencing the mutant chromosome 3A. With these SNPs, we fine-map the *YES-1* locus to 4.3 Mb, containing 59 high-confidence genes.

2.2 Methods

2.2.1 Field trials

2.2.1.1 TILLING population screen

The Kronos TILLING population (N=951 M_4 lines) was initially sown at the John Innes Centre in November 2015, at Church Farm, Bawburgh (52°38'N 1°10'E). All JIC field trials were carried out at Church Farm, though were sown in different fields at the farm each year. The lines were sown in un-replicated single 1 m rows in numerical order (Figure 2-2A), such that lines Kronos0423 was followed by line Kronos0427. For simplicity, TILLING lines will be referred to as KXXXX throughout the chapter (i.e. Kronos0423 as K0423). Wild-type lines (cvs. Kronos, Paragon, and Soissons) were sown as controls, distributed randomly throughout the population. All rows were phenotyped for flag leaf and peduncle senescence as detailed below. From the set of 951 M_4 lines, independent M_4 individuals from the 10 earliest senescence and 11 latest senescing lines were crossed to wild-type Kronos, under glasshouse conditions (see Table 2-5 in section 2.3.1). The resultant F_1 plants were then self-pollinated, producing F_2 seed (Figure 2-7E). Sufficient F_2 seeds were not recovered for three of the mutant lines (K0331, K3085, and K3117), and as a result these

populations were not carried forwards. All original TILLING mutant lines can be obtained through the JIC Germplasm Resources Unit (www.seedstor.ac.uk).

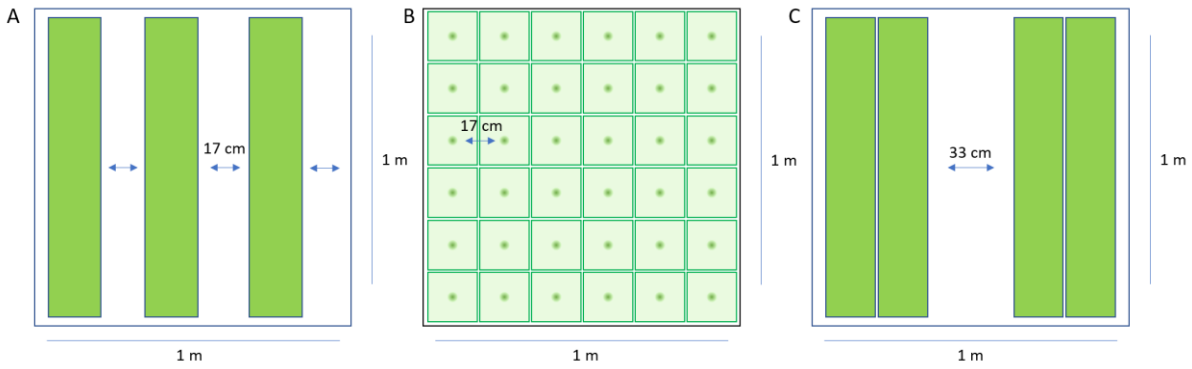

Figure 2-2: Diagram of field trial plot layouts. The 2015 and 2017 trials at the John Innes Centre (JIC) were laid out as in (A), with three single 1 m rows of plants (green) sown per 1 m^2 plot, separated by approximately 17 cm. The 2016 JIC trials were hand-sown into a 6x6 grid in the 1 m^2 plot (B), with approximately 17 cm between each individual plant. (C) The 2018 JIC trial and the 2017 Davis, CA trial were both sown with 2 double rows per 1 m^2 plot, each separated by approximately 33 cm.

2.2.1.2 Recombinant scoring

The F_2 populations, derived from the early and late senescing TILLING lines backcrossed to cv. Kronos, were grown at Church Farm in the 2016 field season. Individual F_2 seeds were hand-sown in 6x6 1 m^2 grids, leaving approximately 17 cm between each plant, as depicted in Figure 2-2B. 116 kg/Ha of Nitrogen was applied to the 2016 trials as fertiliser. 31 F_2 populations were sown, derived from 18 independent TILLING mutant lines (see Table 2-5). For the line K2282, two independent F_2 recombinant populations were sown, referred to as K2282-A and K2282-B. All F_2 populations were phenotyped for flag leaf and peduncle senescence at the individual plant level, as discussed in more detail in the following sections.

F_3 Seeds from the K2282-A and K2282-B populations were taken forward for the remaining field trials in 2017 and 2018. Seeds were selected from F_2 plants that were genotyped as either completely heterozygous across the *YES-1* region or which contained recombination within the mapped interval. A subset of F_3 plants completely mutant and completely wild-type across the *YES-1* region were also grown as controls. In 2017, 30 lines were grown from the K2282-A population, alongside a further 8 lines from the K2282-B population. The F_3 seeds were sown in a randomised block design, with each line replicated between 1 and 4 times depending on seed availability. The seeds were sown as single 1 m rows, with three 1 m rows sown per 1 m^2 plot separated from each other by approximately 17 cm (Figure 2-2A). Within each row, the primary tillers of 12 individual plants were tagged before heading. From these tagged plants, tissue was sampled for DNA extraction and the individual plants were scored for chlorosis and early senescence.

In 2018, 374 individual F_3 seeds were hand-sown in a 1 m^2 grid, as shown in Figure 2-2B. These seeds were derived from 16 F_2 plants which were genotyped as fully heterozygous across the region flanked by the SH467 and SH969 markers. The individual plants were sampled for tissue and scored for individual plant senescence and chlorosis, as in 2016. Nitrogen was applied to the field as fertiliser in both years, 147 kg/ha in 2017 and 124.5 kg/ha in 2018. Precipitation data for the JIC field trials across all years was obtained from a weather station at 52°37'52.29" N, 1°10'23.57" E.

2.2.1.3 Phenotypic characterisation

In 2018, nine lines were grown for more completely phenotypic characterisation. These plants were the progeny of F_3 plants which were genotyped as completely homozygous mutant (4) or homozygous wild-type (5) across the region flanked by markers SH467-SH969. The plants were sown in double 1 m rows separated by approximately 33 cm, within a 1 m^2 plot (Figure 2-2C). The individual lines were replicated three times in a complete randomised design. Three replicates each of wild-type Kronos and M_5 mutant seed from the K2282 line were sown as controls. Within each row, two tillers were tagged at heading from which tissue samples were taken. These individual plants were also scored for chlorosis and senescence and were used for SPAD measurements. The double rows were also scored for senescence.

2.2.1.4 Putative candidate gene phenotyping

In 2018, two independent TILLING lines were grown at the JIC which contained mutations in the putative candidate gene TraesCS3A02G414000. 24 seeds were hand sown per line into single unreplicated 1 m rows. The TILLING lines were chosen using the Ensembl Plants *in silico* TILLING browser to identify lines containing frameshift and deleterious missense mutations (SIFT = 0).

2.2.1.5 Davis, California trial

In November 2016, 91 F_3 lines from the K2282-A population and 55 F_3 lines from the K2282-B population were sown at the University of California (UC) Field Station near Davis, California (38° 31' N, 121° 46' W). The lines were selected if the F_2 parent contained recombination within the region flanked by markers SH467 and SH969, or alternatively if the F_2 parent was fully heterozygous across that region. Additionally, 12 lines from F_2 parents genotyped as fully wild-type or mutant across the region were grown. The lines were sown in double 1 m rows, each separated by an empty row and arranged in a complete randomised design (Figure 2-2C). Within each double row, eight individual plants were tagged at heading for tissue sampling and individual scoring of chlorosis and senescence. 225 kg N/hectare was applied to the trial (as ammonium sulphate), half before planting and half on March 31st (Z30 stage). The trial was treated with appropriate fungicides to prevent stripe rust (*Puccinia striiformis* f. sp. *tritici*). Precipitation data for the Davis trial was obtained from the Davis, California weather station (38° 32' 07'' N, 121° 46' 30'' W).

2.2.1.6 Glasshouse trial

We selected F_3 plants derived from F_2 parents which were genotyped as completely wild-type or completely mutant across the region flanked by markers SH179 and SH969. The seeds were pre-germinated in the dark at 4°C for 48 hours. They were then sown into P96 trays containing 85% fine peat and 15% horticultural grit, before being transplanted into 1 L pots at the three-leaf stage. The 1 L pots contained either a) Petersfield Cereal Mix (Petersfield, Leicester, UK), b) Horticultural Sand (J. Arthur Bower's, Westland Horticulture), or c) Soil taken from the Church Farm site used for JIC field trials (Bawburgh, UK). The plants sown into sand were supplemented with nutrient-replete Hoagland solution (Hoagland and Arnon 1950), with 100 mL applied every 3 days.

The plants were also tested under low water conditions for each of the three soil types listed. In this condition, plants were watered once weekly and then additionally as needed to maintain a soil volumetric water content of approximately 20%, as measured with the Decagon GS3 sensor (ICT International, Armidale, Australia). Under each soil and water condition combination, three plants of each genotype were tested. The plants were visually phenotyped for chlorosis and early senescence phenotypes.

2.2.2 Plant phenotyping

2.2.2.1 Visual colour scoring

Plants were visually scored for the presence of chlorosis, based on their visual appearance in the field. Plants which phenocopied the wild-type Kronos plants were scored as "green", those which phenocopied the TILLING mutant parent (MP) plants were scored as "yellow" (see depiction in Figure 2-7A).

2.2.2.2 Senescence phenotyping

Individual plants were scored for senescence on the main tiller of the plant, which was typically tagged at heading, when the spike was 25% emerged (Zadoks growth stage 57) (Zadoks et al. 1974). Flag leaf senescence was scored when 25% of the leaf showed visual yellowing and necrosis from the tip. Peduncle senescence was scored when the top 3 cm of the main peduncle were completely yellow. Rows were scored considering the progression of senescence across all plants present in the row. Heading was scored when 75% of the main spike reached Zadoks stage 57. Leaf senescence was scored in a similar manner, when 75% of the main flag leaves were 25% senesced. Peduncle senescence was scored when the top 3 cm of 75% of the main peduncles were yellow.

Alongside visual scoring, the SPAD-502 meter (Konica Minolta, Osaka, Japan) was used to obtain non-destructive measurements of leaf chlorophyll content. Individual plant measurements were taken in the field in 2016, 2017, and 2018. Eight individual SPAD readings were taken along the flag leaf on each side of the midrib, which were then averaged to obtain a single SPAD value for each leaf. Row SPAD scores were also obtained in 2018; these were calculated from SPAD

measurements of flag leaves on the two tagged tillers, which were then averaged and taken as the SPAD reading for that row.

2.2.2.3 Chlorophyll quantification

Leaf tissue samples were taken in 2016 and 2018 from the JIC field trials, from which chlorophyll content was directly measured. In 2016, the flag leaves were sampled at heading (N = 3 per genotype), each leaf sampled from an individual plant was treated as an independent biological replicate. In 2018, flag leaf tissue was sampled at anthesis, and third leaf tissue was sampled at Zadoks stage 13-14, approximately 24 days pre-anthesis (Mutant, N = 8, Wild-type, N = 10). In this case, one leaf was sampled per row and was treated as an independent biological replicate. From each leaf, three 1 cm^2 discs were extracted, taken from the base, middle, and tip of the leaf. Chlorophyll was then extracted from the leaf discs as described in (Wellburn 1994). The leaf discs were placed in opaque 15 mL tubes, to which 3 mL of N,N-Dimethylformamide (analytical grade, Sigma Aldrich, UK) was added. The tubes were left at 4°C, shaking, for 48-64 hours until the pigment was completely leached from the tissue. The absorbance of the N,N-Dimethylformamide, which now contains the pigment leached from the leaves, was measured at 664, 647, and 480 nm. The levels of three pigments—chlorophyll A, chlorophyll B, and carotenoids—were calculated using these absorbances as follows:

$$ChlA = 11.65(A_{664}) - 2.69(A_{647})$$

$$ChlB = 20.81(A_{647}) - 4.53(A_{664})$$

$$Carotenoids = (1000*A_{480} - 0.89*ChlA - 52.02*ChlB)/245$$

2.2.2.4 Leaf mineral content

Samples for mineral content quantification were taken from flag leaves of individual plants in the 2018 field trial. Leaf samples of approximately 0.2 g were dried and ground to a fine powder, before being digested by 2 mL of nitric acid (67-69%, low-metal) and 0.5 mL of hydrogen peroxide (30-32%, low-metal) for 12 hours at 95° C. Samples were diluted 1:11 in ultrapure water before analysis with ICP-OES (Vista- PRO CCD Simultaneous ICP-OES; Agilent). Calibration was carried out using standards of Zn, Fe, and Mg at 0.2, 0.4, 0.6, 0.8, and 1 mg/L and Mn and P at 1,2, 3,4, and 5 mg/L.

2.2.2.5 Light microscopy

Thin sections of flag leaves from the 2018 field trial were cut with a razor from mutant and wild-type plants. The sections were imaged using a Leica MZ16 light microscope (Meyer Instruments, Houston, USA; N = 3 per genotype).

2.2.3 Bulked segregant analysis

Individual plants with green and yellow phenotypes from the K2282 F_2 populations sown at the JIC in 2016 were selected for bulked segregant analysis. DNA from seedling tissue was extracted using

the QIAGEN DNeasy Plant Mini Kit. The quality and quantity of the DNA was checked using a DeNovix DS-11 Spectrophotometer, Qubit (High Sensitivity dsDNA assay, Q32854, Thermo Fisher), and by running a sample of the DNA on an agarose gel (1%) to visualise the presence of high molecular weight DNA. Four bulks were assembled by pooling DNA from plants which had been scored as either "yellow" or "green" (K2282-A, N = 75 for yellow, N = 16 for green; K2282-B, N= 33 for yellow, N = 22 for green). Equal quantities of DNA from the individual plants were pooled into each bulk to minimise bias.

Library preparation and sequencing was carried out at the Earlham Institute (Norwich, UK). DNA quality was confirmed using the High Sensitivity Qubit assay, before library preparation was carried out with a KAPA HTP Library Prep Kit (KK8234, Roche). Beckman Coulter XP beads (A63880, Fisher Scientific) were used for size selection, and DNA was sheared to approximately 350 bp using the Covaris S2 Sonicator (Life Technologies). Four libraries were produced, one for each bulk detailed above, which were barcoded and pooled. Five cycles of PCR were carried out on the libraries before carrying out exome capture.

Hybridization to the wheat NimbleGen target capture, previously described in Krasileva et al. (2017), was carried out using the Seq Cap EZ protocol v5.0 (Roche), with the following changes: 2.8 µL of Seq Cap HE Universal Oligo was used, and the Cot-1 DNA was replaced with 14 µL of Seq Cap EZ Developer Reagent (cat. no. 6684335001, Roche). Hybridisation was carried out at 47°C for 72 hours in a PCR machine with a lid heated to 57°C.

The library pool was diluted to 2 nM with NaOH and 10µL transferred into 990µL HT1 hybridisation buffer (Illumina) to give a final concentration of 20pM. This was diluted further to an appropriate loading concentration in a volume of 120 µL and spiked with 1% PhiX Control v3 (cat. no. FC-110-3001, Illumina) before loading onto the Illumina cBot. The flow cell was clustered using HiSeq PE Cluster Kit v4, utilising the Illumina PE_HiSeq_Cluster_Kit_V4_cBot_recipe_V9.0 method on the Illumina cBot. After clustering, the flow cell was loaded onto the Illumina HiSeq2500 instrument following the manufacturer's instructions. The sequencing chemistry used was HiSeq SBS Kit v4. The library pool was run on two lanes with 125 bp paired end reads. Reads in bcl format were demultiplexed using the 6 bp Illumina index by CASAVA 1.8, allowing for a one base-pair mismatch per library, and converted to FASTQ format by bcl2fastq.

2.2.4 Chromosome flow-sorting and sequencing

In order to ensure all parental SNPs were included, we used seeds from the K2282 M_5 mutant line for the chromosome sorting and sequencing. Suspensions of intact mitotic chromosomes were prepared from synchronized root tip meristems as previously described (Vrána et al. 2000). To increase the discrimination of individual chromosomes by flow cytometry, GAA microsatellite loci were fluorescently labelled by Fluorescence In Situ Hybridization in Suspension (FISHIS) (Giorgi

et al. 2013) using FITC-labelled (GAA)7 oligonucleotides (Vrána et al. 2016). Chromosomal DNA was then stained by 2 µg/mL 4′,6-diamidine-2′-phenylindole (DAPI) and the chromosome suspensions were analysed by FACSAria SORP II flow sorter (BD Biosciences, San Jose, USA) at 1000-2000 particles/sec. The population of chromosomes representing chromosome 3A was extracted. The bivariate flow karyotypes of DAPI vs. GAA-FITC were used to estimate the extent of contamination by other chromosomes (Figure 2-3). Briefly, 2000 chromosomes were sorted onto a microscopic slide and probed for GAA microsatellites and Afa-family repeats, allowing analysis by Fluorescence *In Situ* Hybridization (FISH) (Kubaláková et al. 2002). Three batches of 30,000 copies of chromosome 3A, approximately 50 ng of DNA each, were then sorted into PCR tubes containing 40 µL sterile deionized water. Chromosomal DNA was purified and amplified by Illustra GenomiPhi V2 DNA amplification Kit (GE Healthcare, Piscataway, USA) according to Šimková et al. (2008).

Library preparation and sequencing were carried out at Novogene (Hong Kong). DNA integrity was confirmed on 1% agarose gels following which a PCR-free library preparation was carried out, using the NEBNext Ultra II DNA Library Prep Kit for Illumina, following manufacturer's instructions. Libraries were sequenced using a HiseqX platform, generating 150 bp paired end reads.

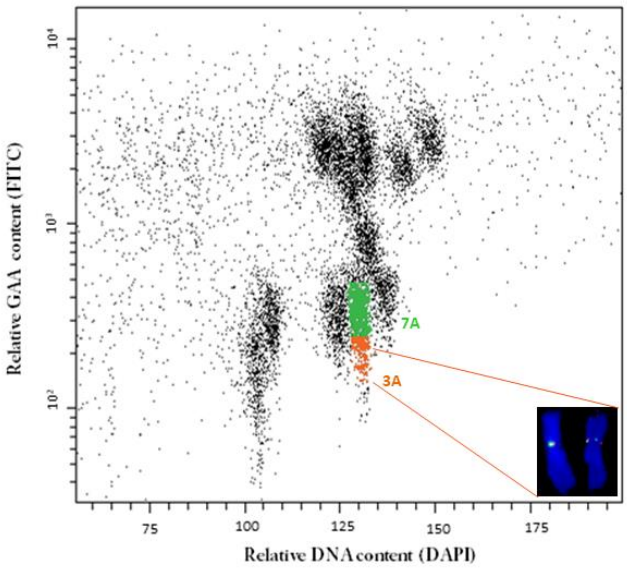

Figure 2-3: Bivariate flow karyotype DAPI vs. GAA-FITC obtained after the analysis of mitotic chromosomes isolated from Triticum durum cv. Kronos. The regions of the karyotype representing chromosomes 3A and 7A are highlighted in orange and green, respectively. Inset: Images of flow-sorted chromosome 3A, which was identified by FISH with probes for GAA microsatellite (green) and Afa-family repeat (red). The chromosomes were counterstained by DAPI (blue).

2.2.5 Sequencing alignments and SNP calling

For the bulked segregant analysis, raw Illumina reads were aligned to RefSeqv1.0 (IWGSC et al. 2018), using bwa-mem (v 0.7.5) with the default settings (-k 20, -d 100) (Li 2013). Alignments

were sorted, indexed, and PCR duplicates removed using SAMtools (v 1.3.1) (Wysoker et al. 2009), and SNPs were called using freebayes (v 1.1.0, default settings) (Garrison and Marth 2012). Depth of coverage was calculated using the exome capture size obtained in the initial analysis (Krasileva et al. 2017) (Table 2-6). Following SNP calling, the output was filtered to include only SNPs that were previously called in the K2282 parent line (Krasileva et al. 2017) using an original script to convert SNP coordinates to the RefSeq v1.0 genome (https://github.com/Uauy-Lab/K2282_scripts; Appendix 8.2.1). The original K2282 SNPs were obtained from the www.wheat-tilling.com website. The relative enrichment of each SNP in the yellow and green bulks was visualised across the wheat genome using the Circos package (Krzywinski et al. 2009). To do this, the AO/DP ratio was calculated for each SNP position in the green and yellow bulks, where AO is the number of reads with the mutant, or "alternate", allele at the position in question, and DP is the depth of reads at that position. A Δ value was then calculated by subtracting the AO/DP ratio for the green bulk from the yellow bulk. A schematic of the pipeline is provided in Figure 2-9. These values were calculated separately for the K2282-A and K2282-B populations.

Following flow-sorting of chromosome 3A from the K2282 M_5 mutant line, reads were aligned to both the RefSeq v1.0 and the Kronos assembly. A schematic of this workflow is provided in Figure 2-17. We were provided with access to the draft Kronos assembly, developed by the Earlham Institute, which was assembled using protocols previously described (Clavijo et al. 2017a, Clavijo et al. 2017b). The Kronos assembly has been made available in advance of publication from Grassroots Genomics (https://opendata.earlham.ac.uk/opendata/data/Triticum_turgidum/). The alignment of raw reads from chromosome 3A was carried out using bwa-mem (v 0.7.5; default settings -k 20, -d 100) (Li 2013). Illumina reads from the wild-type Kronos assembly were aligned to RefSeq v1.0 using hisat (v 2.0.4, default settings with -p 8) (Kim et al. 2015). In all cases, files were sorted, indexed, and PCR duplicates removed with SAMtools (v 1.3.1) (Wysoker et al. 2009). For alignments to RefSeq v1.0, the depth of coverage across the second part of chromosome 3A was calculated using genomic windows of 1 Mb (Table 2-6). Depth of coverage was not calculated for the complete Kronos alignment, as the scaffolds are not associated with a physical chromosome. SNPs were called on the respective alignments using freebayes (v 1.1.0) at default settings in all cases (Garrison and Marth 2012). BCFtools (Wysoker et al. 2009) was used to filter the SNPs based on quality (QUAL $\geq$ 20), depth (DP > 10), zygosity (only homozygous), and whether the SNPs were EMS-like (G/A or C/T SNPs). We only retained homozygous SNPs as the *YES-1* region was genotyped to be completely homozygous in the M_5 TILLING mutants using previous markers.

Following alignment and SNP calling of the chromosome 3A reads against the Kronos genome, the SNPs were manually filtered to remove regions of very high SNP density, which were likely to contain false positives. This was defined as having a greater SNP density than 1 SNP per 20-45 Kb of DNA (Krasileva et al. 2017). We then identified scaffolds from the Kronos genome which were

physically located within the *YES-1* locus in the Chinese Spring RefSeq v1.0 genome using the Basic Local Alignment Search Tool (BLAST) (v2.2.30; default parameters with -max_hsps 1 and -outfmt 6) (Altschul et al. 1990) against the gene sequences annotated within that region, based on the v1.1 gene annotation. All further analysis of the SNP data for mapping and marker design focused solely on the 32.9 Mb *YES-1* region.

Varietal SNPs between Kronos and Chinese Spring were identified by aligning the raw wild-type Kronos reads against RefSeqv1.0 (as detailed above). These varietal SNPs were then removed from the set of SNPs called based on the alignment of the K2282 mutant chromosome 3A reads to RefSeqv1.0. These SNPs were also manually curated to remove any which fell in regions of unreasonably high SNP density, as defined above.

2.2.6 KASP marker genotyping

1.7.1.1 DNA extraction

DNA for genotyping was extracted from wheat leaf tissue as described in (Pallotta et al. 2003). Briefly, tissue samples (approximately 3 cm in length) were ground in 500 mL of extraction buffer (0.1M Tris-HCl pH 7.5, 0.05 M EDTA pH 8.0, 1.25% SDS) in a 96-well plate (Fisher Scientific cat. no. AB0564). After incubation for 1 hour at 65°C, the plates were cooled to room temperature and 250 µL of 6 M ammonium acetate was added to each sample. The plates were cooled for 15 minutes at 4°C, and then centrifuged for 15 minutes at 4000 rpm (Eppendorf® 5810R, cat. no. 5810000017). 600 µL of the supernatant was transferred to a new plate, with 250 µL of ice-cold isopropanol in each well. The supernatant and isopropanol were mixed and left for five minutes to allow DNA precipitation. The plate was then centrifuged for 15 minutes at 4000 rpm to pellet the DNA. The supernatant was removed, and the pellet was washed in 500 µL of ethanol, before being pelleted again (15 minutes, 4000 rpm). After removal of the supernatant, the DNA pellet was left to dry overnight before being resuspended in 100 µL of dH_2O.

1.7.1.2 KASP marker analysis

Markers were designed against SNPs using primarily the PolyMarker pipeline (Ramirez-Gonzalez et al. 2015b). Where markers could not be automatically designed using PolyMarker, we designed markers manually with the aim to make them homoeolog specific. Kompetitive Allele Specific PCR (KASP) genotyping was carried out using standard protocols, as previously detailed in (Ramirez-Gonzalez et al. 2015a). The primer mix and reaction components are provided in Table 2-1 and Table 2-2, and the reaction cycling conditions are provided in Table 2-3. All KASP reactions were carried out in 2.5 µL volume in 384-well plates.

Table 2-1: Primer mix for KASP assays.

Component	**Volume (µL)**
WT Primer (100 µM)	12
Mut Primer (100 µM)	12
Common Primer (100 µM)	40
dH_2O	36

Table 2-2: KASP assay reaction components. DNA is diluted 1:10 from the initial DNA concentration.

Component	**1X (µL)**
KASP® Reagent *or* PACE® reagent	2.5
Primer mix	0.07
DNA (diluted 1:10)	2.5

Table 2-3: KASP assay reaction conditions. Plates can undergo between 26 and 35 cycles of amplification to optimise cluster separation.

Step	**Temperature**	**Time**
Hot-start *Taq* activation	94°C	15 min
Touchdown (x10)	94°C	20 s
	65°C *(Decreasing 0.8°C per cycle to reach 57°C)*	60 s
Amplification (x26-35)	94°C	20 s
	57°C	60 s
Hold	10°C	-

Following KASP reactions, the plates were read using the PHERAstar microplate reader (BMC Labtech, Aylesbury, United Kingdom), and samples grouped into wild-type, mutant, and heterozygotes based on the relative levels of *Vic* and *Fam* fluorescence provided by the primer tails. Markers specific to K2282 are listed in Table 2-4. Markers used for *NAM-A1* genotyping are provided in Chapter 3, Table 3-1 (p.93) and were previously published (Harrington et al. 2019c).

2.2.7 Data Analysis

Appropriate statistical tests for all data analyses were carried out and are detailed explicitly in the results section. If multiple comparisons were being carried out on the same analysis, adjustments for false discovery rate (FDR) were included using the Benjamini-Hochberg adjustment. This is referred to in the results as "adjusted for FDR." All statistical analysis was carried out in R (v3.5.1) (R Core Team 2018), and data was manipulated using packages tidyr (Wickham and Henry 2018) and dplyr (Wickham et al. 2019). Graphs of phenotyping and expression data were produced using ggplot2 (Wickham 2016) and gplots (Warnes et al. 2019), respectively.

Table 2-4: KASP Markers used in mapping the YES-1 locus. Genome specific KASP markers were designed for SNPs identified in the three mapping stages, using the exome capture data and the flow-sorted chromosome 3A data mapped against either the Kronos genome or the RefSeqv1.0 Chinese Spring genome. The SNP position is given for the RefSeqv1.0 genome and was obtained for the Kronos SNPs using BLAST. The HEX tail (WT) and the FAM tail (Mut) are highlighted in red and blue, respectively.

Marker	WT	Mut	Common	SNP Position	SNP Origin
SH266	gaaggtcggagtcaacggattgacatgagccaacagccatc	gaaggtgaccaagttcatgctgacatgagccaacagccatt	gcaggcaaacaaatcaaaatctat	627966266	Original Exome Capture
SH467	gaaggtcggagtcaacggattcttaaacacgttgcagatcagg	gaaggtgaccaagttcatgctcttaaacacgttgcagatcaga	ctccaaatcctcccccaaac	642038467	Original Exome Capture
SH17119	gaaggtcggagtcaacggattctggtttcaaggaaattggaag	gaaggtgaccaagttcatgctctggtttcaaggaaattggaaa	gtttctctactgtcccaagt	644827976	Kronos Genome
SH179	gaaggtcggagtcaacggattcagtcaagtcaggctcc	gaaggtgaccaagttcatgctcagtcaagtcaggctct	gtgctggaacaaatgaa	647452179	Original Exome Capture
SH123480	gaaggtcggagtcaacggattagcatcatatctgggttggtc	gaaggtgaccaagttcatgctagcatcatatctgggttggtt	atgatgaattgaagctattgg	652471882	Kronos Genome
SH044	gaaggtcggagtcaacggattcggcgtgccagggctttctc	gaaggtgaccaagttcatgctcggcgtgccagggctttctt	cagcccatatctgggatga	654781044	RefSeqv1.1
SH838	gaaggtcggagtcaacggattgaatcagcttctacagatgg	gaaggtgaccaagttcatgctgaatcagcttctacagatga	gctcctcgcctacgggat	657616838	Original Exome Capture
SH282	gaaggtcggagtcaacggattgccttgaggacccatggtgg	gaaggtgaccaagttcatgctgccttgaggacccatggtga	ttggatctcactgtgctggg	659031282	RefSeqv1.1
SH59985	gaaggtcggagtcaacggattctagtggtcgttgtccgc	gaaggtgaccaagttcatgctctagtggtcgttgtccgt	cgtgtggcaagcaagagcg	659031658	Kronos Genome
SH858	gaaggtcggagtcaacggatttctatcccggagatcttcctc	gaaggtgaccaagttcatgcttctatcccggagatcttcctt	gtggaagcaccagacataca	663580858	RefSeqv1.1
SH567	gaaggtcggagtcaacggattgtacaacccagaagcgtttaac	gaaggtgaccaagttcatgctgtacaacccagaagcgtttaat	gccaatgcaatctccacgga	667220567	RefSeqv1.1
SH969	gaaggtcggagtcaacggattctggcctgcagaatatagatcac	gaaggtgaccaagttcatgctctggcctgcagaatatagatcat	agcacgtcatcatggcctc	680354969	Original Exome Capture
SH613	gaaggtcggagtcaacggattgggaaccccgacatccac	gaaggtgaccaagttcatgctgggaaccccgacatccat	ctgcgcagccagcttgtgcaa	681464613	Original Exome Capture
SH154	gaaggtcggagtcaacggattccgccgtgatgctgctcgtg	gaaggtgaccaagttcatgctccgccgtgatgctgctcgta	cgtacaccgggtagatga	686162154	Original Exome Capture
SH538	gaaggtcggagtcaacggattgaccagatcaggtcatc	gaaggtgaccaagttcatgctgaccagatcaggtcatt	gccagcgacaagtcc	714034538	Original Exome Capture
SH256	gaaggtcggagtcaacggatttcgtggatgctcgacctcc	gaaggtgaccaagttcatgcttcgtggatgctcgacctct	tactttatcaggctcagctcaggga	731226256	Original Exome Capture

2.3 Results

2.3.1 An initial forward screen of the tetraploid Kronos TILLING population identifies 21 lines with late and early senescing phenotypes.

To identify novel senescence-related mutant phenotypes, we carried out a phenotypic screen of 951 M_4 lines of the Kronos TILLING population at the John Innes Centre (JIC) in 2015. These lines, grown in 1 m plots as depicted in Figure 2-2, were scored for the timing of visual flag leaf and peduncle senescence. We identified a wide range of senescence timings, as shown in Figure 2-4, using wild-type Kronos as a control (purple). From these distributions, we identified the ten earliest senescing lines (yellow) and the eleven latest senescing lines (green). Most of the lines identified were both early or both late for flag leaf and peduncle senescence, though lines K2282 and K2438 showed opposing senescence phenotypes for flag leaf and peduncle senescence (Table 2-5).

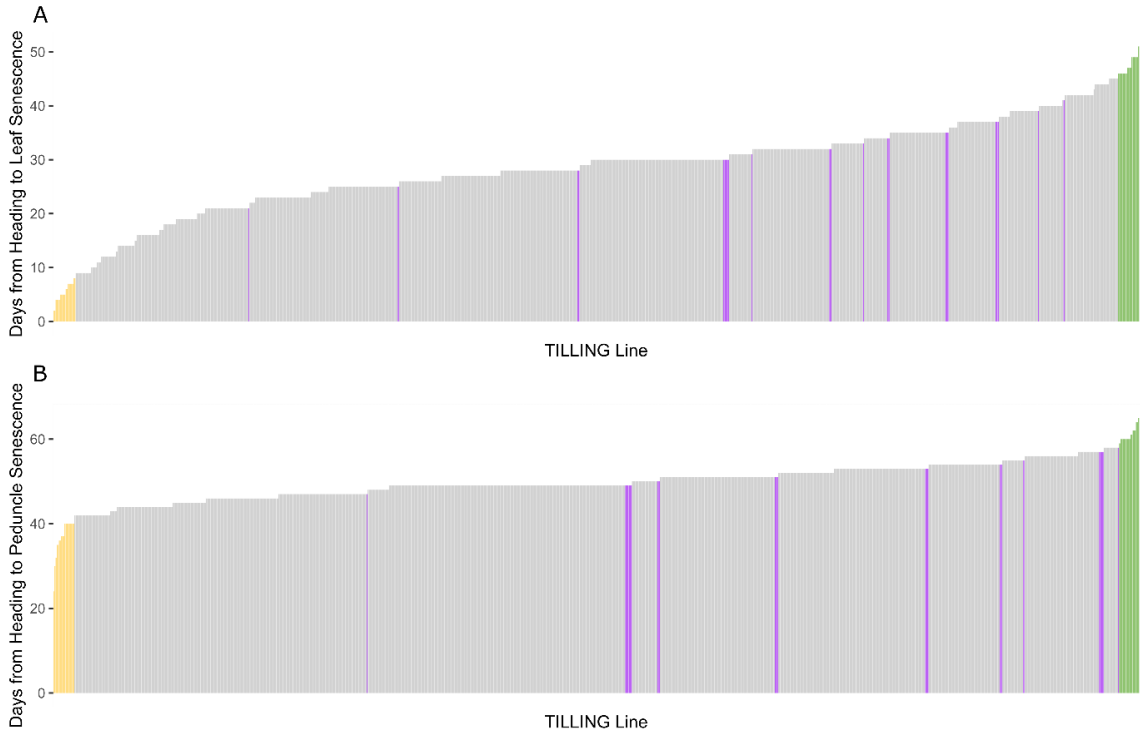

Figure 2-4: A forward screen of the Kronos TILLING population identified early and late senescing lines. 951 Kronos TILLING lines were screened for leaf (A) and peduncle (B) senescence onset timing. The 10 earliest senescing lines are highlighted in yellow, and the 11 latest senescing lines are highlighted in green. Wild-type Kronos was sown as a control, and the senescence timings for the wild-type replicates are shown in purple.

Table 2-5: Kronos TILLING lines selected from 2015 field trial based on senescence timing.

TILLING line	**Flag Leaf Senescence**	**Peduncle Senescence**	**Grown in 2016**	**Number of F_2 Populations**
K0331	Early	Early	No	N/A
K0344	Normal	Early	Yes	1
K0382	Early	Early	Yes	2
K0689	Early	Normal	Yes	2
K0714	Late	Late	Yes	2
K1107	Late	Late	Yes	2
K1426	Early	Early	Yes	2
K2170	Early	Early	Yes	2
K2222	Late	Late	Yes	1
K2235	Early	Early	Yes	2
K2282	Early	Late	Yes	2
K2293	Late	Late	Yes	2
K2425	Late	Late	Yes	1
K2438	Early	Late	Yes	2
K2468	Late	Late	Yes	1
K2711	Late	Late	Yes	2
K2849	Early	Late	Yes	2
K3085	Late	Late	No	N/A
K3117	Early	Normal	No	N/A
K3179	Late	Late	Yes	1
K4054	Late	Late	Yes	2

Of the 21 lines identified in the 2015 field trial, we successfully carried out back-crosses to the wild-type Kronos parent for 18 of the lines. These crosses generated F_2 populations which were then sown in 2016 in the field at JIC (Table 2-5). The F_2 populations were sown as individual plants (see diagram in Figure 2-2) and were thus scored for senescence onset as individual plants. Amongst each F_2 population, individual wild-type Kronos plants were also sown, allowing comparison of senescence timing in the TILLING lines to the wild-type control.

2.3.2 Mutations in the NAC transcription factor NAM-A1 lead to delayed senescence

After scoring the F_2 populations, we were able to identify only two TILLING lines which showed significantly delayed senescence, consistent with their original phenotypes in 2015 ($p < 0.001$, Kolmogorov Smirnov test adjusted for FDR; Figure 2-5). These populations, derived from TILLING lines K1107 and K2711, both contained point mutations in the NAC transcription factor *NAM-A1*. This gene is known to be a positive regulator of senescence, with previous mutant and knock-down lines also exhibiting substantially delayed senescence. We found that the *NAM-A1* mutation segregated with the delayed peduncle senescence phenotype ($p < 0.01$, Tukey's HSD; Figure 2-6), indicating that a mutation in this previously identified senescence regulator was causal for the phenotype observed. From these two lines, we observed that the mutations had a stronger effect on peduncle senescence than flag leaf senescence (Figure 2-5), and that a single missense mutation in *NAM-A1* in the K2711 line could drive a significant delay in senescence. These two observations will be followed up further in Chapter 3.

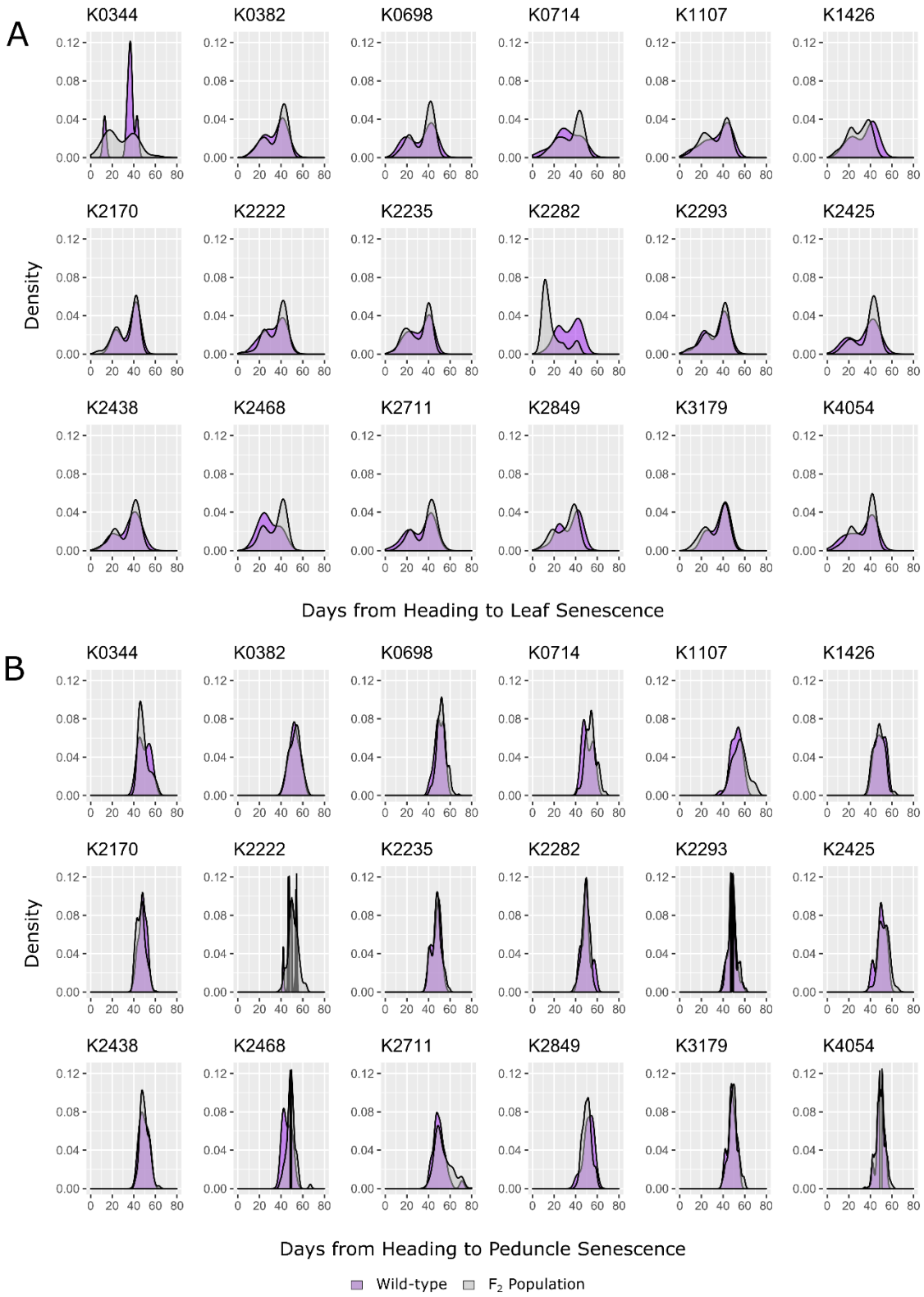

Figure 2-5: Most segregating F_2 populations derived from Kronos TILLING lines fail to show early or late senescence phenotypes. Leaf senescence (A) and peduncle senescence (B) was scored for individual plants in F_2 populations at the JIC in 2016. The distribution of senescence timings for the Kronos wild-type controls (purple) and the F_2 individuals (grey) are shown. The F_2 population only differs significantly from the wild-type for leaf senescence in line K2282 ($p < 0.001$, Kolmogorov Smirnov test adjusted for FDR). Note that the wild-type distributions for each population derive from wild-type plants grown within the respective population; as a result, the wild-type distributions will differ between each population.

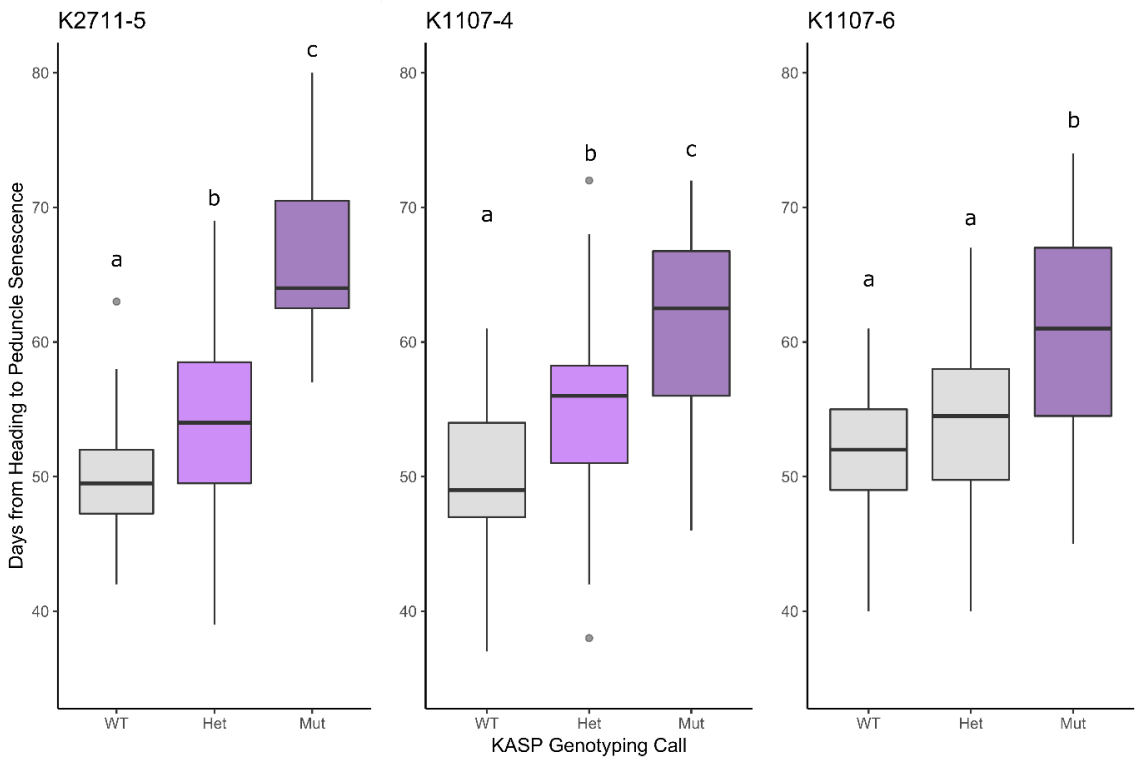

Figure 2-6: Mutations in *NAM-A1* explain the delay in peduncle senescence in K2711 and K1107. Markers specific to the *NAM-A1* mutations in K1107 and K2711 were used to genotype the F_2 populations in 2016. Significant delays in senescence ($p < 0.01$, Tukey's HSD) are indicated with letters; boxplots sharing the same letter (as in K1107-6) are not significantly different in senescence timing. Note that the second K2711 population, K2711-6, did not contain the *NAM-A1* mutation and was not significantly delayed in senescence (data not shown). N ranged from 27 (K2711 Mut) to 116 (K1107-6 Het).

2.3.3 The TILLING line K2282 segregates for a dominant chlorosis and early senescence phenotype.

Excluding the K2711 and K1107 populations, which contained mutations in a known senescence regulator, few of the F_2 populations tested in 2016 exhibited any kind of aberrant senescence phenotype. One that did, K2282, showed an extreme early flag leaf senescence ($p < 0.001$, Kolmogorov-Smirnov test, adjusted for FDR; Figure 2-5). Two independent F_2 populations derived from the K2282 TILLING line, hereafter referred to as K2282-A and K2282-B, both exhibited early flag leaf senescence in comparison to the Kronos wild-type controls. However, this phenotype was not restricted to solely flag leaf senescence as typically defined; that is, a clear progression of leaf chlorosis and necrosis often moving from the leaf tip towards the base. In the case of this mutant phenotype, the plants showed a high level of chlorosis—yellowing—in the flag leaf at anthesis (Figure 2-7A). We carried out chlorophyll quantification to show that the levels of all three main pigments, Chlorophyll A, Chlorophyll B, and carotenoids, are significantly reduced in the early senescing individuals ($p < 0.05$, Student's t-test; Figure 2-7B). This chlorosis phenotype was most pronounced in the interveinal regions of the plants, as shown in Figure 2-8.

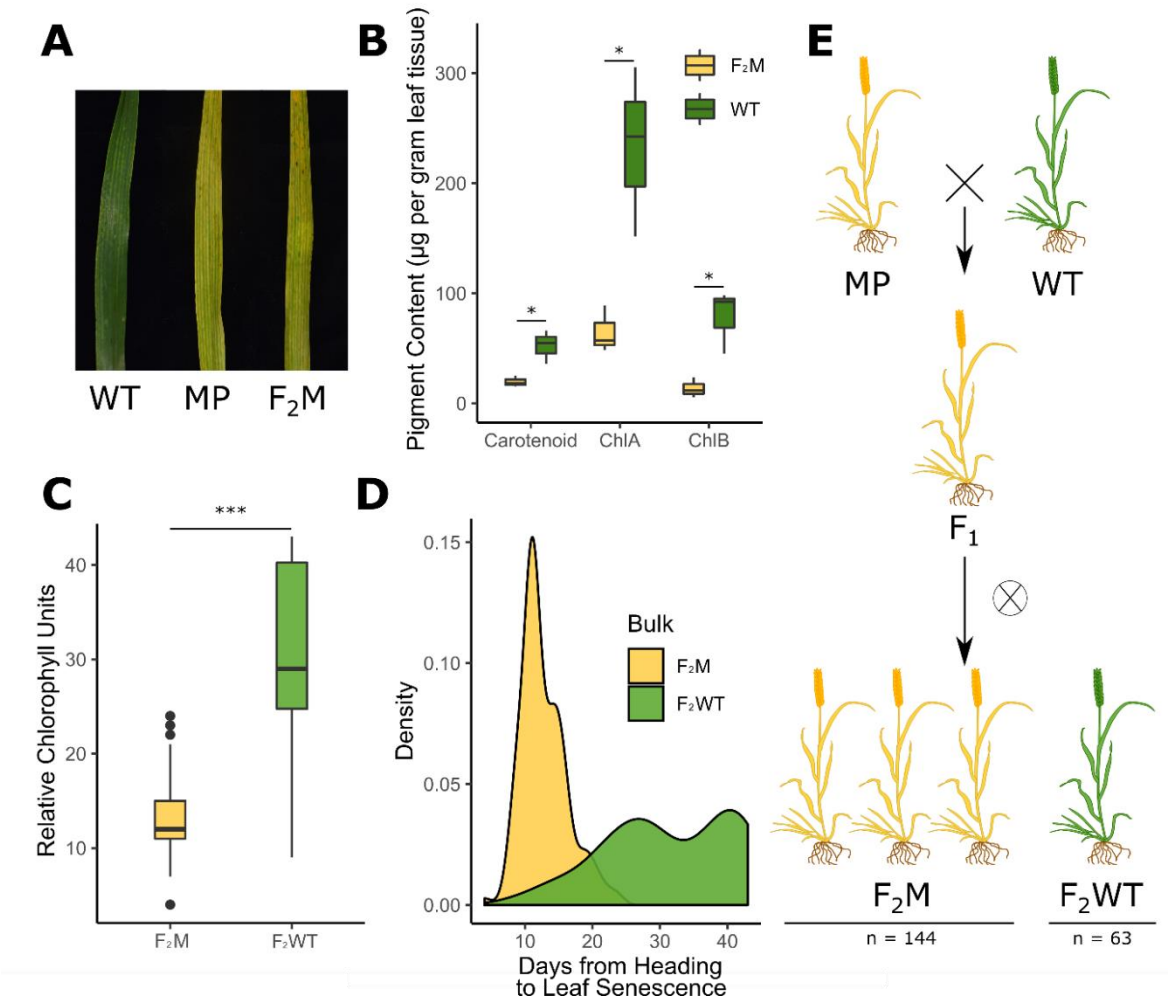

Figure 2-7: A premature yellowing phenotype from the Kronos TILLING population segregates as a single dominant locus. (A) F_2 populations of the K2282 Kronos TILLING line grown at the JIC in 2016 showed an early yellowing phenotype. Pigment content was measured in the yellow mutant plants (F_2M) compared to the wild-type plants (F_2WT) (B; n = 3 per genotype) and was also quantified using SPAD (C; n = 153 F_2M, n = 61 F_2WT). The yellow group (F_2M) senesced significantly earlier than the late bulk (F_2WT) (D; n = 148 F_2M, n = 56 F_2WT). Scoring of the plants for the yellow leaf phenotype demonstrated that the F_2 population was segregating 3:1 for the yellow trait, indicative of a dominant single locus (E; numbers are combined for both populations). F_2M and F_2WT refer to plants which are yellow and green, respectively, and which derive from the F_2 population (see bottom of E), while WT and MP refer to Kronos WT plants or M_4 K2282 plants, respectively (see top of E).

As the appearance of the chlorosis trait was so clear, we were able to assign the F_2 individuals into "yellow" and "green" groups (see Figure 2-7A). The visual scoring of these plants was confirmed using SPAD meter measurements, finding that the relative chlorophyll content of the yellow plants was significantly less than that of the green plants (Figure 2-7C). Having classified the F_2 individuals into these two binary groupings, we then reanalysed the flag leaf senescence data. Doing this, we found that the individuals which had chlorotic flag leaves around anthesis, i.e. were yellow, carried out flag leaf senescence significantly earlier than those individuals which did not exhibit chlorosis ($p < 0.001$, Kolmogorov-Smirnov test; Figure 2-7D). Importantly, to avoid confounding our senescence scores due to the already chlorotic nature of the yellow plants, we

scored flag leaf senescence solely based on the presence of necrosis in the flag leaf, not chlorosis as is typically used. Furthermore, we found that the yellow, chlorotic phenotype segregated in a manner not significantly different from a 3:1 ratio ($p = 0.07$, X^2; Figure 2-7E). This led us to the hypothesis that the chlorosis phenotype is caused by a single dominant locus, which we shall refer to as the *Yellow Early Senescence 1* (*YES-1*) locus.

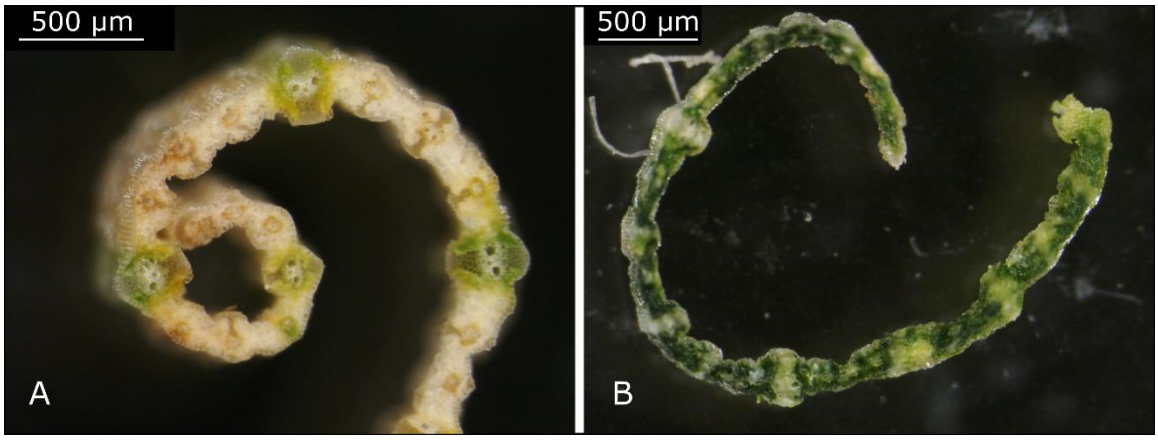

Figure 2-8: K2282 mutants have a striated inter-veinal chlorosis phenotype. Mutant plants (A) showed a characteristic loss of pigment in the inter-veinal regions, retaining green colouration in the main veins of the flag leaf, compared to the wild-type plants (B).

2.3.4 Bulked segregant analysis maps the YES-1 locus to chromosome 3A

Having previously split the K2282 F_2 populations into binary "yellow" and "green" categories, we then used these categories as the backbone of a bulked segregant analysis strategy (Figure 2-9). Tissue was sampled from yellow and green individual plants from each of the K2282-A and K2282-B independent F_2 populations. In order to maximise the fidelity of the analysis, we sampled tissue from as many individual plants as possible, obtaining tissue for 75 yellow individuals and 16 green individuals in K2282-A and 33 yellow and 22 green individuals in K2282-B. We then extracted DNA from the tissue samples and pooled the DNA to form four bulks— K2282-A_yellow, K2282-A_green, K2282-B_yellow, and K2282-B_green. Library preparation, exome capture, and Illumina sequencing were then carried out on these bulks, as described in detail in the methods.

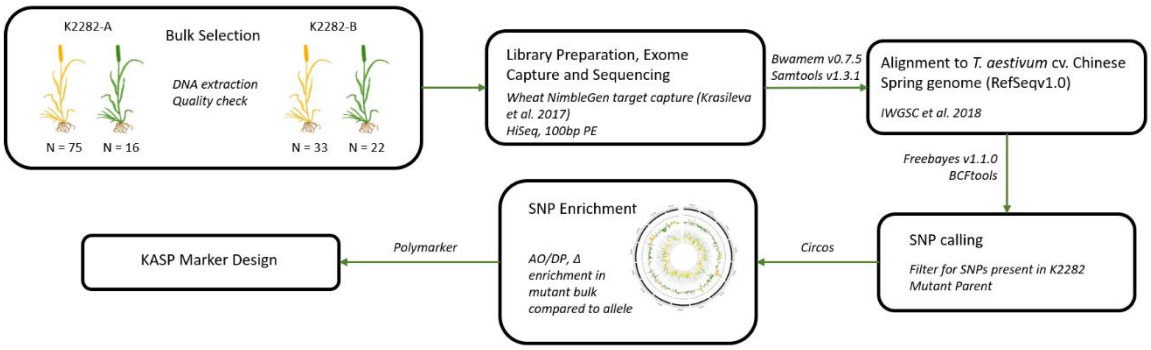

Figure 2-9: Pipeline for analysis of bulked segregant exome capture data. Exome capture was carried out on pooled DNA from yellow and green plants derived from two independent F_2 populations (K2282-A and K2282-B). The enrichment of TILLING SNPs in the yellow bulks was used to identify a region of interest on chromosome 3A. See methods for more details.

After obtaining the raw reads from the exome capture data for each of the four bulks, the reads were aligned against the RefSeqv1.0 genome and SNPs were called against the Chinese Spring reference. We achieved an average of 98% alignment to the RefSeqv1.0 genome, to a sequencing depth of approximately 45 X for three of the four bulks (Table 2-6). Importantly, the reference genome used is derived from a different variety than our TILLING mutants. As a result, many of the SNPs that we obtained from the alignment to RefSeqv1.0 were varietal SNPs, due to differences between the Chinese Spring and Kronos genomes. In order to account for this, we extracted only those SNPs which were identified in the original K2282 mutant line (Krasileva). In our exome capture dataset, derived from a back-crossed F_2 population, we recovered 1,548 of the original 3,060 SNPs—just over 50%. The loss of some SNPs was expected as a proportion of SNPs that were heterozygous in the M_2 generation, which was originally sequenced, would have been lost by the M_4 generation, which was scored in 2015. Additionally, the single backcross to wild-type Kronos would have further reduced the mutation burden in the F_2 population, eliminating approximately 50% of the heterozygous mutations present in the M_4 line.

Table 2-6: Total number of reads and mapped reads for each experiment, following removal of PCR duplicates. Average depth of coverage was calculated across the 119.2 Mb target exome space for the exome capture reads (see (Krasileva et al. 2017)). For the Kronos WT reads and the Chromosome 3A reads mapped to RefSeqv1.0, average coverage was calculated across the length of Chromosome 3A using 1 Mb genomic windows. Note that depth of coverage was not calculated for the Kronos alignment as the scaffolds are not associated to a chromosome.

Sample	**Total Reads**	**Mapped Reads**	**Percentage of reads mapped (%)**	**Sequencing Depth**	**Experiment**
K2282-B Green	45,224,648	44,344,449	98%	47	Exome Capture
K2282-B Yellow	41,112,551	40,353,815	98%	42	
K2282-A Green	44,755,140	43,848,635	98%	46	
K2282-A Yellow	82,272,560	80,607,768	98%	85	
Chromosome 3A	646,829,708	641,183,607	99%	N/A	Mapping to Kronos Assembly
	649,434,109	642,304,308	99%	82	Mapping to RefSeqv1.0
Kronos WT	1,438,873,905	1,208,044,854	84%	30	

We initially focused on the K2282-A population and calculated the ratio of the mutant allele (or alternate allele) over the depth of coverage for each SNP in each bulk. This is the "AO/DP" ratio depicted in Figure 2-10A, inner track. We then normalised the AO/DP ratio in the yellow bulk relative to the green bulk. This Δ value refers to the enrichment of the mutant allele in the yellow bulk compared to the green bulk for each SNP (Figure 2-10A, outer track). Where the Δ value is positive, the mutant allele is more common in the yellow bulk than the green bulk, and vice-versa when the Δ value is negative. As we hypothesize our causal mutation to be a single dominant locus, the yellow bulk should include both homozygous mutants and heterozygotes. As a result, we expected to see a SNP enriched to a delta value of approximately 0.66 in the yellow bulk and depleted to 0 in the green bulk. To reduce our false negative rate, we used a low threshold of 0.5 for the yellow bulk and identified a single region on chromosome 3A that was enriched for the mutant allele in the yellow bulk, and depleted in the green bulk (Figure 2-10B). We repeated this analysis on the second F_2 population, K2282-B, and obtained a consistent result pointing to the long arm of chromosome 3A (Figure 2-11).

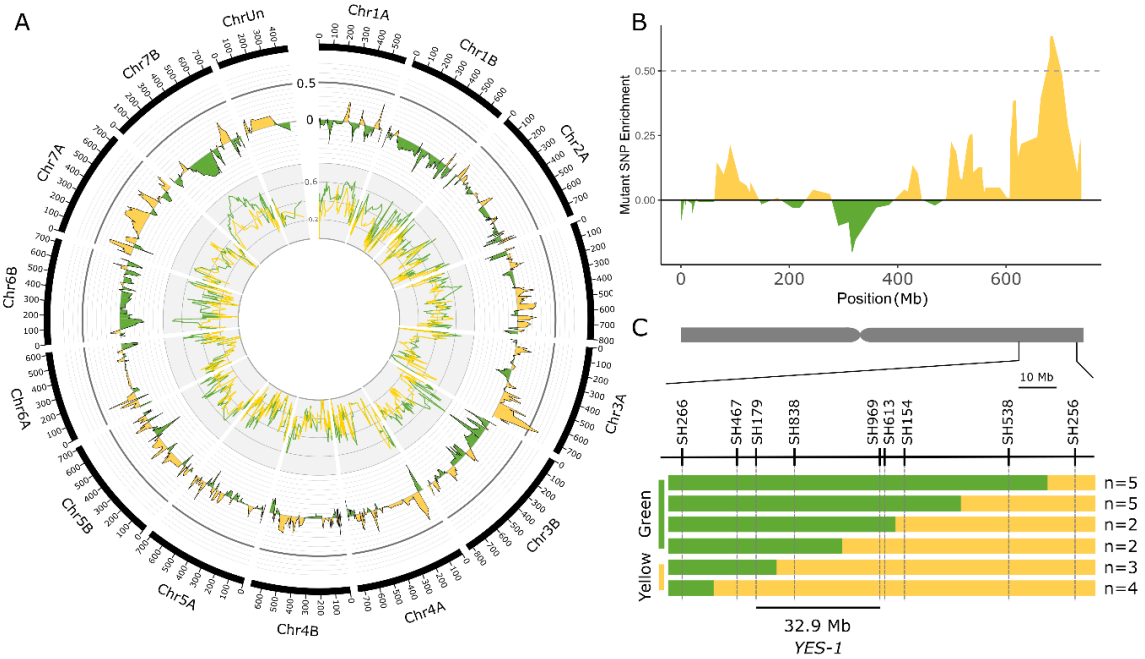

Figure 2-10: Bulked segregant analysis identifies the *YES-1* locus on chromosome 3A. Exome capture was carried out on yellow and green bulks from K2282 x Kronos F_2 populations grown at the JIC in 2016. The K2282-A yellow bulk (yellow line, inner track; smoothed to a moving average of 4) and green bulk (green line) were scored at each SNP locus identified for enrichment of the mutant allele. The level of enrichment in the green bulk was subtracted from the enrichment in the yellow bulk to obtain the Δ value (outer track; smoothed to a moving average of 4). A high Δ value, indicative of a region enriched for mutant alleles within the yellow bulks, was identified on the long arm of chromosome 3A (B; smoothed to a moving average of 4). Markers designed on known TILLING SNPs within this region mapped the *YES-1* locus to a 32.9 Mb interval within the F_2 population (C). Green bars indicate wild-type calls, while yellow bars indicate mutant or heterozygous calls. The numbers of individual plants that fell into each recombination interval are shown to the right. The chromosome scale in (A) is given in Mb.

To confirm this mapping, we developed a series of KASP markers against the SNPs located within and surrounding the region of interest (Figure 2-10C). We used DNA from the individual F_2 plants to confirm that the region of interest maps to the long arm of chromosome 3A. We required at least two independent recombination events to define the mapping interval and were able to narrow the *YES-1* region to 32.9 Mb, between markers SH179 and SH969. This region contains 345 gene models based on the RefSeqv1.1 gene annotation.

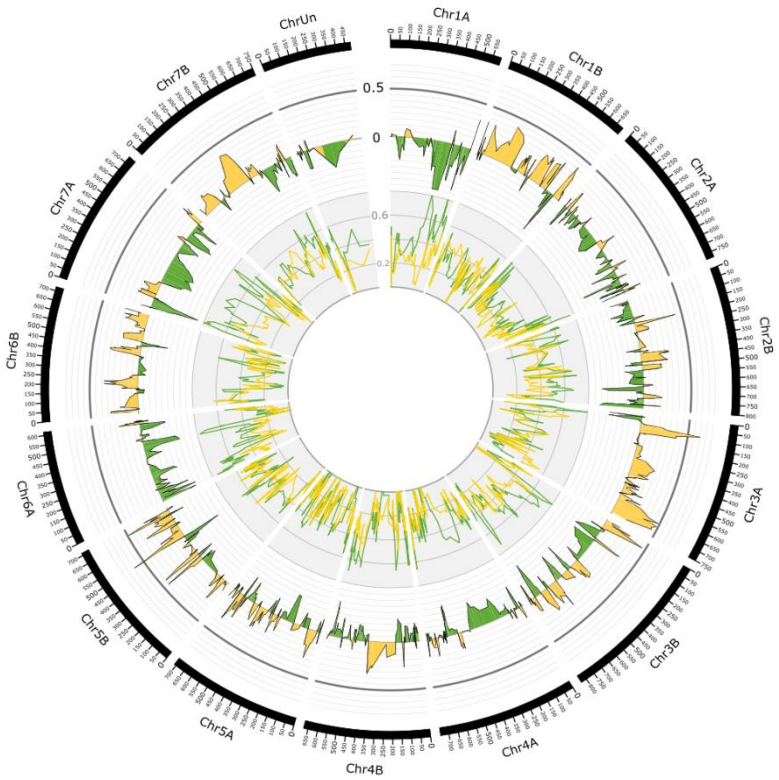

Figure 2-11: Genome-wide enrichment for mutant SNPs from the K2282-B population. The enrichment of the mutant allele at each SNP position (AO/DP) is shown on the inner track, with the green bulk highlighted in green and the yellow bulk highlighted in yellow. The Δ value at each SNP position is shown in the outer track, with regions enriched for the mutant allele shown in yellow, and those enriched for the wild-type allele shown in green. The cut-off of 0.5 for the Δ value is shown as a thick grey line in the outer track. In both the inner and outer tracks, data is presented as the moving average with an interval of 4. The Δ value approached the 0.5 cut-off in both the K2282-B and K2282-A (Figure 2 main text) populations only at the end of chromosome 3A. The chromosome scale is given in Mb.

2.3.5 Chlorosis caused by the YES-1 locus appears before anthesis

After mapping the *YES-1* locus to the long arm of chromosome 3A, we carried out further characterisation of the mutant phenotype. Individual lines which were genotyped as fully wild-type or mutant across the *YES-1* locus depicted in Figure 2-10C were grown at the JIC in 2018, in replicated 1 m rows. We found that the mutant lines contained significantly less chlorophyll A, B, and carotenoid pigment as early as the 3rd leaf stage of wheat development (Zadoks 13-14) ($p < 0.01$, Student's t-test; Figure 2-12A). This difference was larger at anthesis, by which time there was larger variation in pigment levels within the mutant plants ($p < 0.005$, Student's t-test).

A timecourse of chlorophyll content, measured using SPAD units, was carried out on the plants from 14 days pre-anthesis to 39 days post-anthesis. We found that chlorophyll levels were lower in the mutant lines from the beginning of the timecourse until up to 24 days post-anthesis (Figure 2-12B). The variation in pigment levels was immediately visible in the plants in the field, with the mutant lines appearing substantially yellower than the wild-type plants (Figure 2-12C, shown at 20 DPA). Chlorophyll levels peaked in both genotypes at 6 DPA (Figure 2-12B). The first significant

decline in chlorophyll levels from the peak was not seen in the wild-type plants until 24 DPA ($p < 0.01$, pairwise Wilcoxon rank-sum, adjusted for FDR). In contrast, the mutant plants showed a significant decline in their already low chlorophyll levels by 18 DPA ($p < 0.01$, pairwise Wilcoxon rank-sum, adjusted for FDR). However, despite the earlier onset of senescence seen in the mutant plants, they continued to lose chlorophyll until the final timepoint (39 DPA), in line with the wild-type plants. This suggests that while senescence onset is premature in the mutant plants, the progression of senescence may be correspondingly slowed such that complete flag leaf senescence occurs at the expected time.

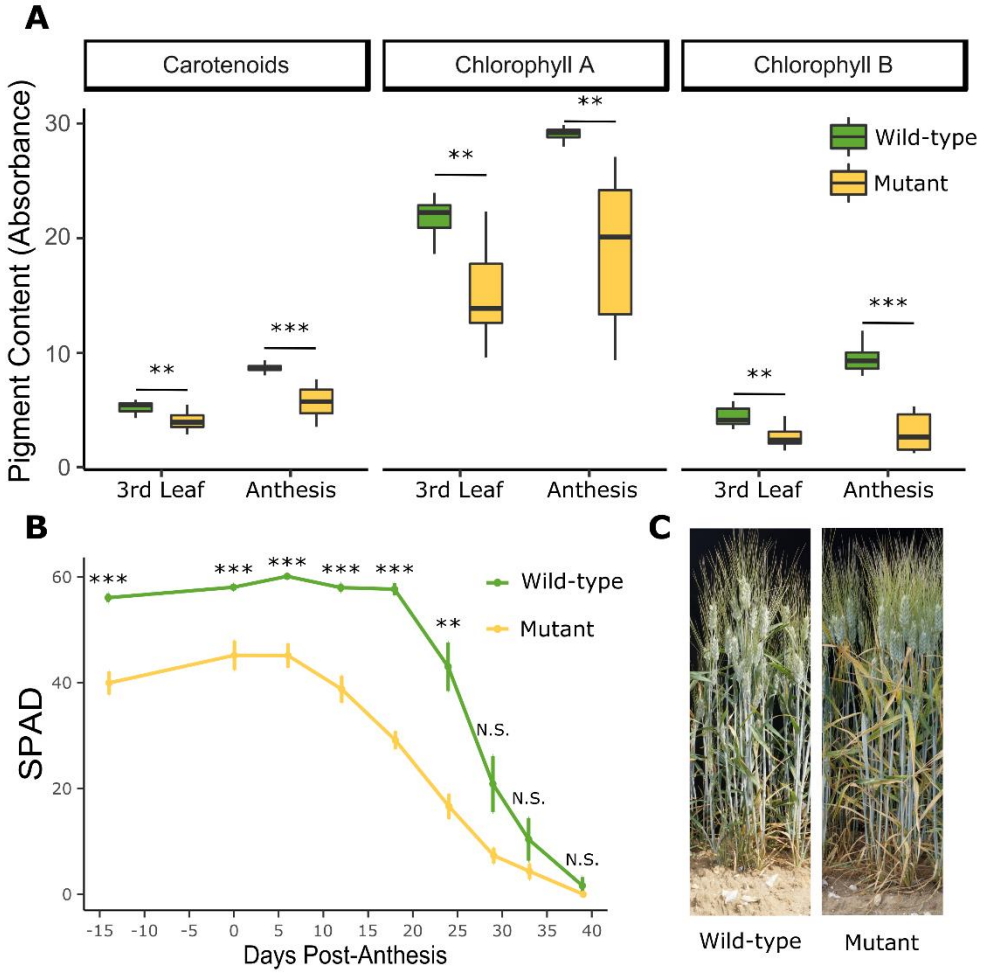

Figure 2-12: The *YES-1* locus causes lower chlorophyll levels before anthesis and earlier onset of senescence. The early chlorosis phenotype first observed at the JIC in 2016 was recapitulated in the JIC 2018 field trials. Pigment content in the mutant lines is significantly lower at the third leaf stage (Zadoks 13-14, 24 days before anthesis) and becomes more extreme by anthesis (A; **, $p < 0.01$; ***, $p < 0.001$ Student's T-test; N = 8 for Mutant and 10 for Wild-type). Relatively chlorophyll content, as measured with a SPAD meter, is significantly decreased in the mutant lines before anthesis, and remains so until 29 days post-anthesis (B; **, $p < 0.01$; ***, $p < 0.001$ pairwise Wilcoxon rank sum, adjusted for FDR; N = 8 for Mutant, 10 for Wild-type; error bars are standard error of the mean). The yellowing phenotype in the leaves was clear in the field at 20 DPA (C).

Alongside further characterisation of the visual symptoms of senescence, we also investigated whether there were any changes in leaf nutrient composition in the mutant plants compared to the wild-type plants. At the 3rd leaf stage (Zadoks 13-14), we found that the chlorotic leaves contained significantly less magnesium than the wild-type leaves ($p < 0.05$, Student's t-test; Figure 2-13A). By anthesis, all four measured minerals were lower in the mutant, chlorotic flag leaves than in the wild-type flag leaves ($p < 0.05$ for Mg, Fe, Zn and $p = 0.05$ for Mn, Student's t-test; Figure 2-13B).

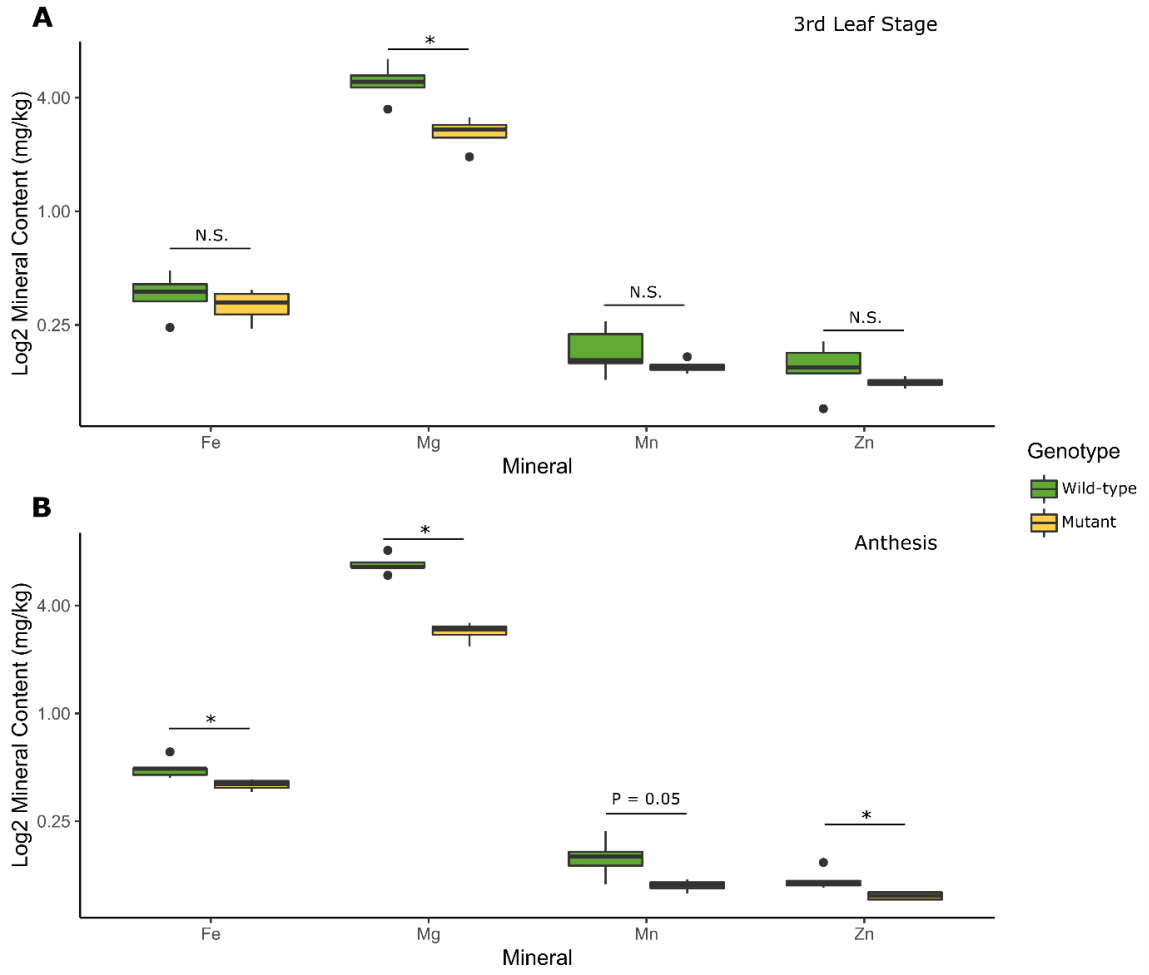

Figure 2-13: Mutant plants carrying the *YES-1* locus have lower leaf mineral content. Measurements of leaf mineral content were carried out at the 3rd leaf stage (A; Zadoks 13-14) and anthesis (B) for wild-type and mutant plants at the JIC in 2018. N = 5 for all genotypes and timepoints.

2.3.6 The chlorosis phenotype is inconsistent across environments.

The same lines were grown at UC Davis, in California, during the 2016-2017 field season. Unlike at the JIC, we saw no chlorosis or senescence phenotype in the mutant lines (Figure 2-14A and B). This suggested that the mutant phenotype we observed at the JIC, underpinning the *YES-1* locus, was environmentally dependent. As we had observed that the mineral content in the mutant flag leaves was lower than in the wild-type (Figure 2-13), we hypothesized that the environmental dependency of the phenotype may be due to variable nutrient levels in the soil. The observed

interveinal chlorosis (Figure 2-8) closely matches that observed in plants under magnesium deficient conditions, supporting this hypothesis (Snowball and Robson 1991).

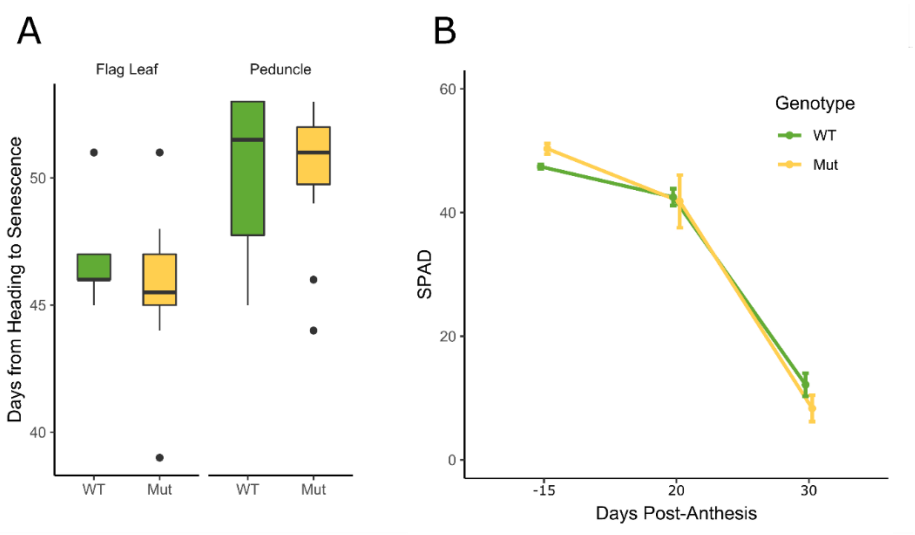

Figure 2-14: No visual senescence or chlorosis was observed in Davis, CA. Rows of F_3 individuals which were either fully wild-type or mutant across the *YES-1* region were phenotyped for visual senescence onset (A) and relative chlorophyll units using SPAD (B; error bars are standard error of the mean). No significant difference in senescence onset or chlorosis was observed at any stage. For visual scores, N = 24 for both genotypes. For SPAD scores, N = 19-24 for Mutant and 49-52 for Wild-type.

To test this, we grew F_3 plants which were fully mutant across the *YES-1* locus under glasshouse conditions in three different soil types—standard cereal mix, soil taken from the JIC field site in 2017, and horticultural sand supplemented with Hoagland solution. If the mutant phenotype was solely due to soil composition and nutrient content, we expected to see the mutant phenotype recapitulated in the JIC soil. However, we observed no signs of chlorosis or premature senescence under any of the three conditions (Figure 2-15A).

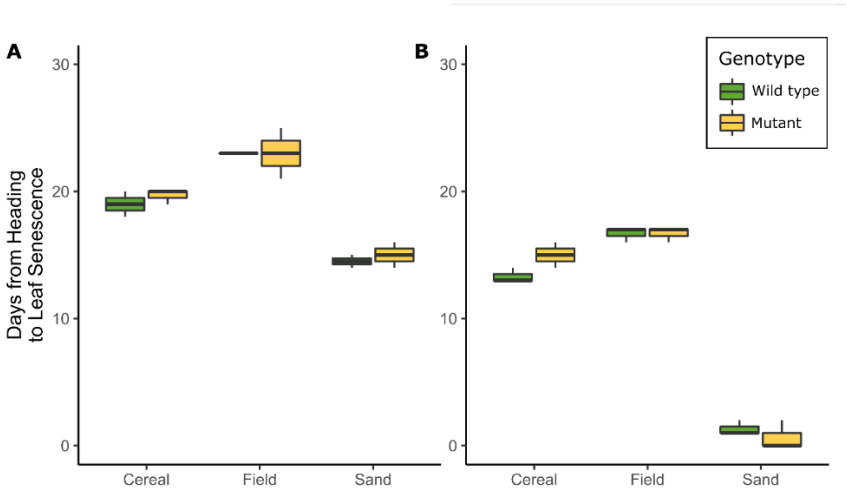

Figure 2-15: No soil or water conditions could recapitulate the early chlorosis phenotype in the glasshouse. Mutant and wild-type plants were grown under glasshouse conditions in three soil conditions (Cereal Mix, JIC 2017 Field soil, and Horticultural Sand supplemented with Hoagland Solution) and under normal (A) and low water (B) conditions. No significant early onset of chlorosis or senescence was observed for the mutant plants in any condition (Student's t-test). N=3 for each genotype and condition combination.

We then considered the possibility that other environmental factors were influencing the presentation of the mutant phenotype. We obtained rainfall and temperature data from both field sites— the JIC for field seasons 2016, 2017, and 2018, when the mutant phenotype was observed, and the UC Davis field site for the 2016-2017 field season, where no mutant phenotype was observed (Figure 2-16). We found that the trials carried out at UC Davis in 2016-2017 received substantially more rainfall between sowing and heading than did any of the JIC trials (Figure 2-16, Table 2-7). The three JIC trials received between 88 and 115 mm of rain in the period between sowing and heading, while the UC Davis trial received 660 mm of rain in the same developmental period.

Overall, from sowing to harvest, the UC Davis trial received between 400 and 600 mm more rain than the JIC trials (Table 2-7). Based on this data, we then hypothesized that the excessive rain experienced by the UC Davis trials may have influenced the mutant phenotype. We then tested the mutant plants under water-replete and water-depleted conditions in the glasshouse, using the same three soil combinations as in the previous experiment. Again, we were unable to recapitulate the early senescence and chlorosis phenotype when the mutant lines were placed under water-depleted conditions (Figure 2-15B). Based on these results, we were unable to determine why the mutant phenotype was not apparent in either the UC Davis field trial or when grown under glasshouse conditions.

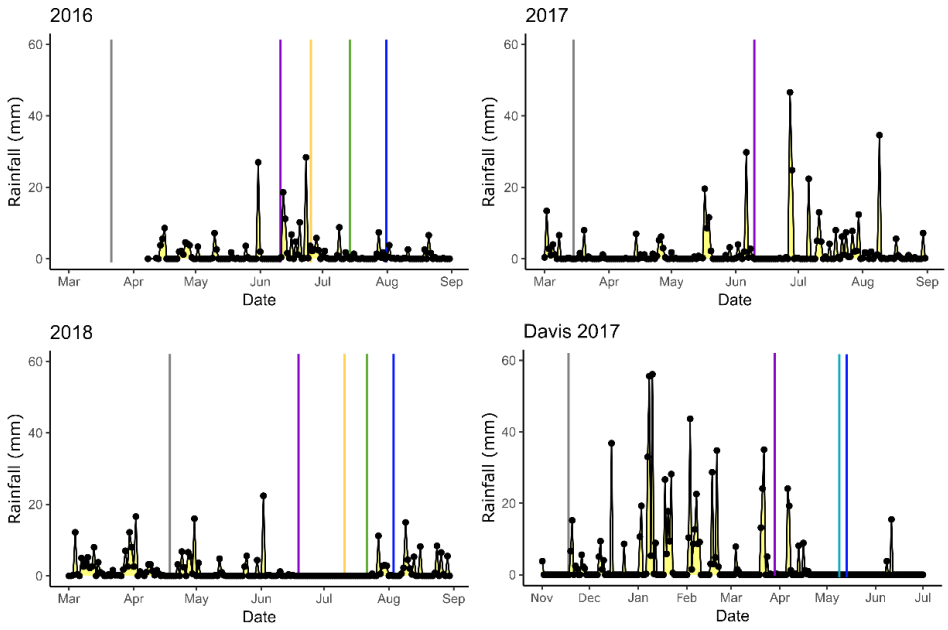

Figure 2-16: Rainfall patterns for field trials at the JIC and Davis, CA. Total rainfall is shown per day for the approximate growing season in each field season (black line with yellow fill). The heading date for each trial is indicated with a purple vertical line; leaf senescence onset is indicated with a green or yellow line, to show the difference in senescence timings for the green and yellow plants, in the JIC, and with a light blue line in Davis, as no difference was observed. Peduncle senescence timing is indicated with a dark blue line. For all years, sowing date is shown as a grey line. Only sowing and heading dates are available for 2017 as senescence onset was not scored for this trial. Rainfall is reported in mm (y-axis). Rainfall data for the 2016 JIC field trials is only available from 08/04/2016.

Table 2-7: Rainfall levels and duration of key developmental stages during the field seasons.

Developmental Stages	Field Trial	Precipitation (mm)	Duration (days)
Sowing to Heading	JIC 2016	104.2	70
	JIC 2017	114.4	68
	Davis 2017	661.1	135
	JIC 2018	88.2	64
Heading to Leaf Senescence	JIC 2016 WT	115.4	33
	JIC 2016 Mut	87	14
	Davis 2017	64.7	47
	JIC 2018 WT	0	32
	JIC 2018 Mut	0	20
Heading to Peduncle Senescence	JIC 2016	128.2	50
	Davis 2017	64.7	51
	JIC 2018	21	43
Sowing to Harvest	JIC 2016	221.8	137
	JIC 2017	331.4	136
	JIC 2018	148.6	129

	Davis 2017	744.8	212

2.3.7 Fine-mapping reduced the YES-1 locus to 4 Mb on chromosome 3A.

Having previously mapped the *YES-1* locus to a 32.9 Mb region, we then looked to fine-map the region further using a new set of recombinants. However, as is made clear in Figure 2-10C, there were few remaining SNPs within the *YES-1* region to aid in fine-mapping. As a result, it was necessary to identify new SNPs between the mutant and wild-type lines. As the first analysis had used exome capture to focus on SNPs present in and near the coding regions of the genome, we now looked to identify SNPs in non-coding regions of the genome. To that end, we collaborated with the group of Jaroslav Doležel (Institute of Experimental Botany, Olomouc, Czech Republic) to purify chromosome 3A from the M_5 K2282 mutant lines, using flow cytometry sorting. The M_4 lines were chosen as they would carry all of the SNPs that were possibly present within the F_2 population, allowing us to maximise the recovery of potentially informative markers. The flow-sorted population of 3A chromosomes partially overlapped on a bivariate flow karyotype DAPI vs GAA-FITC with the population of 7A chromosomes (Figure 2-3). A purity of 80% for the 3A chromosomes was obtained, with the remaining 20% principally made up of the 7A chromosomes. For sequencing, 3 batches of 30,000 chromosomes were prepared via flow cytometry. The independent samples were then amplified via PCR, producing a total of 4.51 µg of DNA for sequencing.

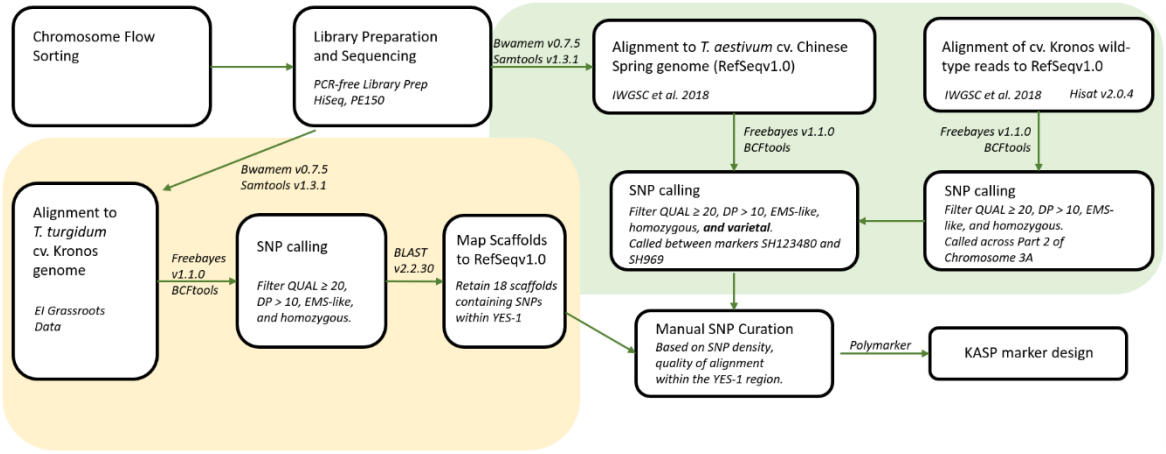

Figure 2-17: Pipeline for analysis of chromosome flow-sorting data. DNA obtained from K2282 chromosome 3A was sequenced using Illumina sequencing. Reads were processed using two approaches: alignment to the cultivar Kronos assembly (yellow) and alignment to the reference Chinese Spring genome (green). SNPs obtained from both methods were used for KASP marker design. Bioinformatics tools used in each step are shown next to the corresponding arrow.

After obtaining the raw reads for the flow-sorted chromosomes, the reads were prepared and aligned to the A and B genomes of the RefSeqv1.0 genome, as detailed in Figure 2-17. We found that, in total, 60.38% of reads aligned successfully to chromosome 3A, with 25.37% aligning to chromosome 7A. This primary contamination matched that expected based on the flow cytometry

results. The remaining 14.25% of reads mapped against other regions of the genome, most likely corresponding to highly repetitive regions present in multiple locations in the genome. On average, we achieved coverage of 82 X across chromosome 3A, when calculated using genomic windows of 1 Mb (Table 2-6).

Having successfully aligned the reads from chromosome 3A against the main reference genome, we then set about identifying new SNPs that would allow us to design markers to refine the *YES-1* locus. To do this, we took two main approaches, detailed in Figure 2-17. Our first approach used the variety-specific Kronos assembly to identify SNPs in non-coding regions, which would therefore be less conserved between Kronos and Chinese Spring (Figure 2-17, yellow). Our second approach utilised the physical contiguity of the RefSeqv1.0 genome to identify new SNPs in genes physically located within and near the *YES-1* locus (Figure 2-17, green). By focusing only on coding regions, which are more conserved across varieties, we would be less likely to identify spurious SNPs that were varietal rather than EMS-induced.

2.3.7.1 SNP calling against the Kronos Assembly

We initially carried out the first approach, aligning reads from the mutant K2282 chromosome 3A against the draft Kronos assembly (10+ Wheat Genomes Project 2016). We called SNPs against the assembly, and filtered for homozygous, EMS-like SNPs which passed minimum quality (QUAL ≥ 20) and depth (DP > 10) thresholds. This identified 75,034 SNPs across the entirety of the Kronos genome. At this stage, we then needed to filter only the SNPs that were located within and near the *YES-1* locus. However, the scaffolds of the Kronos genome are not anchored to physical positions. To isolate only those SNPs within the *YES-1* region, we carried out a BLAST between all of the Kronos scaffolds which contained SNPs, and the gene sequences present within the *YES-1* region in the RefSeqv1.0 genome. We used coding sequences for the comparison as they have a higher level of similarity between varieties, facilitating the BLAST alignments. This method also reduced the chance of obtaining high levels of false positives due to scaffolds aligning to highly repetitive sequences within the *YES-1* region that are also present elsewhere in the genome.

Based on previously identified recombination, we extracted SNPs from genes spanning a region 3 MB upstream and downstream of markers SH179 and SH838, approximately 16 Mb in size. We identified 18 unique scaffolds in the Kronos genome which contained both a homozygous EMS-like SNP and at least one gene located in this physical interval (Table 2-8). Of the 222 genes present within this 16 Mb region, we identified 26 genes within these 18 scaffolds. Some genes may not have been identified if they fell in a scaffold that contained no high-quality SNP, if they were split across multiple scaffolds, or if they were missing from the Kronos genome entirely. The identified SNPs were then manually curated, removing any which appeared in regions of unexpectedly high SNP density (greater than 1 SNP per 25-40 Kb; (Krasileva et al. 2017)), which were likely to be false positives. In particular, we identified a single region in one of the Kronos

scaffolds, number 008254, which contained 66 SNPs within a 47 Kb region, substantially higher than the expected frequency of 23 EMS-induced SNPs per Mb (Krasileva et al. 2017).

This left a final list of 16 scaffolds containing high-quality EMS-like SNPs, covering approximately 2.2 Mb of genome sequence (Table 2-8). Most of the SNPs identified fell outside of annotated genes, suggesting that our approach was successful in identifying novel SNPs in non-coding regions of the *YES-1* locus. We also recovered known SNPs in our dataset, principally the SNPs underlying markers SH838 and SH179. KASP markers were designed for a subset of these SNPs, based on distribution across the *YES-1* region. After running these markers on the original F_2 recombinants, and a second set of recombinants in the F_3 generation, we were able to map the *YES-1* locus further to a 6.6 Mb region, between markers SH123480 and SH59985 (Figure 2-18).

Table 2-8: SNPs identified against the Kronos Genome. Each SNP is indicated with its location on the respective Kronos scaffold, the type of EMS SNP (G>A or C>T), and the depth of coverage at each SNP. The SNP is also show in relation to the genes which were identified on the scaffold using BLAST; some scaffolds contain more than one gene and as a result the SNPs are repeated for each gene. SNPs that were used for KASP marker design are also indicated, as are SNPs that were previously used as KASP markers from the exome capture data (SH179 and SH838).

Kronos Scaffold	Location of SNP	Type	Depth of Coverage	Gene	Gene Start	Gene End	Orientation of Gene	SNP relation to gene	KASP Marker
Triticum_turgidum_Kronos_EIv1.1_scaffold_038440	22962	G>A	132	TraesCS3A01G397300	38722	39595	+	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_051420	17119	G>A	93	TraesCS3A01G397400	50086	50739	-	Downstream	SH17119
Triticum_turgidum_Kronos_EIv1.1_scaffold_051420	17119	G>A	93	TraesCS3A01G397600	16870	17568	+	In Gene	SH17119
Triticum_turgidum_Kronos_EIv1.1_scaffold_011070	133732	C>T	98	TraesCS3A01G399300	113656	113859	+	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_057496	12729	C>T	100	TraesCS3A01G401200	3654	4466	-	Upstream	SH179
Triticum_turgidum_Kronos_EIv1.1_scaffold_012118	44138	G>A	28	TraesCS3A01G401800	145669	146737	+	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_044009	3007	C>T	48	TraesCS3A01G402400	27036	28520	+	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_070196	25155	C>T	76	TraesCS3A01G402700	12487	13271	+	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_019229	128018	G>A	45	TraesCS3A01G403000	109526	110252	+	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_004188	141367	G>A	28	TraesCS3A01G404400	72793	73368	+	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_004188	196608	G>A	32	TraesCS3A01G404400	72793	73368	+	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_004188	269110	G>A	64	TraesCS3A01G404400	72793	73368	+	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_030736	46408	C>T	19	TraesCS3A01G406600	12110	14029	-	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_030736	86495	C>T	37	TraesCS3A01G406600	12110	14029	-	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_000337	3898	G>A	43	TraesCS3A01G407300	433422	434090	+	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_000337	65290	G>A	121	TraesCS3A01G407300	433422	434090	+	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_000337	123480	G>A	182	TraesCS3A01G407300	433422	434090	+	Upstream	SH123480
Triticum_turgidum_Kronos_EIv1.1_scaffold_000337	255090	G>A	410	TraesCS3A01G407300	433422	434090	+	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_000337	358018	G>A	48	TraesCS3A01G407300	433422	434090	+	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_000337	374386	G>A	396	TraesCS3A01G407300	433422	434090	+	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_000337	378224	G>A	19	TraesCS3A01G407300	433422	434090	+	Upstream	

Kronos Scaffold	Location of SNP	Type	Depth of Coverage	Gene	Gene Start	Gene End	Orientation of Gene	SNP relation to gene	KASP Marker
Triticum_turgidum_Kronos_EIv1.1_scaffold_000337	412023	G>A	23	TraesCS3A01G407300	433422	434090	+	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_000337	514840	G>A	191	TraesCS3A01G407300	433422	434090	+	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_426139	344	C>T	15	TraesCS3A01G409000	862	1542	-	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_064761	12835	G>A	102	TraesCS3A01G411800	16203	17105	-	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_047726	25045	C>T	191	TraesCS3A01G413900	21581	24114	-	Upstream	SH838
Triticum_turgidum_Kronos_EIv1.1_scaffold_047726	40197	C>T	19	TraesCS3A01G413900	21581	24114	-	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_047726	50838	C>T	31	TraesCS3A01G413900	21581	24114	-	Upstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_016263	59609	C>T	29	TraesCS3A01G416800	68079	69020	-	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_016263	59985	C>T	32	TraesCS3A01G416800	68079	69020	-	Downstream	SH59985
Triticum_turgidum_Kronos_EIv1.1_scaffold_016263	59609	C>T	29	TraesCS3A01G416600	54461	54952	+	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_016263	59985	C>T	32	TraesCS3A01G416600	54461	54952	+	Downstream	SH59985
Triticum_turgidum_Kronos_EIv1.1_scaffold_013527	159103	C>T	45	TraesCS3A01G418100	18253	19646	+	Downstream	
Triticum_turgidum_Kronos_EIv1.1_scaffold_013527	159103	C>T	45	TraesCS3A01G418200	20113	20469	+	Downstream	

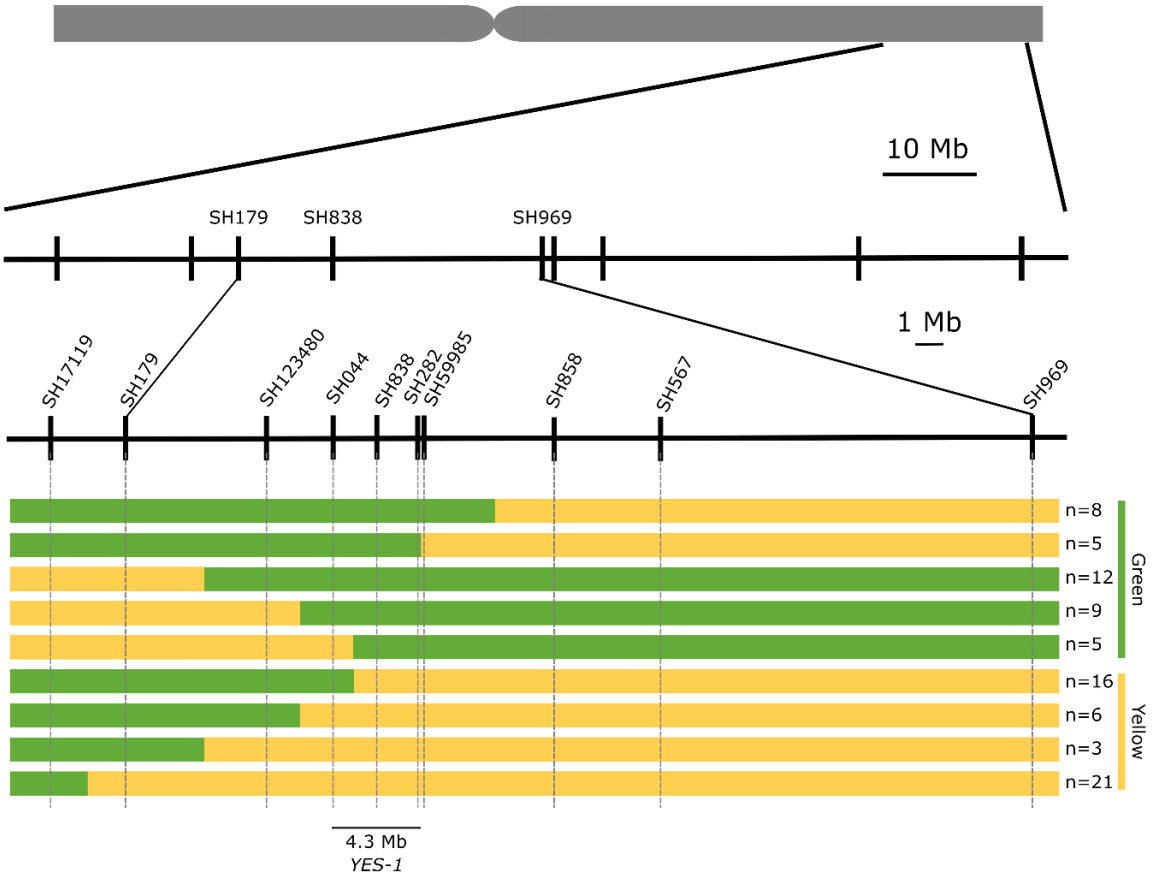

Figure 2-18: The *YES-1* locus fine-maps to a 4.3 Mb region. New markers were developed based on SNPs identified in non-coding regions of the M_5 K2282 mutant. Phenotypic data for flag leaf chlorosis from the JIC 2017 and 2018 field trials were used to classify recombinant lines as "green" or "yellow". These markers mapped *YES-1* to a 4.3 Mb interval between markers SH044 and SH59985.

2.3.7.2 SNP calling against the Chinese Spring RefSeqv1.0

We then carried out our second approach for identifying novel SNPs, using the Chinese Spring-derived RefSeqv1.0 genome (see pipeline schematic in Figure 2-17, green). Our first task was to obtain a set of varietal SNPs that we could use to filter out unwanted varietal SNPs from our dataset. To do this, we aligned reads derived from wild-type Kronos to the RefSeqv1.0 genome, obtaining coverage of approximately 30 X across chromosome 3A. SNP calling was then carried out, obtaining a total of 968,482 homozygous varietal SNPs meeting minimum quality and depth parameters across the second part of chromosome 3A. This segment of the chromosome includes the *YES-1* locus, and with this set of varietal SNPs, we could begin to call SNPs directly between the mutant chromosome 3A, derived from Kronos, and the RefSeqv1.0 genome, derived from Chinese Spring.

After filtering for quality and depth, we then filtered the SNPs to remove any that were shared with the set of varietal SNPs obtained above. Following this, we retained a set of 7,153 SNPs between markers SH123480 and SH969, encompassing approximately 30 Mb. This was a substantially

higher SNP density than would be expected in the Kronos TILLING lines, which have a known mutation density of 23 mutations/Mb (Krasileva). However, upon mapping the physical locations of the identified SNPs, we observed that the SNPs were present in a highly irregular pattern. This suggested that many of the SNPs were obtained as a result of mis-mapping to highly repetitive regions, present in the non-coding regions of the genome.

In order to eliminate the noise presented by mis-mapping to these highly repetitive regions, we extracted SNPs present within 1 Kb up or downstream of an annotated gene in the *YES-1* region. This subset was then manually curated to identify a small set of 15 SNPs that had very high quality and high depth of coverage (Table 2-9). Three of the fifteen SNPs were located within coding regions, while the remaining SNPs fell in either introns (5), promoters (5) or the 3' UTR (2). Two of the coding region variants, SH838 and SH858, caused missense mutations with low SIFT scores (0.00 and 0.03, respectively), while the third, SH567, encoded a synonymous mutation. Using markers designed from a subset of these SNPs, we were then able to fine-map the *YES-1* region to a final 4.3 Mb interval, between markers SH044 and SH59985 (Figure 2-18).

Table 2-9: SNPs identified between the K2282 chromosome 3A sequence and RefSeqv1.0. Each SNP is provided alongside its location on chromosome 3A, the depth of coverage at each SNP, and the type of EMS SNP (G>A or C>T). The nearest gene(s), within 1 Kb of the SNP, are also listed. Where a SNP was within 1 Kb of more than one gene, both genes are listed here. SNPs that were used for KASP markers are indicated.

SNP Position	Type	Quality	Depth of Coverage	Nearest Gene	Gene Start	Gene End	Orientation of Gene	SNP relation to gene	KASP Marker
654214125	C>T	1742.28	53	TraesCS3A02G409800	654214362	654216953	+	Promoter	
654781044	C>T	2525.84	74	TraesCS3A02G410700	654781914	654785094	+	Promoter	SH044
655634840	G>A	109.768	15	TraesCS3A02G411700	655635825	655637330	-	3' UTR	
657616838	C>T	6099.96	191	TraesCS3A02G413900	657613374	657615907	-	Promoter	
657616838	C>T	6099.96	191	TraesCS3A02G414000	657616083	657620810	-	1st exon	SH838
659031282	C>T	960.391	29	TraesCS3A02G416700	659026954	659038375	+	2nd intron	SH282
659031658	C>T	1058.26	32	TraesCS3A02G416700	659026954	659038375	+	2nd intron	SH59985
660869609	C>T	6383.87	218	TraesCS3A02G419100	660864796	660872768	+	7th intron	
663580858	C>T	5425.45	167	TraesCS3A02G422800	663579033	663599409	-	3rd exon	
664911627	C>T	2943.42	92	TraesCS3A02G423300	664912319	664912531	+	Promoter	
667220567	C>T	1980	68	TraesCS3A02G424600	667220385	667220843	+	1st exon	SH567
667220567	C>T	1980	68	TraesCS3A02G424500	667219018	667220345	+	Downstream of 3' UTR	SH567
671647297	C>T	2379.2	72	TraesCS3A02G428000	671639069	671646487	-	Promoter	
672089627	C>T	6236.04	182	TraesCS3A02G429000	672088991	672089101	+	3' UTR	
674263850	C>T	4796.97	138	TraesCS3A02G432900	674255517	674279708	+	1st intron	
676234801	C>T	2844.75	93	TraesCS3A02G434400	676207398	676243866	-	4th intron	
680354969	C>T	1959.56	62	TraesCS3A02G436500	680352382	680359377	-	4th exon	SH969

2.3.8 59 high-confidence gene models are located within the fine-mapped YES-1 locus.

After fine-mapping the *YES-1* locus to 4.3 Mb, we then turned our attention to the genes located within this region. We identified 59 high confidence gene models in the RefSeqv1.1 gene annotation, listed in Appendix 8.3.1, Table 8-1. Of these 59 genes, one gene contained a missense mutation, the SNP underpinning marker SH838. Of the remaining genes, based on the high-quality SNPs identified above, only two contain SNPs near the gene body. One, TraesCS3A02G411700, contains a SNP within the annotated 3' UTR, while the missense mutation underpinned by marker SH838 is also located within 5 Kb upstream of TraesCS3A02G413900 (Table 2-9). Of the remaining 56 genes, no SNPs were identified either in or around the gene sequences with the remainder of the identified SNPs falling in non-coding regions (Table 2-8).

To determine which of the 59 genes were promising candidate genes for the chlorosis and early senescence phenotype, we initially investigated their expression profiles. Using a developmental timecourse of RNA-Seq expression data (Ramírez-González et al. 2018), we identified 25 genes within this region which are expressed above 0.5 transcripts per million (TPM) in at least one stage of leaf or shoot development, which would be consistent with our observation of a leaf chlorosis phenotype (Figure 2-19; Table 8-1). 18 of the 25 were expressed over 0.5 TPM in both vegetative and reproductive developmental stages. This set of 18 genes includes the putative magnesium transporter TraesCS3A02G414000 (Gebert et al. 2009). This gene contains the missense mutation underpinning the SH838 marker, which is predicted to be highly deleterious (SIFT = 0). This gene is the only gene within the 4.3 Mb fine-mapped region which contained a coding region SNP, positioning it as a strong candidate gene. As a magnesium transporter, it also correlates with the observed magnesium deficiency phenotype in the mutant plants (Figure 2-13) and the interveinal chlorosis which is also characteristic of magnesium deficiency (Figure 2-8). However, we grew two additional lines containing mutations in the TraesCS3A02G414000 gene and observed no chlorosis or early senescence phenotype in the field (Table 2-10; Figure 2-20).

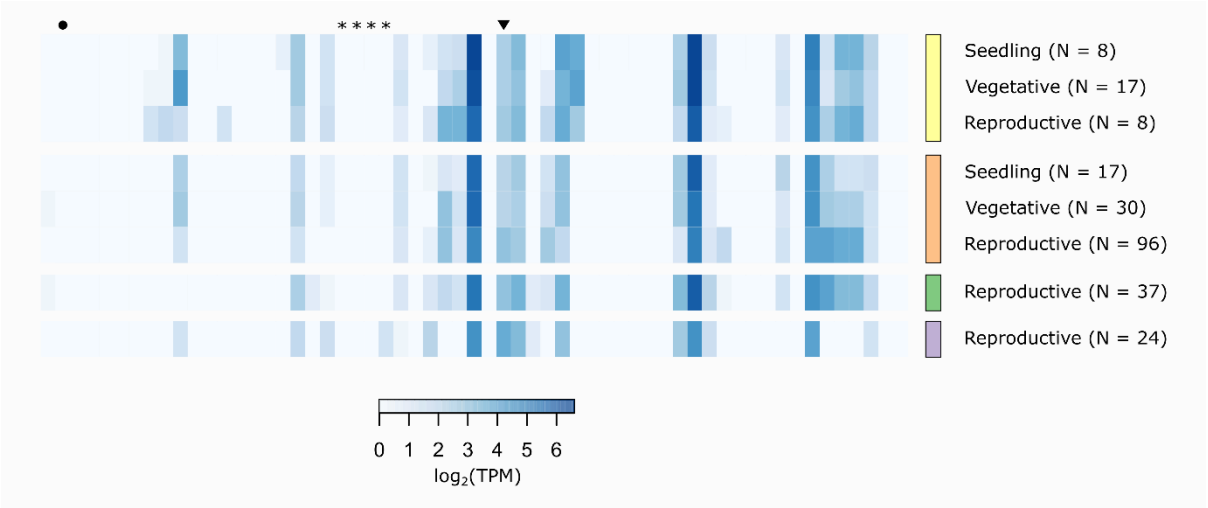

Figure 2-19: The 59 genes within the *YES-1* region show variable expression levels across tissues and development. Gene expression data derived from an RNA-Seq developmental timecourse (RRG 2018) displays expression levels in root (yellow, top), leaf/shoot (orange, second from top), spike (green, second from bottom) and grain (purple, bottom) tissues. Genes mentioned in the text are highlighted by an asterisk (TraesCS3A02G412900 to TraesCS3A02G413200; OsSAG12-1 orthologs), a circle (TraesCS3A02G410800; Tryptophan Decarboxylase 2) or an inverted triangle (TraesCS3A02G414000; putative Mg^{2+} transporter).

Of the remaining genes in the region, we found that the rice orthologs of five of the 59 genes have senescence-related functions. Four genes are present in a tandem duplication, TraesCS3A02G412900 to TraesCS3A02G413200, and are orthologous to the rice gene *OsSAG12-1*. *Os SAG12-1* is a negative regulator of senescence (Singh et al. 2013), so conceivably a mutation in this gene could lead to premature senescence. The fifth gene, TraesCS3A02G410800, is orthologous to *Tryptophan Decarboxylase 2*, a gene which increases serotonin levels and delays leaf senescence when over-expressed (Kang et al. 2009). All five of these genes are expressed at very low levels, if at all, throughout the developmental and tissue stages considered. The *OsSAG12-1* orthologs are expressed exclusively in ripening grains in the developmental timecourse, whereas TraesCS3A02G410800 is only expressed in anthers and ovary tissues at anthesis (Ramírez-González et al. 2018). The majority of the genes within the fine-mapped region lack any functional annotation or orthologs in rice or Arabidopsis, hindering the prediction of any function related to the early senescence and chlorosis phenotype observed in the mutant plants.

Table 2-10: TILLING Lines with mutations in a putative Mg^{2+} transporter, TraesCS3A02G414000, had no phenotype in the field.

Line	EMS Mutation	Amino Acid Mutation	Predicted Consequence	Presence of chlorosis?
K2282	G>A	G378R	Missense mutation, SIFT = 0.01	Yes
K2091	G>A	NA	Splice acceptor variant; frameshift and premature stop codon.	No
K2468	C>T	P187S	Missense mutation, SIFT = 0	No

Figure 2-20: Independent mutations in the putative Mg^{2+} transporter TraesCS3A02G414000 do not recapitulate the chlorosis phenotype from K2282. Individual flag leaves were photographed ten days after heading, during the JIC 2018 field trial. Only plants from lines homozygous for the mutation in TraesCS3A02G414000 were photographed.

2.4 Discussion

2.4.1 Induced SNP variation can lead to novel dominant phenotypes

As a result of the homoeolog redundancy present in the wheat genome, a dominant mutation in a single homoeolog can have a much greater impact on plant phenotype than a recessive mutation. We find that many of the critical domestication alleles in polyploid wheat are derived from dominant mutations, regulating traits including flowering time variation and free-threshing (Yan et al. 2004, Fu et al. 2005, Simons et al. 2006, Greenwood et al. 2017). Dominant alleles have retained their importance in modern breeding, forming the backbone of the Green Revolution through the semi-dominant *Rht-1* dwarfing allele (Peng et al. 1999, Borrill et al. 2015a). However, the pool of standing dominant variation within natural populations of wheat is small, particularly

for traits of agronomic interest. Most often, forward screens of natural populations identify multiple quantitative trait loci (QTL) that each contribute towards a small portion of the variation in that trait. When combined with the difficulty presented by homoeolog redundancy, it becomes particularly tricky to work with those complex traits where recessive mutations must be present in two or three homoeologs for a phenotype to be made visible (Figure 2-1) (Borrill et al. 2015a, Borrill et al. 2019b).

In this chapter, we have identified a novel source of dominant genetic variation underpinning chlorosis and premature senescence in wheat. This variation comes not from natural variation within wheat progenitors or landraces but rather from a chemically-mutagenised TILLING population. The fact that the only strong, novel mutant phenotype we identified from the TILLING screen is caused by a dominant mutation speaks to the low-likelihood of obtaining recessive mutations in both homoeologs in the TILLING lines. Our identification of the *NAM-A1* mutations also corroborates this finding as the *NAM-B1* homoeolog is non-functional in Kronos. As a result, a single mutation in the A genome homoeolog can lead to a complete knock-out of *NAM* gene function, and a strong delayed senescence phenotype (Figure 2-6) (Pearce et al. 2014, Harrington et al. 2019c).

We suggest, therefore, that the Kronos TILLING population is well-suited for further identification of novel dominant alleles governing phenotypes of agronomic importance. Already the population has been used to identify a deletion of the semi-dominant *Rht-B1* dwarfing allele (Mo et al. 2018). The hexaploid Cadenza TILLING population was also used to successfully identify mutant lines producing stable seminal root number phenotypes (Shorinola et al. 2019). Here we have mapped a dominant allele controlling chlorosis and early senescence to a region of chromosome 3A that contains no previously characterised genes. This emphasizes the potential of using populations with induced variation, such as the TILLING populations, both in a tetraploid and hexaploid background, to identify novel dominant alleles.

2.4.2 The advent of cultivar-specific assemblies will greatly assist the identification of non-genic SNPs

Having identified the dominant chlorosis phenotype segregating within the K2282 TILLING line, we then utilised exome-capture to localise the underpinning variation to chromosome 3A. Yet after reaching this point, we found that none of the TILLING SNPs present in this region associated completely with the mutant phenotype. This then suggested that the causal SNP may be located outside of the exome-captured region, in non-coding and potentially regulatory sequence. Dominant variation seen in wheat is often due to variation in regulatory sequences, such as for the flowering time gene *VRN1*. Here, variation in the promoter and intron sequence control the transition from winter to spring growth habit (Yan et al. 2004, Fu et al. 2005). CRISPR editing of

promoter sequences in tomato has demonstrated the wide range of morphological variation that can be obtained through mutations in regulatory sequences (Rodríguez-Leal et al. 2017).

Yet while we know that substantial phenotypic variation can be caused by variation in regulatory regions, most of the sequencing methods used in wheat fail to capture variation in non-coding regions. Due to the size of the wheat genome, reduced representation methods of sequencing are often used to limit the amount of the genome that is sequenced. Yet a pitfall of many of these methods, such as exome capture (Mamanova et al. 2010, Krasileva et al. 2017) and R-gene enrichment sequencing (Jupe et al. 2013, Steuernagel et al. 2016), is that they focus primarily or entirely on the coding regions of the genome. The *in-silico* TILLING database only includes SNPs within or near coding regions (Krasileva et al. 2017). While new capture methods have been designed to include more promoter and intron sequence (Gardiner et al. 2019), alongside the coding sequence, these methods still fail to capture variation in the more distal intergenic regions and the full intronic variation.

An alternative approach is to reduce the size of the genome by sequencing only one of the chromosomes. This not only provides coverage of non-coding regions of the genome, but it also reduces the influence of pre-existing bias, as the enrichment doesn't rely on pre-designed bait libraries for sequence enrichment. Here we successfully used flow cytometry to isolate and sequence chromosome 3A from the K2282 TILLING line (Doležel et al. 2012). This provided us with a wealth of information with which to identify new SNPs in non-coding regions, and then to design new markers on these SNPs to allow fine-mapping of the *YES-1* locus.

However, at this stage we were faced with the problem that the sequenced chromosome was from the tetraploid cultivar Kronos, while the gold-standard reference genome is based on sequence from the hexaploid cultivar Chinese Spring (IWGSC et al. 2018). How could we accurately call SNPs that were specific to the TILLING line when there is substantial variation between cultivars, particularly in non-coding regions? Here we were able to take advantage of a cultivar-specific assembly for Kronos, produced as part of the wheat pangenome initiative (10+ Wheat Genomes Project 2016). By calling SNPs specifically against the Kronos assembly, we were able to identify a set of high-quality SNPs that enabled us to fine-map the *YES-1* locus.

We were in part limited by the lack of physical positioning assigned to the Kronos assembly scaffolds (10+ Wheat Genomes Project 2016), however this can now be overcome by anchoring the Kronos scaffolds against the physical positions of the Chinese Spring RefSeqv1.0 assembly (IWGSC et al. 2018). These positions will not be perfectly correct—it's quite possible that deletions, inversions, or other structural arrangements will have occurred between the two cultivars. However, this provides a first approximation of physical positions for the Kronos scaffolds which can facilitate the identification of SNPs across a specific physical interval. The position of these SNPs could then be validated through genetic mapping in the F_2 populations.

Following the completion of this analysis the physically-contiguous genome of the tetraploid cultivar Svevo was released indicating the trend towards more contiguous assemblies for more cultivars (Maccaferri et al. 2019). Until all genome assemblies are fully contiguous, however, methods such as this for anchoring scaffolds to likely physical positions will be essential when working between the reference genomes and less contiguous varietal genomes.

2.4.3 The dominant chlorosis phenotype caused by YES-1 is environmentally dependent

Our initial identification of the chlorosis phenotype present in TILLING line K2282 was dependent upon the phenotype appearing in both the 2015 and 2016 field seasons at the John Innes Centre. However, despite the consistency of the phenotype across all four years of field seasons in the United Kingdom, we saw no evidence of the mutant phenotype when the TILLING lines were grown in Davis, California. This discrepancy suggested that an environmental variable was likely influencing the presentation of the mutant phenotype.

We found that the plants received substantially more rainfall in the Davis field trial between sowing and heading than in any of the JIC field trials (Table 2-7). That winter, 2016-2017, was unusually wet in California, with the state receiving an average rainfall of 781 mm between October 2016 and March 2017, its second wettest winter on record (NOAA National Centers for Environmental Information 2017). This suggested that the chlorosis phenotype may appear at the JIC in response to higher levels of water stress, however we could not recapitulate the mutant phenotype under low-water conditions in the glasshouse. It was also possible that the chlorosis phenotype may be influenced by differences in temperature or photoperiod conditions in the UK compared to California. Similarly, the fact that the lines were grown as winter wheat in California, but as spring wheat at the JIC, may also have influenced the appearance of the chlorosis phenotype. However, as we were unable to recapitulate the chlorosis phenotype when grown in the UK under glasshouse conditions, in which the temperature, photoperiod, and growth duration were similar to those present in the JIC field, it appears unlikely that these variables are causal.

As chlorosis can be a symptom of nutrient deficiency, we then investigated whether variation in the mutant phenotype may be due to differences in the soil mineral profiles. This would correlate with the observed reduction in leaf mineral content and interveinal chlorosis in the mutant plants (Figure 2-8, Figure 2-13). It would also correspond with the high levels of precipitation observed in the Davis field trial, as rainfall can lead to the leaching and loss of nutrients from the soil. Within the *YES-1* region, we found that the single missense mutation within the region was in a putative Magnesium transporter. This gene was then considered to be a promising candidate; however mutant plants grown in the glasshouse in soil taken from the JIC field site still failed to reproduce the phenotype, suggesting that soil mineral content alone was not sufficient to induce the mutant phenotype. Compounding this, we were unable to recapitulate the chlorosis phenotype with three

independent TILLING mutations in the gene, including frameshift and missense mutations (Table 2-10).

It is possible that the specific missense mutation in TraesCS3A02G414000, the putative Mg^{2+} transporter, is uniquely able to induce a dominant chlorosis phenotype. Perhaps this mutation alters protein function in such a way to produce a dominant negative effect, rather than a simple loss-of-function. This transporter is predicted to function as a pentamer (Eshaghi et al. 2006, Lunin et al. 2006, Payandeh and Pai 2006), based on orthology to related transporters (Gebert et al. 2009), suggesting that perhaps this missense mutation allows the formation of the pentamer with other wild-type alleles of the transporter, but that the presence of even a single copy of the mutant protein is sufficient to inhibit the transporter activity. Such a dominant-negative effect has been observed with the yeast Alr1p and Alr2p Mg^{2+} transporters, where single-base substitutions in the transmembrane domains were sufficient to inhibit growth on low-Mg^{2+} media suggesting that the mutant monomers were able to inhibit wild-type protein function (Wachek et al. 2006). The mutations identified in this system were located within the transmembrane domain of the Alr1p protein, similar to the location of the TILLING missense mutation at the edge of the transmembrane domain of the putative Mg^{2+} transporter. In this way heterozygous plants could present an equally strong mutant phenotype as homozygous mutant plants, as observed in our study. If this were the case, specific base-editing, perhaps driven by a Cas9 system, would be required to validate the effect of this mutation in an independent background (Zong et al. 2017).

Given our inability to recapitulate the phenotype in field soil in the glasshouse, however, it's also possible that a completely unrelated SNP located in a non-coding region of the genome is causal for the dominant mutant phenotype. This SNP may be regulating the Mg^{2+} transporter itself, or another gene entirely. In our analysis here, we highlighted the 25 genes within the fine-mapped *YES-1* region which were expressed during leaf development (Figure 2-19). However, if the chlorosis phenotype is driven by a dominant regulatory SNP, it is possible that the effect of the mutation is to induce ectopic expression of a gene that is not normally expressed in the leaf and shoot tissue. Regulatory mutations inducing this kind of ectopic expression in wheat have previously been identified, including many of the *VERNALISATION* genes (Yan et al. 2003, Yan et al. 2006) and the *Ppd-D1* alleles (Beales et al. 2007).

There are few examples of dominant chlorosis phenotypes in the literature; one, in *Brassica napus*, only appears before budding (the equivalent developmental stage as heading in cereals) (Wang et al. 2016). In contrast, our observed phenotype appears to become stronger following heading. Another example is the *Ygm* mutant in wheat, named as such for the yellow-green leaves of the plants (Wu et al. 2018). This mutant shows a semi-dominant phenotype, with the heterozygotes presenting as intermediate between the yellow mutants and the green wild-types. This phenotype is underpinned by abnormal chloroplast development, and the mutants show differential expression of genes involved in chlorophyll biosynthesis and carbon fixation, amongst other traits. In order to

ascertain whether a similar pathway may be acting in the K2282 mutant, further work characterising the cellular and molecular basis of the phenotype will be required. In tandem with further fine-mapping of the *YES-1* locus, this work has the potential to shed light not only on the regulation of chlorosis in wheat, but also on mechanisms governing dominant traits in wheat as a whole.

2.5 Summary

In this chapter, we have shown that the new genomic and germplasm resources available in wheat, and particularly in tetraploid wheat, can facilitate the identification and mapping of novel mutant phenotypes. We screened the Kronos TILLING population and identified a dominant chlorosis and early senescence mutant. Using bulked segregant analysis and exome capture, we were able to map the causal locus, *YES-1*, to chromosome 3A. We then used flow cytometry to isolate the mutant chromosome 3A, aligning the sequence to both the Chinese Spring and the Kronos assemblies. With these new, non-coding region SNPs, we were able to fine-map the *YES-1* locus to a 4.9 Mb region, encompassing 59 high-confidence genes. We hypothesise that a missense mutation in the putative Mg^{2+} transporter TraesCS3A02G414000 may be causal for this phenotype.

3 Conserved residues in the wheat (Triticum aestivum) NAM-A1 NAC domain are required for protein binding and when mutated lead to delayed peduncle and flag leaf senescence.

This work was previously published in BMC Plant Biology (Harrington et al. 2019c). The cloning of the *NAM-A1* missense alleles and preliminary yeast two-hybrid screens were carried out by Lauren Overend during her placement on the JIC International Summer School in 2017, under the supervision of the author. The initial phenotyping of the 2015 trial was carried out by Dr. Philippa Borrill at the JIC. Dr. Nicolas Cobo facilitated the field trials in UC Davis, sowing the plants and assisting with phenotyping and sampling alongside the author.

3.1 Introduction

3.1.1 The NAC transcription factor family

3.1.1.1 NAC transcription factors function in plant development and stress responses

Specific to the Plant Kingdom, the NAC transcription factor family has been shown to have roles regulating both fundamental developmental processes and many different stress responses. The first NAC transcription factor, *NAM*, was described in Souer et al. (1996) and was found to be essential for pattern formation in early development in Petunia. During this analysis, they also identified *ATAF1/2*, genes that were later found to be involved in ABA signalling in response to a biotrophic pathogen (Jensen et al. 2008). Shortly after, the *CUC1/2* genes were found to have roles in organ separation and were orthologous to the *NAM* gene in *Petunia* (Aida et al. 1997). Together, these three genes gave their name to a new family of plant-specific transcription factors, the NACs (*NAM*, *ATAF1/2*, and *CUC*).

Since these initial discoveries, many NAC transcription factors have been shown to regulate responses to environmental stresses, including ATAF1/2 as discussed above. They may be induced in response to wounding, such as with *ATAF1/2* and the potato (*Solanum tuberosum*) NAC *StNAC* (Collinge and Boller 2001). Other NAC transcription factors act in response to drought, inducing the expression of stress-responsive genes and increasing the drought tolerance when over-expressed in Arabidopsis (Tran et al. 2004). In wheat, a variety of NACs have been shown to regulate responses to biotic stresses such as stripe rust (*TaNAC30* (Wang et al. 2018a), *TaNAC21/22* (Feng et al. 2014)) and to abiotic stresses such as salt stress and drought tolerance (*TaNAC29* (Huang et al. 2015b, Xu et al. 2015), *TaNAC2* (Mao et al. 2012), *TaNAC69* (Xue et al. 2006), *TaNAC2a* (Tang et al. 2012), *TaNAC67* (Mao et al. 2014)). Other NAC transcription factors have been shown to regulate both abiotic and biotic stress responses in wheat (*TaNAC8* (Xia et al. 2010b), *TaNAC4* (Xia et al. 2010a)).

NACs also act across the spectrum of development, from the earliest developmental stages through to the final stages of a plant's life. In Arabidopsis, the *NAC1* gene regulates the development of lateral roots by acting within the auxin signalling pathway (Xie et al. 2000). More recently, a wheat NAC transcription factor, *TaRNAC1*, was found to increase root length and biomass when over-expressed in wheat roots, as well as improve drought tolerance (Chen et al. 2018). NAC transcription factors have also been found to regulate senescence across the plant kingdom. In wheat specifically, the NAM NAC transcription factors are positive regulators of both senescence and grain protein content (Uauy et al. 2006b). The *TaNACS* transcription factor was found to be a negative regulator of senescence in wheat, delaying senescence upon over-expression (Zhao et al. 2015). Other NAC transcription factors have been found to regulate not only developmental processes, but also a variety of stress responses. *TaNAC1*, a wheat NAC transcription factor, has been found to regulate both stripe rust and *Pseudomonas syringae* resistance and also promote lateral root development when expressed heterologously in *Arabidopsis thaliana* (Wang et al. 2015). The wide range of developmental and stress responses regulated by NAC transcription factors suggests that they play a key role in plant growth and survival in many contexts.

3.1.1.2 NAC transcription factors contain a highly conserved NAC domain

NAC transcription factors are united by a shared set of structural motifs that are unique to this family. They are characterised by the presence of a conserved N-terminal domain, the NAC domain, which itself consists of five highly conserved subdomains, hereafter numbered i through v (Figure 3-1A). Alongside a NAC domain, NAC transcription factors also contain an intrinsically-disordered C-terminal region (Figure 3-1A). This region is involved in the transcription activation or repression activity of the transcription factors (Xie et al. 2000, Yamaguchi et al. 2010, Welner et al. 2012). The structure of the NAC domain was first obtained based on the NAC domain of the Arabidopsis ANAC019 protein (Ernst et al. 2004). This NAC domain was found to consist of a unique fold motif, specifically a twisted anti-parallel β-sheet surrounded by a long α helix on the N-terminal side of the sheet and a shorter α helix on the opposing side. Diffraction data from this initial experiment suggested that the NAC domain forms dimers through interactions between highly conserved residues within the NAC domain, particularly residues present near the N-terminal end of the protein. This correlated with other research which had shown that NAC transcription factors primarily formed dimers in solution (Olsen et al. 2005a).

Further work on the structure of the NAC transcription factors began to reveal the specific roles of key residues of the NAC domain. The crystal structure of the NAC domain of ANAC019 binding DNA was resolved in 2012, demonstrating that ANAC019 could bind DNA as a homodimer (Figure 3-1B)(Welner et al. 2012). This identified a set of regions and residues of the NAC domain that were involved in the DNA-protein interface, located in subdomains iii, iv, and v. In the same study, the impact of specific residues on DNA-protein binding was tested using an electrophoretic mobility shift assay (EMSA); of the residues tested, which were selected based on the predicted

likelihood of their role in DNA-binding, only one, arginine 88, was found to be required for DNA binding *in vitro* (Welner et al. 2012). This specific residue had previously been identified as one of a pair of residues (arginine 85 and arginine 88) that were required for DNA binding by ANAC019 *in vitro* (Olsen et al. 2005a). In this same paper, a second pair of residues in ANAC019, arginine 19 and glutamic acid 26, were found to be required for protein dimerisation.

More recently, further work on the protein dimerisation properties of NAC transcription factors has identified a set of specific residues in subdomain i of the NAC domain which are predicted to be involved in pH-dependent stabilisation of the NAC dimer (Kang et al. 2018). In particular, when a histidine residue located in subdomain 3 (H135 in ANAC019) is not protonated, a series of interactions are formed between residues in subdomain i of each monomer, including salt-bridge, sidechain-backbone, and backbone-backbone hydrogen bonds (Figure 3-1C). In this state, the unprotonated H135 residue is able to form a backbone-sidechain interaction with the proline residue at position 16 in the other NAC monomer, contributing to the stabilisation of the dimer. However, under more acidic conditions, the H135 residue is protonated and instead forms a salt bridge with the aspartic acid residue at position 24 in its own monomer (Figure 3-1C). This prevents the formation of salt bridge interactions between two different residues, arginine 19 and glutamic acid 26, which are likely to be essential in stabilising the dimer. These were the two residues which were previously found in (Olsen et al. 2005a) to be essential for protein dimer formation.

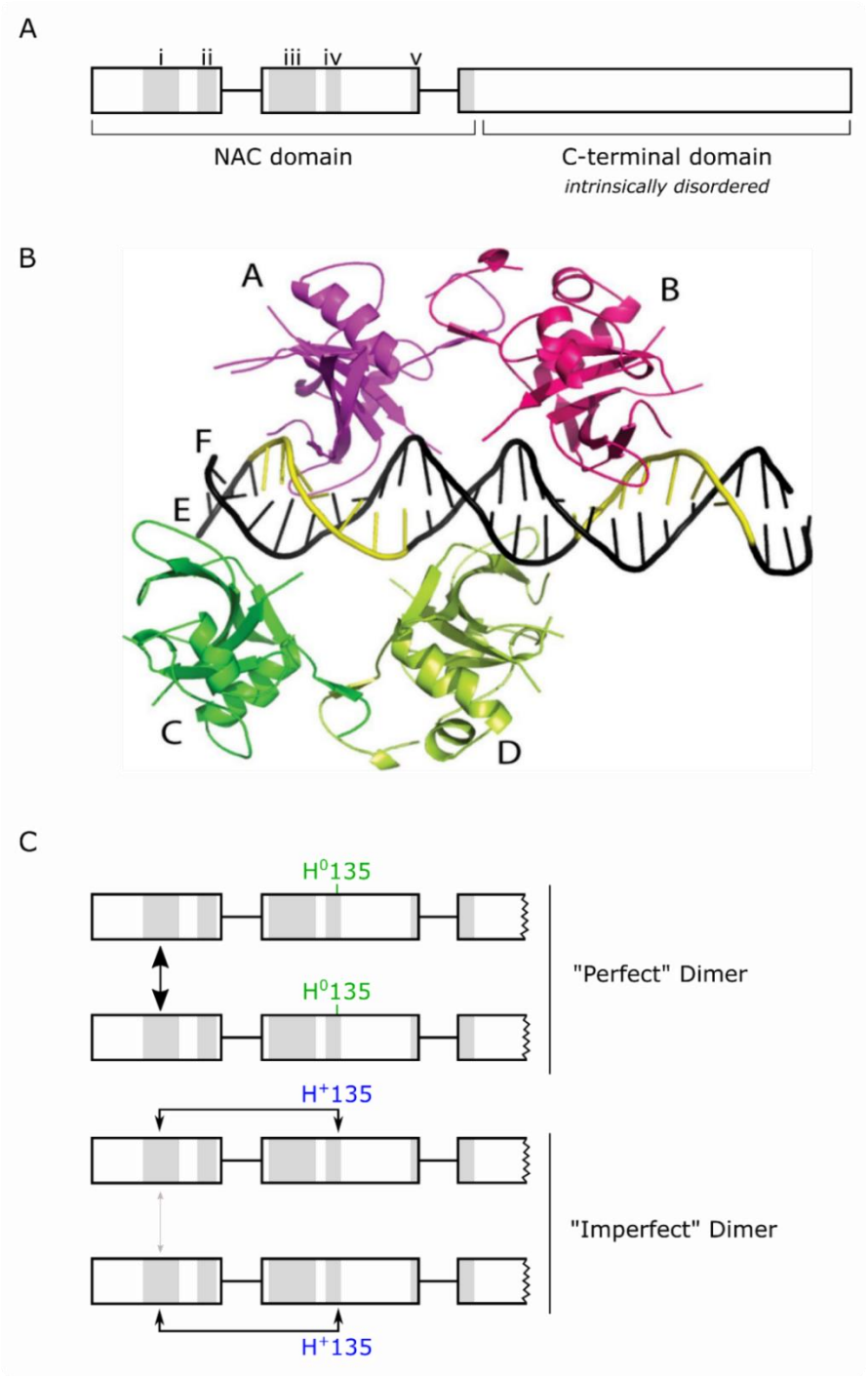

Figure 3-1: The structure of and interactions between NAC transcription factors. (A) A schematic of a canonical NAC transcription factor highlights the N-terminal NAC domain and the C-terminal intrinsically disordered domain. Subdomains i through v of the NAC domain are highlighted in light grey. (B) Crystal structures of ANAC019 demonstrate that NAC transcription factors bind DNA as a dimer, shown here as dimers AB and CD binding to the major clefts of the DNA strand (EF), with binding sites shown in yellow. (C) Two forms of the NAC dimer, the "perfect" and "imperfect" dimers, depend on the protonation status of a histidine residue in subdomain iv. Unprotonated H135 allows the formation of strong interactions between NAC monomers via residues in subdomain i. Protonation of H135 leads to the formation of a salt bridge interaction with a residue in subdomain i, changing protein conformation and reducing the strength of the interactions between NAC monomers. The C-terminal domain of the schematic is truncated for visualisation purposes, indicated by the jagged line. Section B is reproduced from Welner et al. (2012), with permission from the publisher, and section C is after Kang et al. (2018).

This structural work has been essential in increasing our understanding of the mechanisms by which NAC transcription factors interact with each other, and with the promoters of the genes they regulate. However, to our knowledge, no investigations have been carried out into the phenotypic impact of mutating these and other conserved residues of the NAC domain *in planta*. Over the past few years, the understanding of NAC transcription factors in wheat has increased dramatically. Not only are the roles of more individual NAC transcription factors being characterised, as previously introduced, but the global landscape of NAC transcription factors in wheat has now been investigated. We previously carried out a genome-wide analysis of the NAC transcription factor family in wheat, characterising their expression patterns and structural similarities (Borrill et al. 2017). We were able to resolve a phylogeny of wheat, rice, and barley NAC transcription factors that closely recapitulated the eight NAC subfamilies previously identified in a study of 11 different species, including Arabidopsis, rice, and poplar (Shen et al. 2009). Within these subfamilies, we found that gene expression patterns across a timecourse of development were more similar than compared to NACs in other subfamilies.

3.1.2 Induced variation in polyploid wheat generates novel missense alleles

This analysis, alongside the development of new genomic, transcriptomic, and genetic resources in wheat, suggested to us that we could use wheat as a system to investigate the role of specific residues in the NAC domain in protein function, and that this information could be applicable across multiple species. We could use the *in-silico* TILLING population to identify EMS-induced point mutations in highly conserved residues of the NAC domain (Krasileva et al. 2017). Having recently been re-mapped against the RefSeqv1.0 genome, almost 99% of all sequenced wheat genes contain at least one missense mutation and on average 30 distinct missense alleles within the TILLING lines. This high mutation density was possible due to the polyploid nature of wheat, whereby the multiple homoeologous copies of each gene, which are often functionally redundant, act as buffers against the severe and lethal mutant phenotypes that would otherwise arise (Wang et al. 2012).

However, while the polyploidy of the wheat genome is advantageous in allowing an increased mutation load, it correspondingly increases the difficulty in obtaining a mutant phenotype. Many of these homoeologs are functionally redundant, and consequently require mutations in all three copies to induce a complete mutant phenotype (Borrill et al. 2019b). In a hexaploid background, this would require two rounds of crossing to obtain a triple mutant in the gene of interest. To address this limitation, we can turn to the Kronos TILLING population. This population was developed in a tetraploid wheat cultivar, containing only two genomes and thus typically two homoeologs of each gene (Uauy et al. 2009, Krasileva et al. 2017). As a result, only two homoeologs need to be mutated in order to obtain a mutant phenotype. Yet while this is indeed useful when solely testing gene knock-outs, in order to evaluate the impact of a specific point

mutation on gene function it would be most useful if only one gene needed to be mutated to produce an interpretable mutant phenotype.

3.1.3 NAM-A1 is a positive senescence regulator that is present in a single functional copy in tetraploid wheat

One of the NAC transcription factors previously mentioned, the *NAM-1* gene, was originally identified through its B-genome homoeolog, *NAM-B1*, which underpinned a high grain protein locus that had been introgressed into modern bread wheat from the *Triticum turgidum* ssp. *dicoccoides* accession FA15-3 (Olmos et al. 2003, Uauy et al. 2006a). Notably, the B-homoeolog was found to be non-functional in most modern cultivars, either due to a frame-shift mutation or a complete deletion of the homoeolog (Uauy et al. 2006b). The A and D genome homoeologs of *NAM-B1*, *NAM-A1* and *NAM-D1*, are functional in modern cultivars. Additionally, a separate paralogous set of homoeologs on chromosome 2 were identified during characterisation of *NAM-B1*, referred to as the *NAM-2* genes. Initially only the B and D genome copies of *NAM-2* were identified; more recently, work has also shown that *NAM-A2* is present and functional in wheat (Borrill et al. 2019a).

Further work in both hexaploid and tetraploid wheat demonstrated that a premature truncation mutant in the A-genome homoeolog, *NAM-A1*, was sufficient to induce a significant delay in senescence. In hexaploid wheat, mutations in *NAM-A1* were sufficient to induce a delay in flag leaf senescence of approximately five days, while retaining more chlorophyll in the flag leaf than the wild-type control from 30 days after anthesis until senescence (Avni et al. 2014). This delay in senescence was similar, if not slightly larger, than that seen for the mutation in *NAM-D1*, though both were less severe than the double mutant. The *NAM-A1* mutation also led to decreases in grain protein content of between 1 and 7% compared to the control, though this was both less severe than the double mutant and, in some cases, the single D-genome mutant. In a tetraploid background, where only the A-genome homoeolog is present and functional, a premature stop codon mutation led to a substantial delay in flag leaf senescence onset of approximately 27 days (Pearce et al. 2014). This mutant also had significantly less grain protein, iron, and zinc content. In this case, the effect of *NAM-B2* on senescence was also tested and found to be significantly weaker than that of *NAM-A1*. In both the tetraploid and hexaploid experiments these results were consistent across different environments in California and Israel indicating that the effect of *NAM-A1* on senescence and grain quality is highly stable. This presented *NAM-A1* as an ideal candidate to interrogate the impact of specific point mutations in the NAC domain on protein function *in planta*.

3.1.4 Aims and hypotheses

Given the presence of *NAM*-A1 as a single-copy gene in tetraploid wheat, we hypothesized that we could use this system to study the effect of missense mutations in the NAC domain of *NAM-A1* on protein function and plant phenotype. In this chapter, we identified a set of highly conserved

mutations in the NAC domain of *NAM-A1* that were present in the Kronos TILLING population. We then investigated whether TILLING lines containing these mutations were able to recapitulate the expected delayed senescence phenotype observed previously. We also tested the impact of these mutations on protein-protein interactions in yeast and investigated the effect of these mutations on heterologous induction of cell death in *N. benthamiana*. Through this work, we identified four residues in the NAC domain which, when mutated, lead to both delayed senescence and a loss of protein binding. Based on their high levels of conservation, these residues are likely to be essential for the function of NAC transcription factors across the Plant Kingdom.

3.2 Methods

3.2.1 Plant growth and phenotyping

3.2.1.1 Glasshouse experiments

Kronos TILLING mutant lines (Uauy et al. 2009, Krasileva et al. 2017) which contained mutations in the NAM-A1 gene were selected initially using the www.wheat-tilling.com website. We identified lines which contained missense mutations predicted or known to affect protein function, alongside control lines containing synonymous, non-conserved missense, or splice-variant mutations in *NAM-A1* (Table 3-14) (Ernst et al. 2004, Olsen et al. 2005a). M_5 seeds of the selected TILLING lines were pre-germinated on moist filter paper for 48 h at 4°C. The sample size for each experiment varied, and is detailed in the relevant figure caption. Seeds were then sown into P96 trays in 85% fine peat with 15% horticultural grit. At the two to three leaf stage, individual plants were transplanted to 1 L round pots containing Petersfield Cereal Mix (Petersfield, Leicester, UK). The plants were grown in standard glasshouse conditions, with 16:8 hours of light:dark cycles. At the seedling stage, they were genotyped to identify homozygous mutant and wildtype lines (see genotyping methods below). The primary spikes were tagged at heading, when 75% of the spike had emerged from flag leaf sheath (Zadok's growth stage 57) (Zadoks et al. 1974). The onset of flag leaf senescence was scored when 25% of the main flag leaf had lost chlorophyll and appeared visually yellow. Peduncle senescence was scored when the top inch of the main peduncle became fully yellow. The glasshouse phenotyping experiments were repeated twice under the same conditions (referred to as Experiment 1 and 2 in the text).

3.2.1.2 *Nicotiana benthamiana* growth

N. benthamiana seedlings were sown into half seed trays in F2 soil (Peat with 2.5 kg/m3 Dolimite, 1.3 kg/m3 Base fertiliser, 2.7 kg/m3 Osmcote 3-4 months, 0.25 kg/m3 Wetting agent, and 0.3 kg/m3 Exemptor). The seedlings were transplanted 11 days after sowing into FP8 pots. Plants were grown under a 16:8 hours light:dark regime under sodium lamps (400W), with temperature set to 25°C day and 22°C night.

3.2.1.3 JIC field trials

Two lines (K2711 and K1107) were selected for further evaluation in the field. The M_5 mutant lines were previously scored for senescence phenotypes in 2015, as detailed in Chapter 2, section 2.3.1. Following this, for each line homozygous *NAM-A1* mutant plants were crossed to the wild-type original parent, Kronos. F_1 plants were self-pollinated to produce F_2 populations segregating for the mutant or wild-type *NAM-A1* allele. We recovered two independent populations for K1107 (referred to as K1107-4 and K1107-6) and one for K2711. The F_2 populations were sown at Church Farm, Bawburgh (52°38'N 1°10'E), in March 2016 as detailed in Chapter 2, section 2.2.1.

To score segregating F_2 plants for phenotype and genotype, individual seeds were hand-sown in 6x6 1 m^2 grids of 36 individual plants with approximately 17 cm between each plant as depicted in Figure 2-2B, Chapter 2 (p.41) (K2711, n = 158 F2 plants; K1107-4, n = 218; K1107-6, n = 212). The individual plants were scored for heading date (two tillers with spikes 75% emerged), leaf senescence (two tillers with flag leaves 25% senesced from tip), and peduncle senescence (two peduncles, top inch senesced). Senescence scoring was consistent with that used in the glasshouse experiments. Following identification of homozygous mutant and wild-type F_2 individuals (see genotyping methods below), we selected between 13-16 individual plants for each line and genotype to take forward for further phenotyping in 2017 (Davis). A subset of these selected homozygous lines was also used for phenotyping in 2018 (JIC). Only individuals genotyped as homozygous wild-type or mutant were included in data analysis for 2016.

In April 2018, homozygous lines derived from independent F_2 plants in the 2016 field trial were sown at Church Farm for K2711 (WT *NAM-A1*, n = 5; Mutant *NAM-A1*, n = 5) and K1107 (WT *NAM-A1*, n = 5; Mutant *NAM-A1*, n = 5). Each of the lines was treated as an independent biological replicate of the mutant or wild-type genotype, thus accounting for the effect of segregating background mutations derived from the original TILLING mutants. Each biological replicate was sown in double 1 m rows with each row containing 20-30 plants, separated by a single empty row, as depicted in Figure 2-2C, Chapter 2 (p.41), arranged in a complete randomized design. Double rows were scored as "senesced" when 75% of the main tillers exhibited the leaf or peduncle senescence phenotype. A total of 124.5 kg/ha of Nitrogen was applied to the 2018 field trials.

3.2.1.4 Davis field trial

Homozygous lines derived from individual F_2 plants were sown for K2711 (WT *NAM-A1* n = 13, Mutant *NAM-A1* n = 15) and K1107 (WT *NAM-A1* n = 14, Mutant *NAM-A1* n = 16) in November 2016 at the University of California Field Station, Davis, California (38° 31' N, 121° 46' W). As in 2018, each of the lines was treated as an independent biological replicate of the relevant genotype. K2711 and K1107 were sown as two independent complete randomized design trials. Each biological replicate was sown as double 1 m rows, with each row containing 20-30 plants, separated by one empty row, identical to the JIC 2018 trials, again as depicted in Figure 2-2C, Chapter 2 (p.41). Senescence was scored in the same manner as for the 2018 JIC trials. A total of

147 kg/ha were applied as ammonium sulphate, with half at pre-planting and the rest at the beginning of jointing (Zadoks 30, March 31 2017). Both the Davis and JIC trials were sprayed with appropriate commercial fungicides to avoid disease incidence.

3.2.1.5 Chlorophyll content

We used the SPAD-502 meter (Konica Minolta) to non-destructively measure relative chlorophyll content in the flag leaves. For glasshouse trials, readings were taken eight times along the flag leaf blade, four on each side of the mid-vein, and averaged to obtain a single reading for each leaf (n = 5 for each line and genotype). For the 2017 and 2018 field measurements, 2 individual leaves per biological replicate (n = 5 for each line and genotype) were measured as above and averaged for each time point.

We also measured the total chlorophyll content of sampled peduncle tissue in the JIC 2018 field trial. The top 3 cm of peduncles was sampled at 33 and 49 days after anthesis and was cut into small segments (approximately 0.5 cm in length). Two peduncles were sampled per biological replicate and pooled before chlorophyll extraction. Five biological replicates were sampled for each homozygous genotype (K2711 WT, K2711 mutant, K1107 WT, and K1107 mutant). Plant tissue was soaked in 3 mL N,N-Dimethylformamide (analytical grade, Sigma Aldrich, UK) for 4-5 days at 4 ˚C until all pigment had leached from the tissue. Chlorophyll and carotenoid content were quantified as previously described in Chapter 2, section 2.2.2.3.

3.2.1.6 Grain size

Grain samples from individual plants (Glasshouse and Field 2016) and pooled samples (Field 2017 and 2018) were analysed using the MARVIN seed analyser (GTA Sensorik GmbH). Grain morphometric parameters, including area, length, and width, alongside approximate thousand grain weight (TGW) were obtained.

3.2.1.7 Grain protein content

Grain protein content measurements were carried out using the Perten DA7250 NIR analyser. Samples of between 20 – 700 grains were tested for protein content (Protein dry basis %). Samples with fewer than 300 grains were analysed using a static cup; samples with more than 300 grains were analysed using a rotating cup. No significant variation in moisture content was observed between the genotypes and was thus not accounted for in further analysis. Nor did the choice of rotating cup or static cup lead to any significant variation in recorded protein content.

3.2.2 Genotyping

KASP genotyping was used at all stages of the project to confirm the presence or absence of the expected *NAM-A1* mutations in mutant and wild-type TILLING lines, respectively. The KASP reactions and DNA extractions were carried out as previously described in Chapter 2, section 2.2.6. The list of KASP primers used for each of the TILLING lines used in this study is presented in Table 3-1. Where possible, KASP primers were designed using the PolyMarker tool (Ramirez-

Gonzalez et al. 2015b). Where appropriate KASP markers could not be automatically designed, we manually identified primer sets that ensured homoeolog and paralog specificity. All wild-type (WT) primers contain the HEX tail ("GAAGGTCGGAGTCAACGGATT"), while all mutant (Mut) primers contain the FAM tail ("GAAGGTGACCAAGTTCATGCT").

Table 3-1: KASP Primers for NAM-A1 TILLING mutants.

Primer	Sequence (5′ - 3′)
K3661_Mut	GAAGGTGACCAAGTTCATGCTgtggaaccggaagcccga
K3661_WT	GAAGGTCGGAGTCAACGGATTgtggaaccggaagcccgg
K3661_Com	ggctcggcgcaaaaagca
K2734_Mut	GAAGGTGACCAAGTTCATGCTcgaccagctcctcgtcca
K2734_WT	GAAGGTCGGAGTCAACGGATTcgaccagctcctcgtccg
K2734_Com	aagcagcgcggcatcagcatg
K3741_Mut	GAAGGTGACCAAGTTCATGCTctactggaaggccaccga
K3741_WT	GAAGGTCGGAGTCAACGGATTctactggaaggccaccgg
K3741_Com	cgccttcttgacgccgag
K2060_Mut	GAAGGTGACCAAGTTCATGCTtttctaccgcgggaagccgct
K2060_WT	GAAGGTCGGAGTCAACGGATTtttctaccgcgggaagccgcc
K2060_Com	tggtggtggtggagccagaca
K3186_Mut	GAAGGTGACCAAGTTCATGCTtcttctaccgcgggaagcca
K3186_WT	GAAGGTCGGAGTCAACGGATTtcttctaccgcgggaagccg
K3186_Com	tggtggtggtggagccagaca
K2615_Mut	GAAGGTGACCAAGTTCATGCTggagctcccaccgggcttcca
K2615_WT	GAAGGTCGGAGTCAACGGATTggagctcccaccgggcttccg
K2615_Com	aagcagcgcggcatcagcatg
K2551_Mut	GAAGGTGACCAAGTTCATGCTtggcctcggggacggggta
K2551_WT	GAAGGTCGGAGTCAACGGATTtggcctcggggacggggtg
K2551_Com	aagcagcgcggcatcagcatg
K2711_Mut	GAAGGTGACCAAGTTCATGCTagcccgacgtcgccgccca
K2711_WT	GAAGGTCGGAGTCAACGGATTagcccgacgtcgccgcccg
K2711_Com	tccatccatcagagaaggcg
K1107_Mut	GAAGGTGACCAAGTTCATGCTagcacccagtcgtccaat
K1107_WT	GAAGGTCGGAGTCAACGGATTagcacccagtcgtccaac
K1107_Com	tcgcacggtatagctagcag

3.2.3 NAC domain alignment

All proteins in *Arabidopsis thaliana, Zea mays, Populus trichocarpa*, and *Physcomitrella patens* which contained the NAC domain, defined using the Interpro ID IPR003441, were extracted from EnsemblPlants using BioMart (Kersey et al. 2017). NAC proteins present in *Hordeum vulgare* and *Oryza sativa* were obtained from the curated list in Borrill et al. (2017). The NAC-domain containing proteins from *Triticum aestivum* were obtained from the gene annotations for RefSeqv1.1 (IWGSC et al. 2018).

The sequences of the NAC domain, ending after the final "CRVFKK" sequence which marks the end of subdomain v, were extracted from all of the collated proteins. These sequences were then aligned using Clustal Omega v1.2.0, and manually curated in Jalview v2.10.5 (Sievers et al. 2011). Following alignment, residues with less than 10% occupancy across the complete set of NAC transcription factors were excluded from further analysis as in Borrill et al. (2017). Based on this curated alignment, the Jensen-Shannon Divergence score at each residue was calculated as in Capra and Singh (2007), using the "Protein Residue Conservation Prediction" web server.

3.2.4 NAM-A1 *and* NAM-B1 *allele cloning*

TILLING mutations in the *NAM-A1* gene (TraesCS6A02G108300) were identified using the TILLING database at www.wheat-tilling.com, as previously detailed (Krasileva et al. 2017). We selected alleles for further study based on conservation of the residue in the NAC domain (PSSM viewer, pfam02365) and predicted SIFT scores (Ng and Henikoff 2003). For simplicity, the Kronos mutant lines are labelled as 'K' followed by their four-digit identifier (e.g. Kronos 2711 is referred to as K2711).

3.2.4.1 Gene synthesis and TOPO cloning

The *NAM-A1* coding sequence from *T. aestivum* (Transcript TraesCS6A02G108300.2) and the *NAM-B1* coding sequence from *T. turgidum* subsp. *dicoccoides* (GenBank accession DQ869673.1) were synthesized into the pUC57 vector (Genewiz). Gene sequences were amplified using primers available in Table 3-2 using Q5 polymerase, a proof-reading PCR enzyme, to reduce errors in gene amplification. Alongside the full-length coding sequences, truncated versions of *NAM-A1* and *NAM-B1* were also amplified, referred to as *NAM-A1*(1-217) and *NAM-B1*(1-215). These truncated versions correspond to the NAC domain of the proteins, without the C-terminal intrinsically disordered domain. The specific sequences for the full-length and truncated fragments of *NAM-A1* and *NAM-B1* are available in Appendix 8.1.1, and the TOPO vectors containing the full-length and NAC domain sequences can be ordered from https://www.addgene.org/Cristobal_Uauy/. All synthesized genes were diluted to approximately 1 ng/µL before use in PCR. The reaction cycling conditions are given in

Table 3-3, and the reaction components are given in Table 3-4.

Table 3-2: Primers used to amplify the full length and NAC domain fragments of wild-type NAM-A1 and NAM-B1.

Construct	Forward	Reverse	Size of fragment (bp)
NAM-A1	ATGAGGTCCATGGGCAG	TCAGGGATTCCAGTTCACG	1227
NAM-B1	ATGGGCAGCTCCGACTCA	CGTGAACTGGAATCCCTGA	1218
NAM-A1(1-217)	ATGAGGTCCATGGGCAG	TCACTGCTGATCTCCGG	651
NAM-B1(1-215)	ATGGGCAGCTCCGACTCA	CCGGCGATCAGCAGTGA	645

Table 3-3: Thermocycling conditions for amplification of gene (fragments) using Q5 polymerase.

Step	**Temperature**	**Time (s)**
Initial Denaturation	98°C	30
Amplification (x30)	98°C	10
	54°C	30
	72°C	60
Final Extension	72°C	120
Hold	10°C	∞

Table 3-4: Reaction setup for amplification of gene (fragments) using Q5 polymerase.

Reagent	**1x (µL)**
Q5 Buffer (NEB)	5
10mM dNTPs	0.5
GC enhancer (NEB)	5
Forward Primer (10µM)	1.5
Reverse Primer (10µM)	1.5
DNA	1
H_2O	10.5

Following amplification, PCR reactions were run on a 1% agarose gel, cut-out and DNA was extracted (QIAquick Gel Extraction Kit, QIAGEN). Deoxyadenosine triphosphate (dATP) overhangs ("A-tail") were added onto the purified DNA fragments as described in Table 3-5.

Table 3-5: Reaction set up for A-tail addition to gel-purified DNA fragments.

Reagent	**1x (µL)**
10X PCR Buffer (NEB)	1
dATP (2 mM)	1
Taq Polymerase	0.5
Gel-purified PCR product	7.5

The reaction (10 µL final volume) was incubated at 72°C for 20 minutes. Following A-tail addition, the fragments were ligated into the pCR™8/GW/TOPO® vector detailed in Table 3-6.

Table 3-6: Reaction set up for ligation of DNA fragments into pCR8.

Reagent	**1x (µL)**
Salt solution	1
TOPO vector	1
A-tailed PCR product	4

The total volume of 6 µL was mixed and incubated at 25°C for 1 hour and 30 minutes. Following ligation, the circularised vectors were transformed into library-efficient competent DH5α (Invitrogen, cat. no. 18263012) as follows:

1. Thaw 50 µl aliquots of competent cells on ice
2. Warm plates and SOC media at 37°C
3. Add 2 µl of ligation to cells and mix by gently tapping the tube
4. Incubate on ice for 30 mins
5. Heat shock cells at 42°C for 30 secs
6. Return to ice for 2 mins
7. Add 250 µl pre-warmed SOC media
8. Incubate 1 hour at 37°C, 225 rpm shaking
9. Spread 10 µl and 100 µl of culture on pre-warmed LB plates containing Spectinomycin (100 µg/mL)
10. Incubate overnight at 37°C

As the ligation of the A-tailed DNA fragments into the pCR8 vector is not oriented, the cloned pCR8 vectors with gene inserts were screened to identify colonies with the correct orientation of the gene insert. All colonies were screened using two primer pairs: M13F and a gene-specific reverse primer (Rev1), and M13R and Rev1. Only colonies that have a PCR band in the M13F/Rev1 combination have the gene insert in the correct orientation. The protocol was carried out as detailed in Table 3-7 and Table 3-8.

Table 3-7: Reaction conditions for colony PCR on pCR8 transformants.

Reagent	**1x (µL)**
5X GoTaq® Buffer with 7.5 mM $MgCl_2$ (Promega M791B)	1.5
2 mM dNTPs	1.5
Taq Polymerase (NEB)	0.07
M13F or M13R (2 µM)	1.5
Rev1 (10 µM)	1.5
DNA (colony in 50 µL water)	7.5

Table 3-8: Thermocycling conditions for colony PCR on pCR8 transformants.

Step	**Temperature**	**Time**
Initial Denaturation	95°C	5 min
Amplification (x29)	95°C	30s
	50°C	15s
	72°C	2 min 30s
Final Extension	72°C	10 min
Hold	10°C	-

Following colony PCR, the reaction products were run on a 1% agarose gel and screened for the presence or absence of an appropriately sized band in the M13F/Rev1 condition. Those colonies which did contain such a band were selected into overnight cultures, grown with Spectinomycin selection (100 µg/mL in LB). 2 mL of culture was used for Minipreps of each transformant (QIAprep Spin Miniprep Kit, ID: 27104). The concentration of purified DNA preps was quantified using a DeNovix Spectrophotometer (DS-11).

A BigDye™ Terminator v3.1 cycle sequencing reaction (Invitrogen, Catalogue Number 4337455) was then carried out in the isolated Minipreps, as detailed in Table 3-9 and Table 3-10.

Table 3-9: Reaction conditions for BigDye™ Terminator sequencing.

Reagent	**1x (µL)**
BigDye Sequencing Buffer	1.5
BigDye Reagent	1
Sequencing Primer (2 µM)	1.75
DNA (150 ng/µL)	1
dH_20	4.75

Table 3-10: Thermocycling conditions for BigDye™ Terminator sequencing.

Step	**Temperature**	**Time**
Initial Denaturation	96°C	2 min
Amplification (x25)	96°C	10s
	55°C	30s
	60°C	4 min
Hold	15°C	-

The completed sequencing reactions were then sent to Eurofins for Sanger sequencing to verify the correct sequence of the gene inserts in the pCR8 vector. Sequencing was carried out using both the M13F and M13R primers to ensure adequate coverage of the entirety of the gene sequence. Their sequences are, in the 5′ to 3′ direction, ‘CAGGAAACAGCTATGAC’ for M13R and ‘TGTAAAACGACGGCCAGT’ for M13F.

3.2.4.2 Site-directed mutagenesis

In order to obtain the allelic series of mutations in *NAM-A1,* site-directed mutagenesis was carried out using the primers in Table 3-13. These allelic variants included the TILLING mutations, as well as the corresponding alanine mutations. The site-directed mutagenesis was carried out on the full length and truncated versions of *NAM-A1* in the Gateway donor vector pCR8. The reaction conditions are given in Table 3-11 and Table 3-12.

Table 3-11: Reaction components for site-directed mutagenesis PCR.

Component	**1x (µL)**
Q5® Reaction Buffer (5X; NEB M0491)	5
dNTPs (10 mM)	0.5
Forward primer (10 µM)	1.25
Reverse primer (10 µM)	1.25
Q5® High Fidelity DNA Polymerase (NEB M0491)	0.25
Q5® High GC Enhancer (5X; NEB M0491)	5
Template DNA (1 ng/µL)	1
dH_2O	10.5

Table 3-12: Reaction cycling conditions for site-directed mutagenesis.

Step	**Temperature**	**Time**
Initial Denaturation	95°C	5 min
Amplification (x25)	95°C	30 s
	70°C	15 s
	72°C	2 min
Final Extension	72°C	10 min
Hold	10°C	-

5 µL of the PCR reaction was then run on a 1% agarose gel to confirm the presence and correct size of the PCR product. 0.5 µL of *Dpn*I was added to the remaining PCR reaction and the digest was carried out for one and a half hours at 37°C. The digested reaction was then directly transformed into chemically-competent DH5α, as detailed in the previous section. Following recovery at 37°C, the cells were plated onto LB + Spectinomycin plates to select for the pCR8 resistance. Colonies were then cultured overnight and the pCR8 vectors isolated via QIAprep Spin Miniprep (QIAGEN, cat. no. 27104). The success of the site-directed mutagenesis was confirmed using the BigDye terminator kit for Sanger sequencing, as detailed previously (Table 3-9, Table 3-10).

Table 3-13: Primer sequences for site-directed mutagenesis of the *NAM-A1* alleles.

	ID	Line	SNP	Primer F	Primer R	Codon Change	Residue Change
TILLING Mutants	SD1_P38S	Kronos3661	C/T	CGGAGCTCCCATCGGGCTTCCGG	CCGGAAGCCCGATGGGAGCTCCG	Ccg/Tcg	P/S
	SD1_R41Q	Kronos2615	G/A	TCCCACCGGGCTTCCAGTTCCACCC	GGGTGGAACTGGAAGCCCGGTGGGA	cGg/cAg	R/Q
	SD1_T45M	Kronos2734	C/T	TTCCACCCGATGGACGAGGAGCTGGT	CGACCAGCTCCTCGTCCATCGGGTG	aCg/aTg	T/M
	SD3_R110W	Kronos2711	C/T	CGCGGCCGAACTGGGCGGCGACG	CGTCGCCGCCCAGTTCGGCCGCG	Cgg/Tgg	R/W
	SD3_G121D	Kronos3741	G/A	TACTGGAAGGCCACCGACACGGACAA	TTGTCCGTGTCGGTGGCCTTCCAGTA	gGc/gAc	G/D
	SD4_P154L	Kronos2060	C/T	AAGCCGCTCAAGGGCCTCAAAACCA	TGGTTTTGAGGCCCTTGAGCGGCTT	cCc/cTc	P/L
	SD4_P153P	Kronos3186	G/A	CGGGAAGCCGCCCAAGGGCC	TGAGGCCCTTGGGCGGCTTC	ccG/ccA	P
	NC_C133Y	Kronos2551	G/A	ACGGGGTACGGCCTGGTCCG	AGGCCGTACCCCGTCCCCGA	tGc/tAc	C/Y
	ID	**Line**	**SNP**	**Primer F**	**Primer R**	**Codon Change**	**Residue Change**
Alanine Mutants	SD1_P38A	Kronos3661	C/G	CGGAGCTCCCGTCGGGCTTCCGG	CCGGAAGCCCGCTGGGAGCTCCG	Ccg/Gcg	P/A
	SD1_R41A	Kronos2615	CG/GC	TCCCACCGGGCTTCGCGTTCCACCC	GGGTGGAACGCGAAGCCCGGTGGGA	CGg/GCg	R/A
	SD1_T45A	Kronos2734	A/G	TTCCACCCGGCGGACGAGGAGCTGGT	CGACCAGCTCCTCGTCCGCCGGGTG	Acg/Gcg	T/A
	SD3_R110A	Kronos2711	CG/GC	CGCGGCCGAACGCGGCGGCGACG	CGTCGCCGCCGCGTTCGGCCGCG	CGg/GCg	R/A
	SD3_G121A	Kronos3741	G/C	TACTGGAAGGCCACCGCCACGGACAA	TTGTCCGTGGCGGTGGCCTTCCAGTA	gGc/gCc	G/A
	SD4_P154A	Kronos2060	C/G	AAGCCGGCCAAGGGCCTCAAAACCA	TGGTTTTGAGGCCCTTGGCCGGCTT	Ccc/Gcc	P/A
	NC_C133A	Kronos2551	TG/GC	ACGGGGGCCGGCCTGGTCCG	AGGCCGGCCCCCGTCCCCGA	TGc/GCc	C/A

3.2.5 Co-Immunoprecipitation

The methods used to test protein binding *in planta* via co-immunoprecipitation are presented in detail in Chapter 5, where the method is predominantly used. In brief, to test for protein interactions between NAM-A1 and NAM-B1, we cloned the full-length coding sequences of each gene into, respectively, the pGWB21 (N-terminal Myc tag) and pGWB12 (N-terminal FLAG tag) Gateway Binary Vectors (available from https://www.addgene.org/Cristobal_Uauy/) (Nakagawa et al. 2007). These constructs were then transformed into the *Agrobacterium tumefaciens* strain LBA4404 using electroporation and were then plated on selective LB plates with Rifampicin (50 µg/mL) and Kanamycin (100 µg/mL).

Following transformation, individual colonies containing the desired constructs were cultured overnight. After resuspension of the cell pellet in MMA, the constructs were mixed to produce the desired co-infiltration mixture, in this case 10xMyc:*NAM-A1*, FLAG:*NAM-B1*, and the silencing inhibitor P19. With each individual component diluted to an OD_{600} of 0.2, the mixture was then infiltrated into leaves of 3-week-old *N. benthamiana* using the needleless syringe method. Leaves were harvested for protein extraction three days after infiltration, and snap-frozen in liquid N_2.

Protein was extracted from the harvested leaves using a standard protocol, as detailed in section 5.2.4.2.3 of Chapter 5. The extracted protein was then added to washed magnetic anti-FLAG beads, for co-immunoprecipitation. After one hour of incubation at 4°C, shaking, the corresponding input, unbound, and bound samples were isolated. The protein samples were boiled in 1X SDS loading dye to denature, and then run on two 12% acrylamide gels to separate the proteins by weight. Following transfer to a nitrocellulose membrane and blocking, the membranes were incubated with, respectively, an anti-FLAG HRP-conjugated antibody (Abcam, cat. no. AB49763) and an anti-cMyc HRP-conjugated antibody (Abcam, cat. no. AB62928) and left to incubate overnight at 4°C, shaking. The membranes were washed the next day, and then probed with the Pierce™ ECL Western Blotting Substrate (Thermo Scientific™, cat. no. 32109) to visualise the protein bands.

3.2.6 Yeast two-hybrid

To test protein binding affinities, the NAC domains of *NAM-B1*(1-215) and the various *NAM-A1*(1-217) alleles were cloned into the pDEST22 and pDEST32 vectors for expression in yeast. pDEST22 fuses the GAL4 activation domain to the N-terminal of the inserted gene, while pDEST32 fuses the GAL4 DNA binding-domain to the N-terminal of the inserted gene. The cloned alleles included those mutations present in the TILLING lines, as well as the corresponding alanine mutations (Table 3-13). Only the NAC domains of the wild-type and mutant genes were used to prevent auto-activation due to the transcriptional activation properties of the NAC C-terminal domain. The establishment of the yeast two-hybrid system, including protocols for yeast transformation and media recipes, are provided in detail in section 5.2.3 of Chapter 5. We will describe the relevant protocols in brief here.

To test the binding affinities of the NAM-A1 alleles against NAM-B1, we initially co-transformed chemically competent yeast of the appropriate MaV203 strain with the *NAM-A1*(1-217) alleles in either the pDEST22 or pDEST32 vector, alongside the wild-type *NAM-B1*(1-215) allele in the remaining vector. Auto-activation controls were also produced, consisting of either the pDEST22 or pDEST32 vector containing the *NAM-A1*(1-217) alleles alongside the opposing empty vector (i.e. pDEST32_NAM-A1(1-217) and pDEST22). Following transformation, the positive transformants were recovered on SC-Leu-Trp agar plates.

An initial drop assay was carried out to determine whether any binding was seen between the NAM-A1(1-217) alleles and NAM-B1(1-215). Single transformed colonies were picked and resuspended in 200 μL of dH_2O. 2 μL of this dilution was then plated on both control media, SC -Leu -Trp, and on selective media, SC -Leu -Trp -His + 10mM 3-amino-1,2,4-triazole (3AT). Growth of the plated yeast on the selective media was scored as indicative of protein interactions. We tested six independent colonies for each transformation combination and repeated the experiment three times with consistent results.

We also carried out serial dilution assays to determine the relative strength of the protein interactions. For this, single colonies were picked and grown overnight in a liquid culture of SD -Leu -Trp, to a final OD_{600} of 0.1. From this initial culture, serial 1:10 dilutions were carried out in dH_2O. These dilutions, produced for three independent transformed colonies for each combination tested, were plated on both SC -Leu -Trp as a control, and SC -Leu -Trp -His + 10 mM 3AT to test for protein interactions. Based on the previous results from the drop assay, we only tested combinations with *NAM-B1*(1-215) in pDEST32 and *NAM-A1*(1-217) in pDEST22. We repeated this experiment twice and obtained consistent results.

3.2.7 Cell-death scoring in N. benthamiana

To test the impact of the mutations on *NAM-A1* function, we tested their ability to induce cell death in *N. benthamiana*. The full-length alleles were cloned into the Gateway binary vector pGWB21, with an N-terminal 10X Myc tag. They were then transformed into *A. tumefaciens* strain LBA4404 via electroporation and then cultured and prepared for infiltration as described in detail in section 5.2.4.2 in Chapter 5, and briefly above in section 3.2.5.

The infiltration cultures, consisting of solely the gene of interest at an OD_{600} of 0.2, were then patch infiltrated into 3-week-old *N. benthamiana* plants. Patch infiltration was carried out using the needle-less syringe method previously described in the co-immunoprecipitation method. In this case, however, only small patches were produced in the leaf, as depicted in Figure 3-2. The patches were scored for cell death five days following infiltration. The adaxial side of the leaves was photographed under white light to image the level of chlorosis. The abaxial side of the leaves was photographed under ultra-violet light to image the level of auto-fluorescence, indicative of cellular breakdown and lysis.

Figure 3-2: Schematic of patch infiltration in N. benthamiana. Each circle represents a separate infiltration on the leaf, carried out using a syringe containing the Agrobacterium infiltration mixture

The level of cell death, in terms of both chlorosis and auto-fluorescence, was scored based on a discrete scale previously established in Maqbool et al. (2015), where 0 is no cell death and 6 is extreme cell death (Figure 3-3).

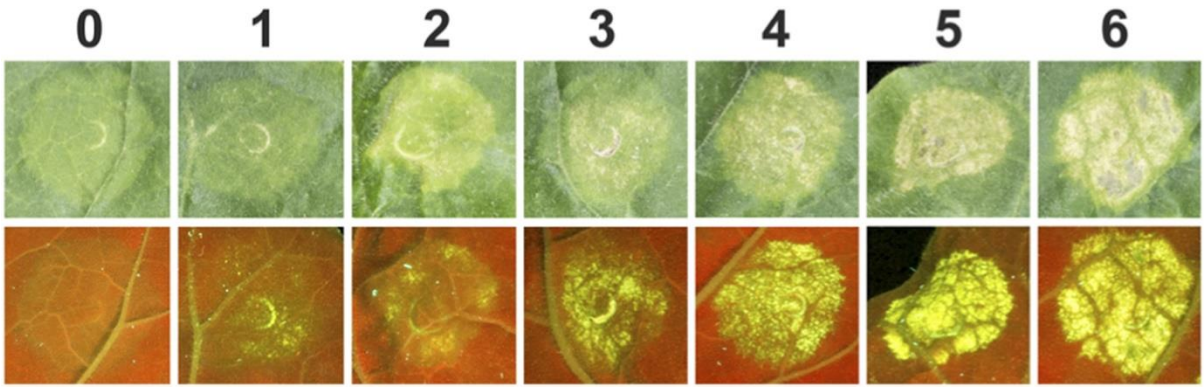

Figure 3-3: Representative images depicting the level of cell death associated with the discrete 0-6 scale. Cell death as visualised under white light (upper) and UV light (lower) is shown in response to co-expression of Pikp and AVR-Pik alleles in various combinations, inducing the hypersensitive response. The figure is reproduced from Maqbool et al. (2015), under a CC-BY-4.0 license.

3.2.8 Data analysis and visualisation

All analysis of the data in this chapter was carried out using R (v3.5.1) (R Core Team 2018), with graphs produced using ggplot2 (Wickham 2016). Appropriate statistical analyses were carried out for the data and are detailed within the results section. We used non-parametric statistical tests where assumptions for normality could not be met during the analysis of glasshouse and field trials phenotypic data. To account for the discrete nature of the data obtained by the *N. benthamiana* cell death experiment we carried out comparisons of the mutant alleles against wild-type *NAM-A1* by fitting a linear model using the Poisson distribution.

3.3 Results

3.3.1 The NAC domain of NAC transcription factors is highly conserved throughout the plant kingdom.

NAC transcription factors are present throughout the Plant Kingdom and are characterised by a highly conserved N-terminal NAC domain. Yet while this domain is necessary for protein function, it is not yet clear which specific residues are required for protein function. We hypothesized that residues which are highly conserved in NAC transcription factors across the plant kingdom are more likely to be essential for protein function. To investigate this hypothesis further, we extracted all protein sequences which contained a NAC domain from seven plant species. This set of species included monocots (*Triticum aestivum, Hordeum vulgare, Oryza sativa var. japonica, Zea mays*), dicots (*Populus trichocarpa, Arabidopsis thaliana*), and a bryophyte (*Physcomitrella patens*). These species were selected as they span a broad range of angiosperm and plant kingdom diversity and have a well-annotated set of NAC transcription factors. We identified a total of 1404 NAC transcription factors from the seven species, ranging from 35 in *P. patens* to 537 in *T. aestivum* (Harrington et al. 2019c).

We then extracted the conserved NAC domain from each of these sequences, encompassing all sequence from the start codon up to and including the canonical "CRVFKK" sequence which marks the end of subdomain v. The NAC domains were then aligned, and the conservation of each residue was calculated using the Jensen-Shannon divergence (JSD) method (Figure 3-4) (Capra and Singh 2007). We focussed only on residues that were present in at least 10% of sequences, a cut-off previously used in NAC transcription factor analyses (Borrill et al. 2017). From the resulting alignment we obtained an average JSD conservation score of 0.21 across the entire NAC domain. Within the NAC domain, five regions are particularly highly conserved, the subdomains i through v. These regions, highlighted in grey in Figure 3-4, were substantially more conserved than the NAC domain as a whole. The most heterogeneous and largest region, subdomain iii, had a mean JSD score of 0.24, while the most conserved subdomain i had a mean JSD value of 0.39.

We then identified residues in the NAC domain that had previously had their role in protein-protein or protein-DNA interactions tested (Olsen et al. 2005a, Welner et al. 2012). Residues which had been shown to inhibit these interactions when mutated are underlined in red in Figure 3-4, while those which had no effect are underlined in blue. Only three of the tested residues were essential for these interactions; two in subdomain i which were required for protein-protein interactions (Olsen et al. 2005a), and one in subdomain iii which was required for protein-DNA interactions (Olsen et al. 2005a, Welner et al. 2012).

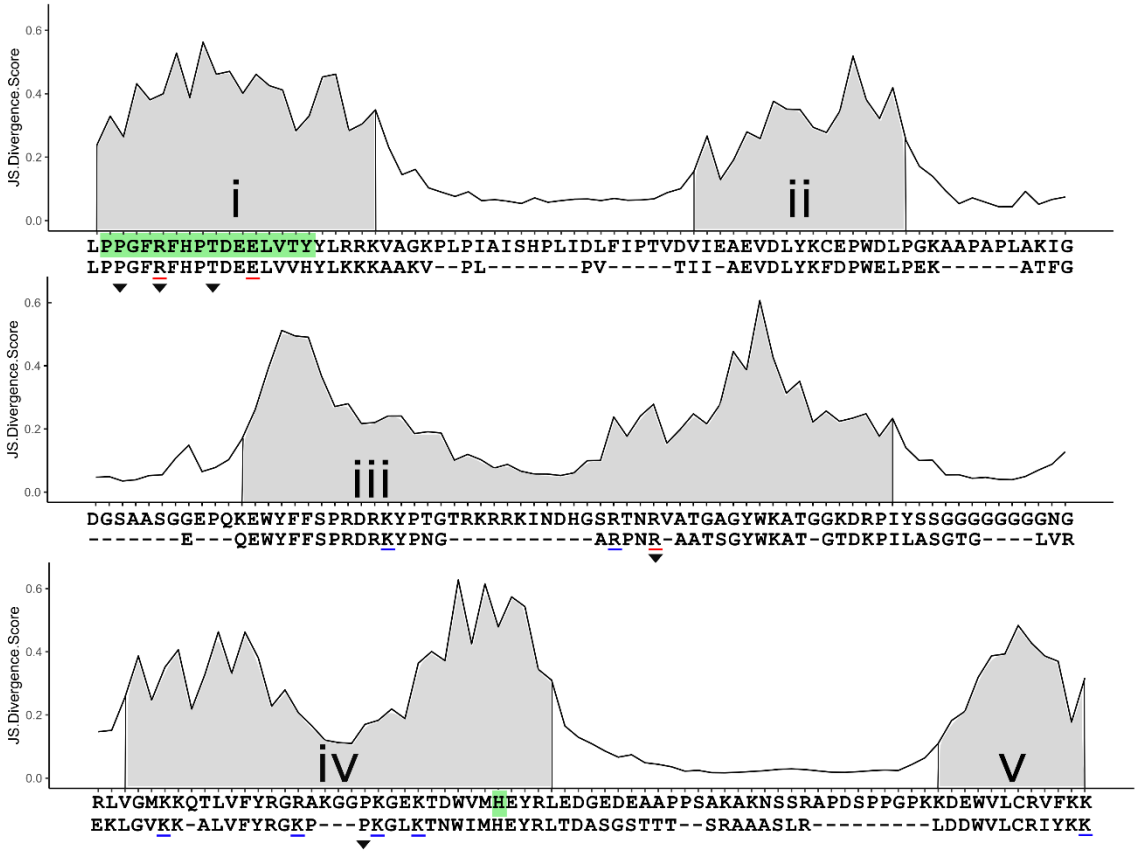

Figure 3-4: Identification of highly conserved residues of NAM-A1 in the cultivar Kronos TILLING population. The Jensen-Shannon Divergence scores derived from a consensus alignment of 1404 NAC-domain-containing proteins from seven plant species across the Plant Kingdom show that the five NAC subdomains, highlighted in grey (i-v), are highly conserved. The consensus alignment (first row) shows the most frequent amino acid at that position; note that residues which were present in less than 10% of the proteins were excluded from the consensus sequence. Residues that were found to be required for protein-protein interactions (i) or protein-DNA interactions (iii – v) are underlined in red, those not required are underlined in blue (Olsen et al. 2005a, Welner et al. 2012). A histidine residue in subdomain iv is known to play a role in protein dimer stabilisation through interactions with part of the protein dimerization domain (subdomain i; both highlighted in green) (Kang et al. 2018). The aligned NAC-domain of NAM-A1 is shown below the consensus sequence. Residues for which missense mutations were identified in the Kronos TILLING population are highlighted with an inverted triangle (black). Note that the non-conserved residue (C133Y) containing a mutation in K2551 is not present in the consensus alignment as it is present in fewer than 10% of the NAC proteins tested.

3.3.2 The Kronos TILLING population contains missense mutations in highly conserved residues of the NAM-A1 NAC domain

The alignment of the 1404 NAC transcription factors from species across the Plant Kingdom demonstrated that the amino acid residues within the five NAC subdomains are the most highly conserved within the entire NAC domain. This high level of conservation in the NAC subdomains is clear when comparing the consensus NAC sequence, derived from the most common amino acids in the 1404 aligned sequences, with the individual sequence of *NAM-A1*. We then investigated whether missense mutations in highly conserved residues of the NAC domain existed within the Kronos TILLING population. As the Kronos population is tetraploid, and the B-genome

homoeolog of *NAM-A1* is non-functional, we could test the effect of missense mutations on *NAM-A1* in a diploid-like context.

We initially used the www.wheat-tilling.com website (Krasileva et al. 2017) to identify the missense mutations present in *NAM-A1* in the Kronos population, and confirmed the presence and effect of these mutations in the Ensembl Plants variation browser (Kersey et al. 2017). We identified five missense mutations located in highly conserved residues of the NAC domain which were predicted to be highly deleterious based on SIFT scores (Figure 3-5A, Table 3-14). Three of the residues are located in the highly conserved subdomain i, one of which (K2615) is the same residue that was previously shown to have a role in protein binding (Figure 3-4) (Olsen et al. 2005a). A fourth residue (K2711) is located in the middle of subdomain iii, in a residue that was found to be essential for DNA binding (Welner et al. 2012). The fifth residue, K2060, is located in subdomain iv. We found that each of the five positions in the NAC domain were highly enriched for a specific amino acid, with the K2711 arginine residue showing the highest conservation, present in 85% of the NAC sequences tested (Figure 3-5).

Alongside the missense mutations in highly conserved residues, we also identified a set of control lines that would allow us to tease out the impact of these mutations on NAM-A1 function. We selected a line containing a synonymous mutation in NAM-A1 (K3186) as a negative control which should have no effect on protein function. We also identified a missense mutation in a highly non-conserved amino acid (K2551) between subdomains iii and iv. Not only did this mutation have a high SIFT score of 0.78, suggesting that the mutation is highly likely to not be deleterious (Table 3-14), but we also found that the residue itself is highly non-conserved among NAC transcription factors. Only 139 of the 1404 NAC sequences contained a residue at this position, and of these sequences, the cysteine identified in *NAM-A1* was only present in 2 of the sequences (Figure 3-5).

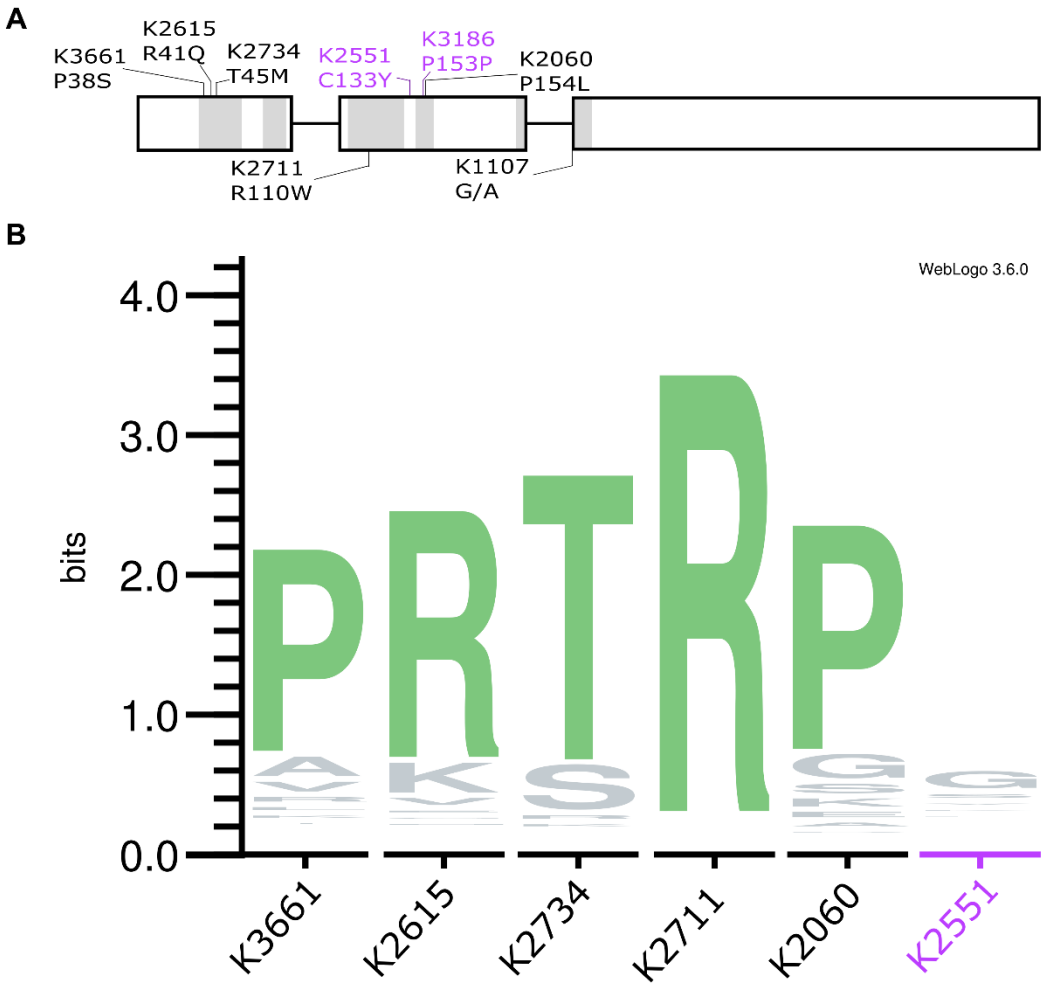

Figure 3-5: The Kronos TILLING population contains missense mutations in highly conserved residues of the NAM-A1 NAC domain. (A) Missense mutations, as well as non-conserved and synonymous controls (purple) and the splice acceptor variant K1107, are shown on a schematic of the NAM-A1 structure. (B) The missense mutations are located in highly conserved residues of the NAC domain, with the wild-type residue the primary residues found in the consensus NAC domain. The non-conserved control, K2551 (purple), is present at a location with very low residue conservation. The y-axis, measured in bits, is proportional to the sequence conservation at each site with a maximum score of 4.32 corresponding to a perfectly conserved residue. The overall height of each stack is proportionate to the level of sequence conservation, while the height of each individual letter is proportionate to the observed frequency of the amino acid at that position. The sequence logo was created using WebLogo v3.6.0 from weblogo.berkeley.edu.

Table 3-14: NAM-A1 mutation location and characterisation. Details on the individual TILLING SNPs are provided, alongside the heterozygosity of the SNP in the M2 sequencing bulk, predicted consequence, and predicted impact based on conservation and SIFT scores. Links to the variant on EnsemblPlants are also provided. Predicted region is derived from (Ernst et al. 2004).

Line	SNP	Codon mutation	Het/Hom	Consequence	cDNA position	CDS position	Amino acid mutation	SIFT	JSD Score	Region	RefSeqv1 chr6 (bp)	Ensembl Link
Kronos3661	C/T	Ccg/Tcg	Hom	Missense	339	112	P38S	0	0.26	Dimerisation	77100016	https://bit.ly/2GLlY71
Kronos2615	G/A	cGg/cAg	Het	Missense	349	122	R41Q	0	0.4	Dimerisation	77100006	https://bit.ly/2X6gdX3
Kronos2734	C/T	aCg/aTg	Het	Missense	361	134	T45M	0	0.46	Dimerisation	77099994	https://bit.ly/2S2xUTL
Kronos2711	C/T	Cgg/Tgg	Het	Missense	555	328	R110W	0	0.27	DNA interface	77099581	https://bit.ly/2X1ZuUW
Kronos2551	G/A	tGc/tAc	Het	Missense	625	398	C133Y	0.78	-	-	77099511	https://bit.ly/2Ig556S
Kronos2060	C/T	cCc/cTc	Het	Missense	688	461	P154L	0	0.17	DNA interface	77099448	https://bit.ly/2SXh2lS
Kronos3186	G/A	ccG/ccA	Het	Synonymous	686	459	P153P	-	0.16	-	77099450	https://bit.ly/2N6RCwK
Kronos1107	G/A	N/A	Het	Splice Acceptor Variant	N/A	N/A	N/A	-	-	DNA interface	77099212	https://bit.ly/2GtGqtO

We also selected a TILLING line that contained a mutation we predicted would cause a complete loss of function. Line K1107 contained a splice acceptor variant mutation, where the canonical splice acceptor site for exon 3 was mutated and lost. The nearest predicted splice site would maintain the correct frame of reference, but led to an 11-residue deletion, encompassing the majority of subdomain v. We therefore predicted that this deletion would be sufficient to substantially, if not completely, disrupt the function of NAM-A1.

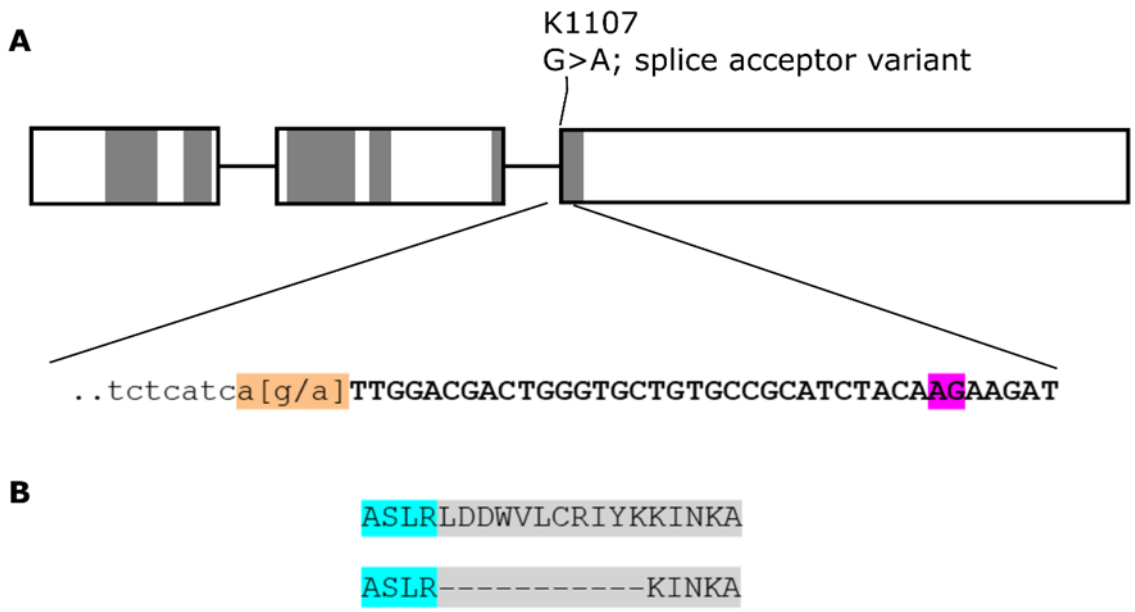

Figure 3-6: The splice acceptor variant mutation in K1107 is predicted to lead to an 11-residue deletion. (A) The Kronos TILLING line K1107 contains a G to A substitution at the splice acceptor site of the third exon of NAM-A1. This mutation (orange) is predicted to lead to the use of an alternate splice site (purple). (B) This predicted splice isoform contains an 11-residue deletion encompassing the majority of subdomain v.

In the remainder of the paper, the TILLING lines will be referred to with both the TILLING line and corresponding mutation (e.g. K2711$_{R110W}$). Line K1107, containing a splice acceptor variant (SAV), will be referred to as K1107$_{SAV}$.

3.3.3 Mutations in conserved residues of NAM-A1 lead to delayed senescence in wheat.

Having identified a set of lines in the Kronos TILLING population that contained mutations in *NAM-A1*, we then began characterising the effect of these mutations on monocarpic senescence. We first carried out a preliminary screen of the mutant lines, growing plants directly from the M_5 seed. The plants were scored for visual flag leaf and peduncle senescence, grown under glasshouse conditions (Figure 3-7, Table 3-15). In the initial screen, we tested only four of the missense mutations (K3661$_{P38S}$, K2734$_{T45M}$, K2711$_{R110W}$, and K2060$_{P154L}$) alongside the three control lines. As there were no wild-type sister lines to K3661, a homozygous mutation, we compared the senescence phenotype in the mutant lines to the senescence phenotypes of all the wild-type lines

derived from the heterozygous mutants ("Combined WT"). In such a way, we aimed to account for variable senescence timings caused by background mutations in the TILLING lines. We also compared the $K1107_{SAV}$ line against this "Combined WT" set as we were unable to grow wild-type sister lines for $K1107_{SAV}$ in the initial experiment.

We found that three of the missense mutation lines were significantly delayed in flag leaf senescence, while all of the missense lines were significantly delayed in peduncle senescence ($p < 0.05$, two-sample Wilcoxon test). The splice junction mutation, line $K1107_{SAV}$, surprisingly did not lead to a significant delay in flag leaf senescence, though it did show a significant delay in peduncle senescence ($p < 0.001$, two-sample Wilcoxon test). The two negative controls had no significant difference in senescence timing between the wild-type and the mutant lines. Overall, we saw the most substantial delay in senescence in the $K2711_{R110W}$ line, where the mutants were delayed in flag leaf senescence by almost 15 days compared to wild-type and in peduncle senescence by over 30 days (Table 3-15). However, this initial screen was hampered by the inconsistent sample sizes for each of the genotypes, with some genotypes such as the $K2711_{R110W}$ mutant itself only containing three replicates. This was due in part to poor germination of the M_5 seeds for some of the TILLING lines.

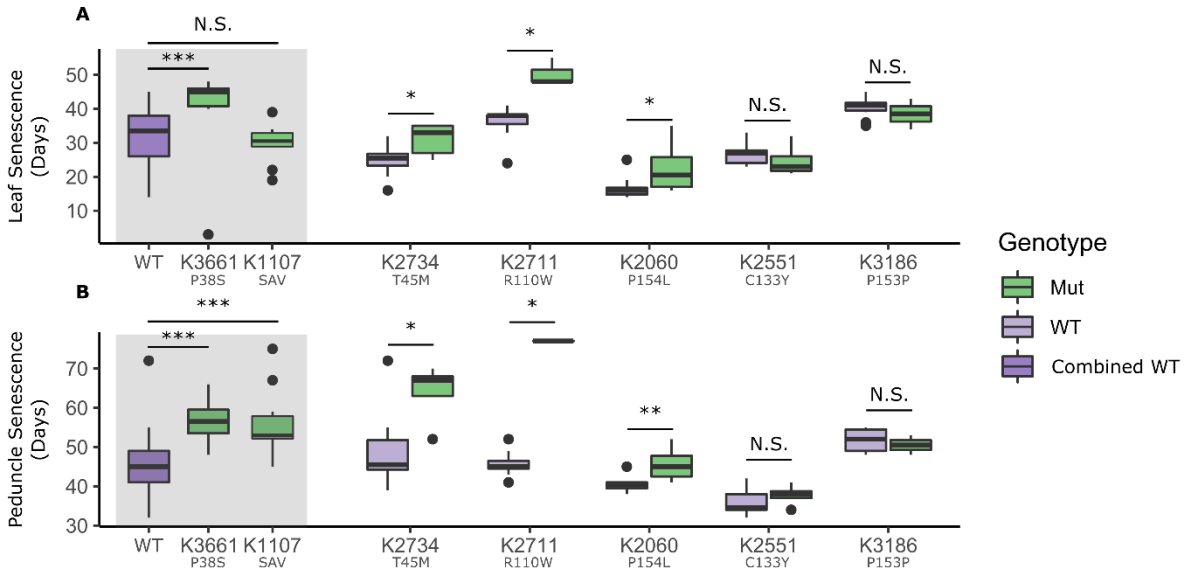

Figure 3-7: Lines $K2711_{R110W}$, $K2734_{T45M}$, and $K3661_{P38S}$ are consistently delayed in peduncle senescence. The selected TILLING lines containing mutations in NAM-A1 were grown and phenotyped for flag leaf (A) and peduncle (B) senescence. The mutant plants for lines $K3661_{P38S}$ and $K1107_{SAV}$ are compared to the "Combined WT" of all WT plants from the remaining lines (see results for details). N is between 3 and 10. Statistical tests were carried out using the two-sample Wilcoxon test; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Table 3-15: Summary statistics for the initial *NAM-A1* mutant screen.

Variable	TILLING Line	Genotype	N	Mean	Standard Deviation
Days from Heading to Leaf Senescence	Average	WT	66	31.6	8.4
	$K3661_{P38S}$	Mut	10	36.5	17.8
	$K1107_{SAV}$	Mut	10	29.8	5.8
	$K2734_{T45M}$	Mut	5	31.0	4.7
		WT	10	24.6	4.3
	$K2711_{R110W}$	Mut	3	50.3	4.0
		WT	11	36.3	4.7
	$K2060_{P154L}$	Mut	8	22.6	7.2
		WT	8	16.9	3.6
	$K2551_{C133Y}$	Mut	4	24.8	5.0
		WT	6	26.8	3.7
	$K3186_{P153P}$	Mut	3	38.5	6.4
		WT	10	40.6	3.3
Days from Heading to Peduncle Senescence	Average	WT	66	45.6	6.4
	$K3661_{P38S}$	Mut	10	56.7	5.7
	$K1107_{SAV}$	Mut	10	56.3	8.7
	$K2734_{T45M}$	Mut	5	64.0	7.2
		WT	10	48.7	8.0
	$K2711_{R110W}$	Mut	3	77.0	0.0
		WT	11	45.6	2.9
	$K2060_{P154L}$	Mut	8	44.4	2.6
		WT	8	40.5	2.2
	$K2551_{C133Y}$	Mut	4	37.8	2.9
		WT	6	36.0	3.7
	$K3186_{P153P}$	Mut	3	50.5	3.5
		WT	10	51.7	2.8

To investigate this further, we repeated the glasshouse experiment. This time we were also able to score the fifth missense mutation, $K2615_{R41Q}$, which had not been tested in the previous experiment. Unlike in the initial experiment, we found no significant delays in flag leaf senescence for any of the lines tested (Table 3-16, Figure 3-8A). This was unexpected, given the known role of *NAM-A1* in controlling flag leaf senescence and the large delays in senescence observed in the previous experiment. We did see a large delay in flag leaf senescence in the $K2615_{R41Q}$ line, which had not been tested in the first experiment, of almost 20 days (Table 3-16). However, we were only able to recover three mutant plants for this line, and as such the result was not statistically significant ($p = 0.07$, two-sample Wilcoxon test).

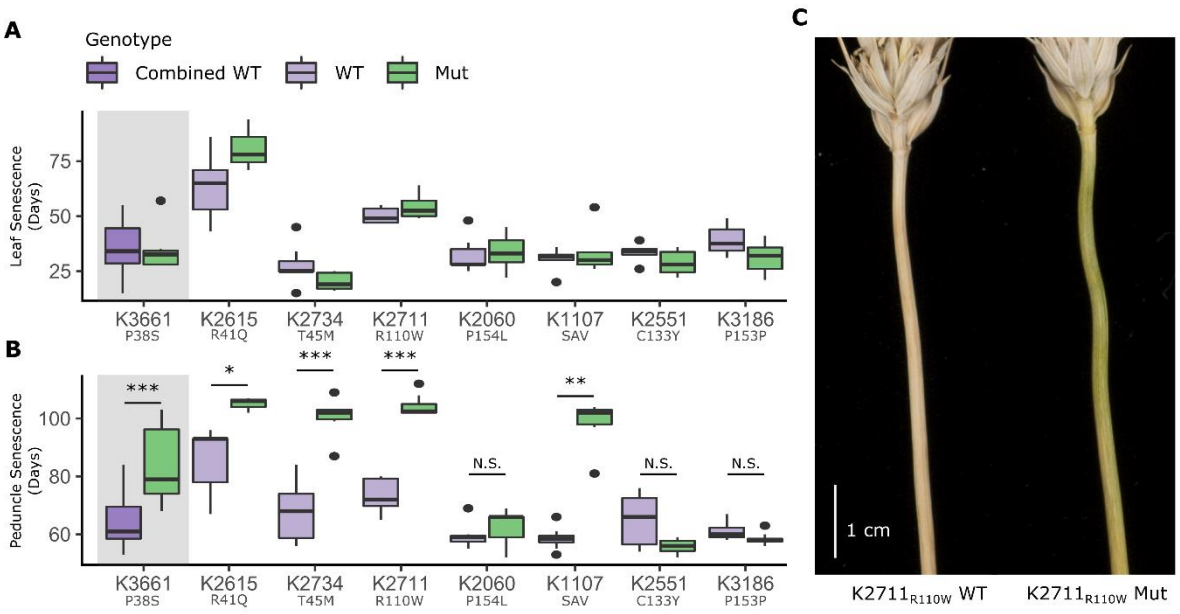

Figure 3-8: Mutations in the NAC domain of NAM-A1 delay peduncle senescence in the glasshouse. The TILLING lines were again scored for flag leaf senescence (A) and peduncle senescence (B) in terms of days after heading. The plants were K3661, which is homozygous for the NAM-A1 mutant allele, was compared to a combination of all the wild-type lines as there were no sibling wild-type lines for direct comparison (see results for details; grey background). The delay in peduncle senescence for $K2711_{R110W}$ is pictured in (C) at 90 days after anthesis. All statistics shown are two-sample Wilcoxon tests; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. N is between 8-10 plants for all lines except K2615 Mut where n = 3.

We did see significant delays in peduncle senescence for all of the missense mutations except $K2060_{P154L}$, as well as for the splice acceptor mutant $K1107_{SAV}$ (Figure 3-8B). Of the missense mutations, lines $K2734_{T45M}$ and $K2711_{R110W}$ were the most delayed, with a difference of 32 and 31 days on average between wild-type and mutant peduncle senescence, respectively (Table 3-16). This slightly smaller than the delay seen in the splice acceptor mutant, where the mutant $K1107_{SAV}$ plants were delayed in peduncle senescence by approximately 40 days. $K2060_{P154L}$ was the only conserved missense mutation that did not lead to significantly delayed peduncle senescence, though the mutant peduncles senesced on average 2.5 days later than the wild-type ($p = 0.31$, two-sample Wilcoxon test; Table 3-16). As expected, the negative controls $K3186_{P153P}$ and $K2155_{C133Y}$ showed no delay in senescence compared to wild-type (Figure 3-8, Table 3-16).

Table 3-16: Summary statistics for the second *NAM-A1* mutant screen.

Variable	TILLING Line	Genotype	N	Mean	Standard Deviation
Days from Heading to Leaf Senescence	3661_{P38S}	Mut	8	34.4	9.6
		WT	49	35.7	9.7
	2615_{R41Q}	Mut	3	81.0	11.8
		WT	11	62.9	13.3
	2734_{T45M}	Mut	8	20.3	3.8
		WT	8	27.6	8.7
	2711_{R110W}	Mut	8	54.0	5.1
		WT	8	50.3	3.7
	2060_{P154L}	Mut	9	33.9	7.5
		WT	8	32.4	8.0
	1107_{SAV}	Mut	6	33.7	10.4
		WT	8	30.6	4.7
	2551_{C133Y}	Mut	8	28.8	5.8
		WT	9	33.4	4.0
	3186_{P153P}	Mut	8	31.1	7.1
		WT	8	39.0	6.7
Days from Heading to Peduncle Senescence	3661_{P38S}	Mut	8	83.3	13.6
		WT	49	64.4	8.3
	2615_{R41Q}	Mut	3	105.0	2.6
		WT	11	86.0	11.9
	2734_{T45M}	Mut	8	100.6	6.3
		WT	8	68.4	10.8
	2711_{R110W}	Mut	8	104.4	3.7
		WT	8	73.4	5.7
	2060_{P154L}	Mut	9	62.4	5.7
		WT	8	59.6	4.5
	1107_{SAV}	Mut	6	98.2	8.8
		WT	8	58.6	3.9
	2551_{C133Y}	Mut	8	55.8	2.6
		WT	9	65.0	8.8
	3186_{P153P}	Mut	8	58.4	2.3
		WT	8	61.0	3.0

We complemented our visual senescence scores with quantitative measures of leaf chlorophyll content using the SPAD-502 meter. Individual plants from each line and genotype were measured at five-day intervals following anthesis until 40 days after anthesis (DAA), at which point the majority of flag leaves from all lines except $K2615_{R41Q}$ had senesced completely (Figure 3-9). We found no significant differences between the wild-type and mutant plants at any timepoint for any line, consistent with the lack of flag leaf senescence phenotype observed using visual scoring (Figure 3-8A). We did observe a trend for the mutant plants in lines $K1107_{SAV}$, $K2711_{R110W}$, and $K2615_{R41Q}$ to have higher chlorophyll measurements across the timecourse, with the gap widening as senescence progressed, however these differences were not significant when adjusted for the

false discovery rate using the Benjamini-Hochberg adjustment. Nevertheless, these results suggest that the *NAM-A1* mutations do lead to a slight delay in flag leaf senescence that wasn't captured in the visual scoring.

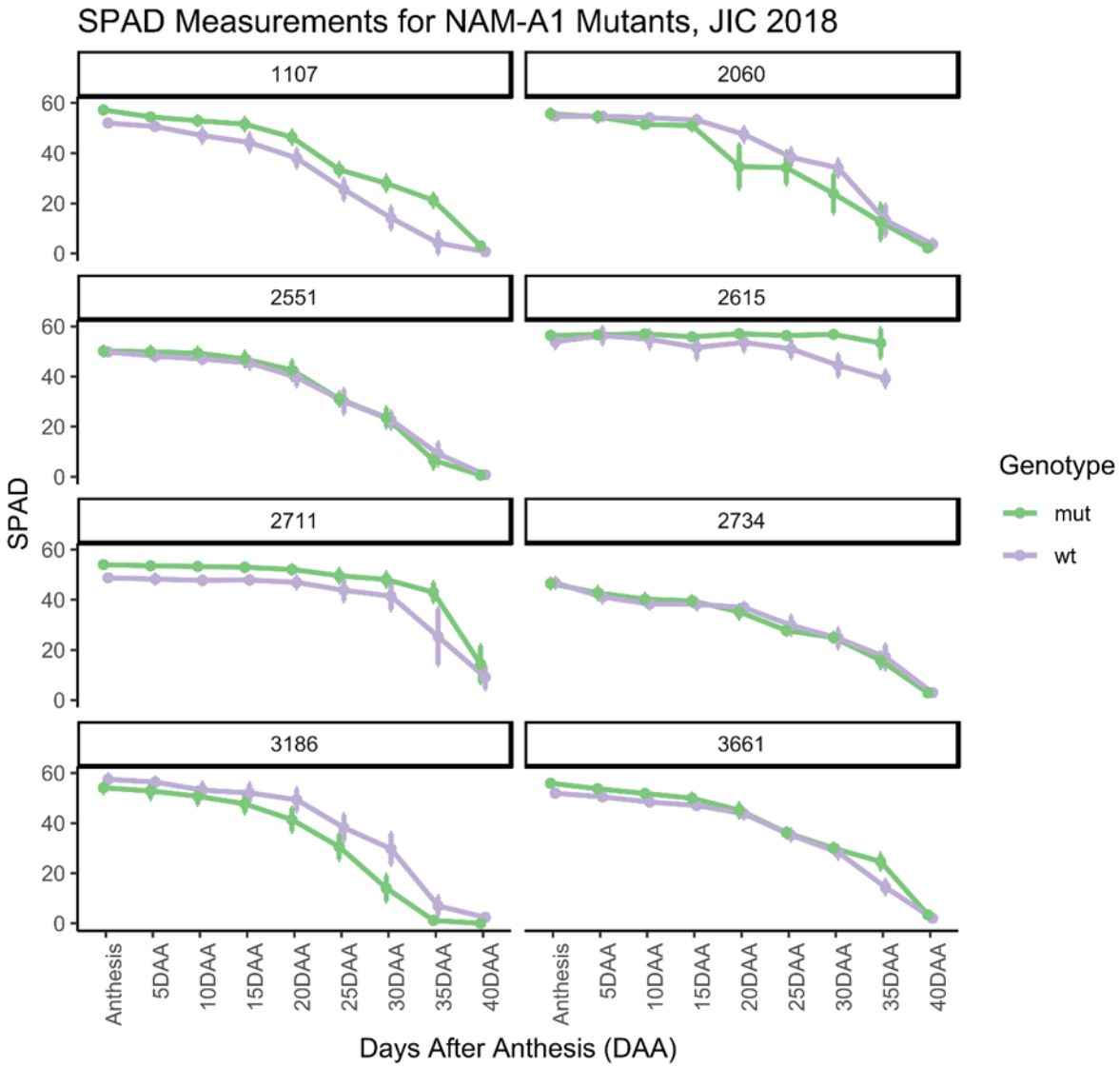

Figure 3-9: No significant differences in chlorophyll content are seen for the *NAM-A1* TILLING mutants. SPAD measurements were taken of the primary flag leaf at intervals from anthesis to 40 days after anthesis. Across all timepoints, no significant differences in SPAD units were observed for the mutant plants compared to wild-type for any of the TILLING lines. N = 5 for all lines, genotypes, and timepoints. Statistical comparisons were carried out using the Kruskal-Wallis Rank Sum test adjusted for FDR. Error bars depict the standard error of the mean.

3.3.4 The effect of the mutations on NAM-A1 function is consistent in the field

From the experiments carried out under glasshouse conditions, we had shown that the mutations in conserved residues of the *NAM-A1* NAC domain were able to recapitulate the delayed senescence phenotype. We then considered what impact these mutations might have on senescence timing under field conditions. We had previously identified the $K1107_{SAV}$ and $K2711_{R110W}$ lines

independently in a screen of the TILLING population for senescence mutants (see Chapter 2). We therefore chose to focus solely on these two mutant lines for further characterisation in the field. These lines were then backcrossed to the wild-type Kronos parent, generating a segregating F_2 population which we could then score for senescence in the field. We showed that the mutations in *NAM-A1* segregated with the corresponding early and late peduncle senescence phenotypes (figure in Chapter 2). From this, we then phenotyped only the plants which contained the mutant or wild-type allele of *NAM-A1*. In the initial phenotypic screen at the JIC in 2016, we found that the F_2 plants containing the mutant *NAM-A1* alleles were significantly delayed in peduncle senescence, corresponding with the observed phenotypes in the glasshouse (Figure 3-10B). We did see a significant delay in flag leaf senescence for the $K2711_{R110W}$ mutants, but there was no observed delay for the $K1107_{SAV}$ mutants (Figure 3-10A).

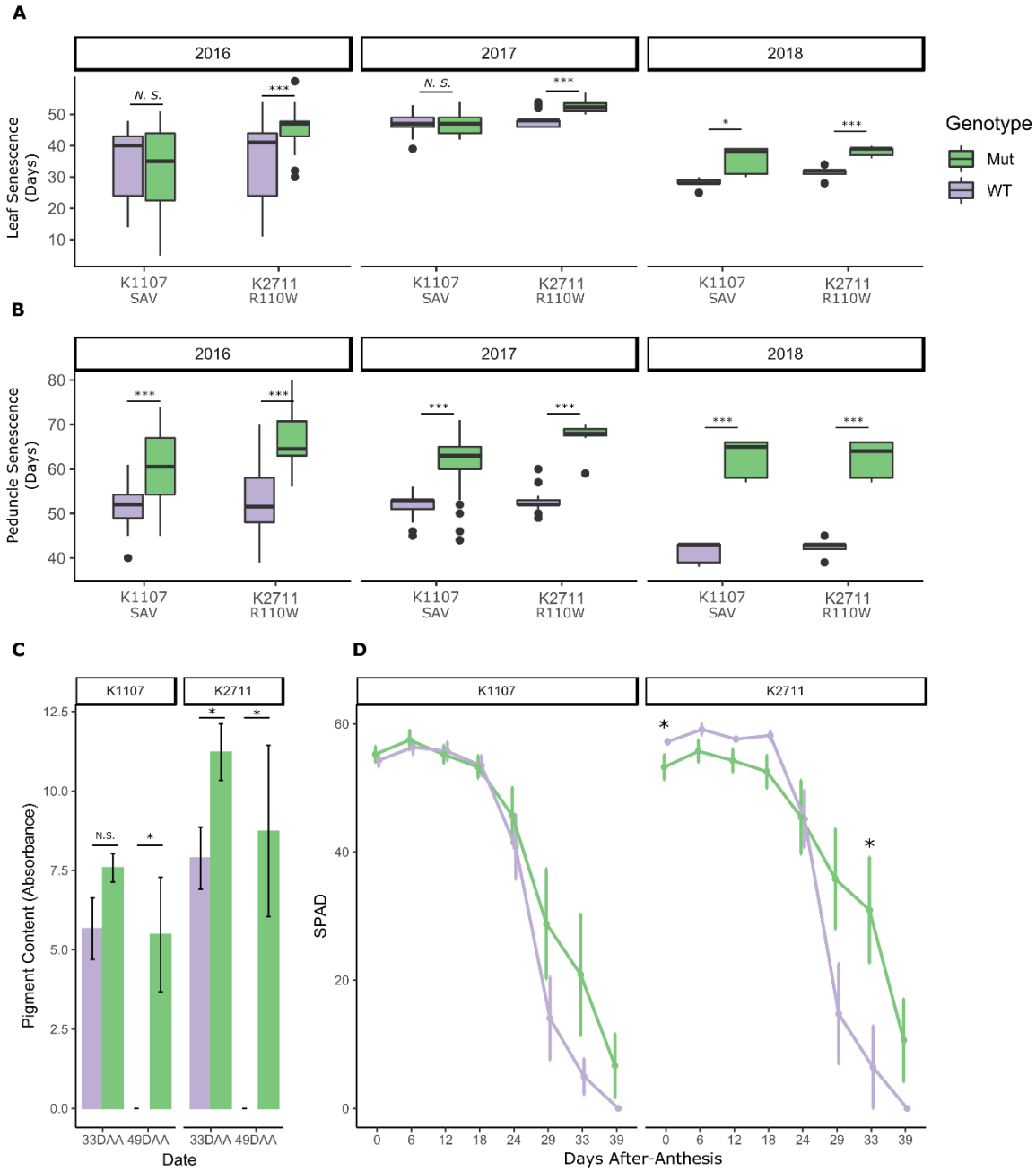

Figure 3-10: Mutations in *NAM-A1* lead to a significant delay in peduncle senescence in the field. TILLING lines $K1107_{SAV}$ and $K2711_{R110W}$ were scored for flag leaf (A) and peduncle (B) senescence, relative to heading, in the field across three years. Peduncle chlorophyll content (C) was quantified at 33 and 49 days after anthesis (DAA) for each genotype of each line in the 2018 field trial. SPAD measurements of flag leaf chlorophyll content (D) were carried out in 2018 at intervals throughout senescence. See methods for the sample size for each year. Statistical comparisons were carried out with a Student's t-test for all panels except the SPAD time course, which used the Kruskal-Wallis Rank Sum test; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

The mutant and wild-type genotypes for each line were characterised across two further field seasons, one at the JIC (2018) and one in California, at the UC Davis field station (2017). We found that the delay in peduncle senescence was consistent across all years and environments, with a highly significant delay in senescence observed for both mutant alleles ($p < 0.001$, Student's t-test; Figure 3-10B). The effect of the mutations on flag leaf senescence was less consistent, corresponding with the inconsistent phenotype observed in the glasshouse trials. The $K2711_{R110W}$ mutant allele consistently led to a significant delay in flag leaf senescence across all trials ($p < 0.001$, Student's t-test; Figure 3-10A). In contrast, the $K1107_{SAV}$ mutant allele only led to a significant delay in flag leaf senescence in the final 2018 trial, and then to a lesser extent than the $K2711_{R110W}$ mutant allele ($p < 0.05$, Student's t-test; Figure 3-10A).

To investigate the effect of the mutant alleles on peduncle senescence further, we quantified the levels of chlorophyll in the peduncle during the 2018 field trial. We found that by 33 DAA, the $K2711_{R110W}$ mutant plants retained more chlorophyll than the wild-type plants ($p < 0.05$, Student's t-test; Figure 3-10C). No significant difference was seen between the $K1107_{SAV}$ genotypes by this stage. However, by 49 DAA, in both lines the wild-type alleles showed no retention of chlorophyll in the peduncle while the mutant plants still retained some chlorophyll ($p < 0.05$, Student's t-test; Figure 3-10C). This corresponded with both mutations leading to delayed peduncle senescence, but with the $K2711_{R110W}$ mutation potentially having a stronger mutant phenotype.

We also quantified the levels of chlorophyll in the flag leaves of the mutant and wild-type plants during the 2018 field trial using a SPAD meter. While the wild-type $K2711_{R110W}$ plants actually contained higher levels of chlorophyll at anthesis, by 33 DAA the mutant lines retained significantly more chlorophyll ($p < 0.05$, Kruskal-Wallis rank-sum test; Figure 3-10D). The $K1107_{SAV}$ mutants did not contain significantly more chlorophyll at any stage of the timecourse, though at 33DAA the SPAD scores of the mutant plants were on average 15 points higher than the wild-type ($p = 0.25$, Kruskal-Wallis rank-sum test; Figure 3-10D). This discrepancy between the SPAD profiles of the two *NAM-A1* mutants corresponds with the stronger visual phenotype of flag leaf senescence seen for line $K2711_{R110W}$.

3.3.5 Mutations in NAM-A1 lead to decreased protein content in the grain

Following characterisation of the senescence profiles of the $K2711_{R110W}$ and $K1107_{SAV}$ mutants in the field, we also investigated whether the delay in senescence observed in the mutants affected grain size and quality parameters. We first measured the size and weight of the grains from both field trials carried out at the JIC and found no consistent effects of the *NAM-A1* mutant alleles (Figure 3-11). All parameters measured were decreased in the $K2711_{R110W}$ mutant lines in 2016, but this was due to a co-segregating dwarfing mutation in the TILLING background. Following selection of only plants without the dwarf phenotype, these reductions in grain size and weight were lots in the following trials. Indeed, by 2018 it appeared that the *NAM-A1* mutations may lead

to a slight increase in grain size and weight, though not length. These differences were relatively small and mostly non-significant; this corresponds with previous work that found no consistent effects of mutations in the *NAM* genes on grain size parameters (Avni et al. 2014).

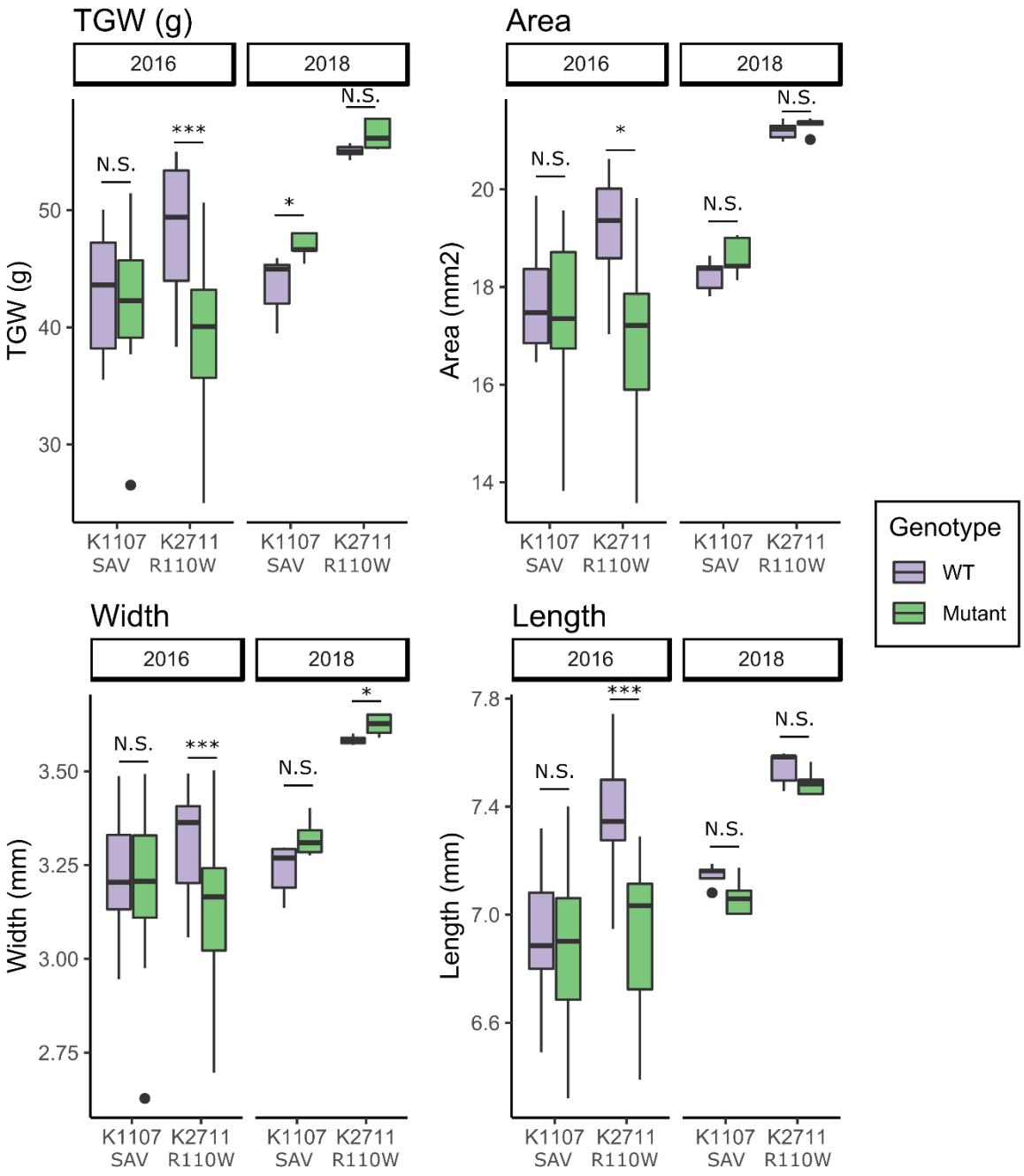

Figure 3-11: No consistent variation in grain size parameters is seen in the *NAM-A1* mutant lines. Grain samples from wild-type and mutant plants of both the $K1107_{SAV}$ and $K2711_{R110W}$ lines were measured in 2016 and 2018 for four morphometric parameters: thousand grain weight (TGW), grain area, grain width, and grain length. Statistical comparisons were carried out using the Kruskal-Wallis Rank Sum test; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

We then measured the protein content of the grains from both field seasons using a Near-Infrared (NIR) spectrometer. Previous work had shown that knock-downs and mutations of the *NAM* genes can lead to reduced grain protein and nutrient content (Uauy et al. 2006b, Avni et al. 2014, Pearce

et al. 2014). We were able to recapitulate that result, showing that the grain protein levels in both of the *NAM-A1* mutants were decreased compared to the wild-type across two years ($p < 0.05$, two-sample Wilcoxon test). The only exception was $K1107_{SAV}$ in 2016, a year when the mutant plants showed no delay in flag leaf senescence. While there was no significant decrease in protein content ($p = 0.24$, two-sample Wilcoxon test), there was a trend towards reduced protein content in the $K1107_{SAV}$ mutant plants (Figure 3-12).

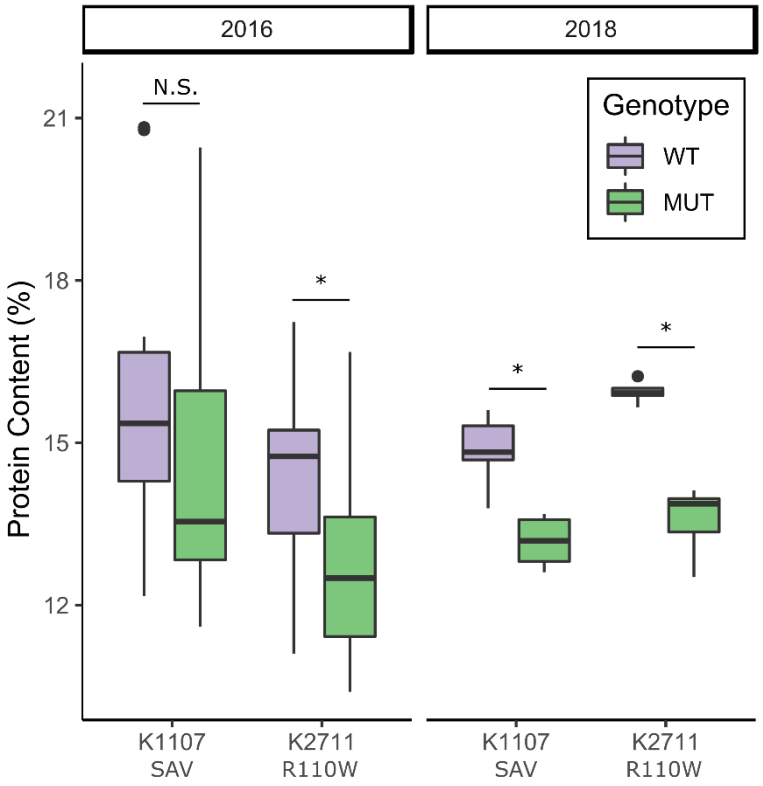

Figure 3-12: Grain protein content (GPC) is decreased in the $K2711_{R110W}$ mutant line under field conditions. GPC was measured from mature grain samples of wild-type and mutant plants from the 2016 and 2018 field trials of lines $K1107_{SAV}$ and $K2711_{R110W}$. Statistical comparisons were carried out with the two-sample Wilcoxon test; *, $p < 0.05$. N = 5 for 2018 field trials and varies between 10 and 22 for the 2016 field trials.

3.3.6 Mutations in the NAC domain of NAM-A1 prevent protein binding

Having demonstrated that the missense mutations in the NAC domain of *NAM-A1* were sufficient to disrupt the function of *NAM-A1*, we then began to investigate how these missense mutations may be affecting protein function. We had previously found in the lab that NAM-A1 is able to physically interact with its B-genome homoeolog NAM-B1; these results will be discussed in detail in Chapter 5. This interaction was initially identified in a yeast two-hybrid (Y2H) library screen, and then validated using co-immunoprecipitation (Figure 3-13). We were able to co-immunoprecipitate a Myc-tagged version of NAM-A1 alongside FLAG-tagged NAM-B1, following heterologous expression in *Nicotiana benthamiana*. This provided us with a system with which we could test the impact of the missense mutations on the ability of the NAM-A1 protein to bind NAM-B1.

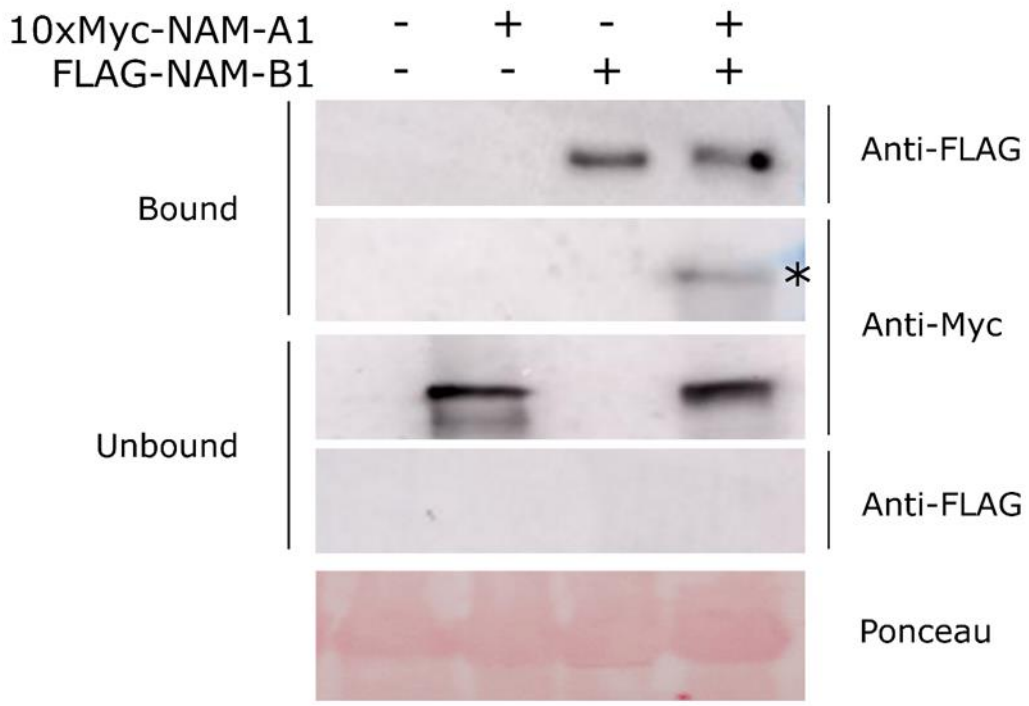

Figure 3-13: NAM-A1 interacts with NAM-B1 in a co-immunoprecipitation. 10xMyc-tagged NAM-A1 is pulled down alongside FLAG-tagged NAM-B1 using anti-FLAG magnetic beads. The asterisk indicates the co-immunoprecipitated 10xMyc-tagged NAM-A1. FLAG-tagged bands are approximately 46 kDa; 10xMyc-tagged bands are approximately 51 kDa. Ponceau stain is of total protein input (Rubisco). Note that this western blot is duplicated in Figure 5-6 in Chapter 5, as the same Co-IP was carried out for both experiments.

We cloned the full set of conserved missense mutant alleles and non-conserved and synonymous mutant alleles into *NAM-A1* using site-directed mutagenesis. We cloned both the TILLING-derived mutant alleles, that is the specific missense mutation present in the TILLING line, and also a series of alanine mutants, where the wild-type residue is mutated into an alanine. This method, sometimes referred to as alanine-scanning mutagenesis (Cunningham and Wells 1989), can be used to determine whether a residue is essential for protein function. Alanine residues have the smallest side chain of any amino acid, consisting of a single methyl group (CH_3). As a result, changing a residue to an alanine allows the elimination of the wild-type side chain while not inducing any changes in main-chain conformation. This method would therefore allow us to distinguish between any negative effects caused by the new allele in the TILLING lines compared to the inherent role of the original residue. E.g. the delay in senescence caused by the $K2711_{R110W}$ mutation may be due to steric changes caused by the sidechain associated with W (tryptophan; indole side chain) rather than because the R residue (arginine; guanidino group side chain) is required for protein function.

The mutant alleles were then cloned into vectors appropriate for use in a Y2H screen, pDEST22 and pDEST32. Having found previously in the lab that the C-terminal domain of *NAM-A1* leads to auto-activation of the reporter gene in the Y2H system, we only cloned the NAC domain of *NAM-*

A1 for use in this assay. Auto-activation in Y2H screens by the C-terminal domain of NAC transcription factors has previously been described and is a not uncommon problem when working with transcriptional activators. Previous work on Arabidopsis NAC transcription factors has found that removal of the C-terminal domain does not substantially alter protein dimerisation properties (Kang et al. 2018). With this knowledge, we proceeded to test the ability to the mutant alleles of the *NAM-A1* NAC domain to form heterodimers with the NAC domain of *NAM-B1*.

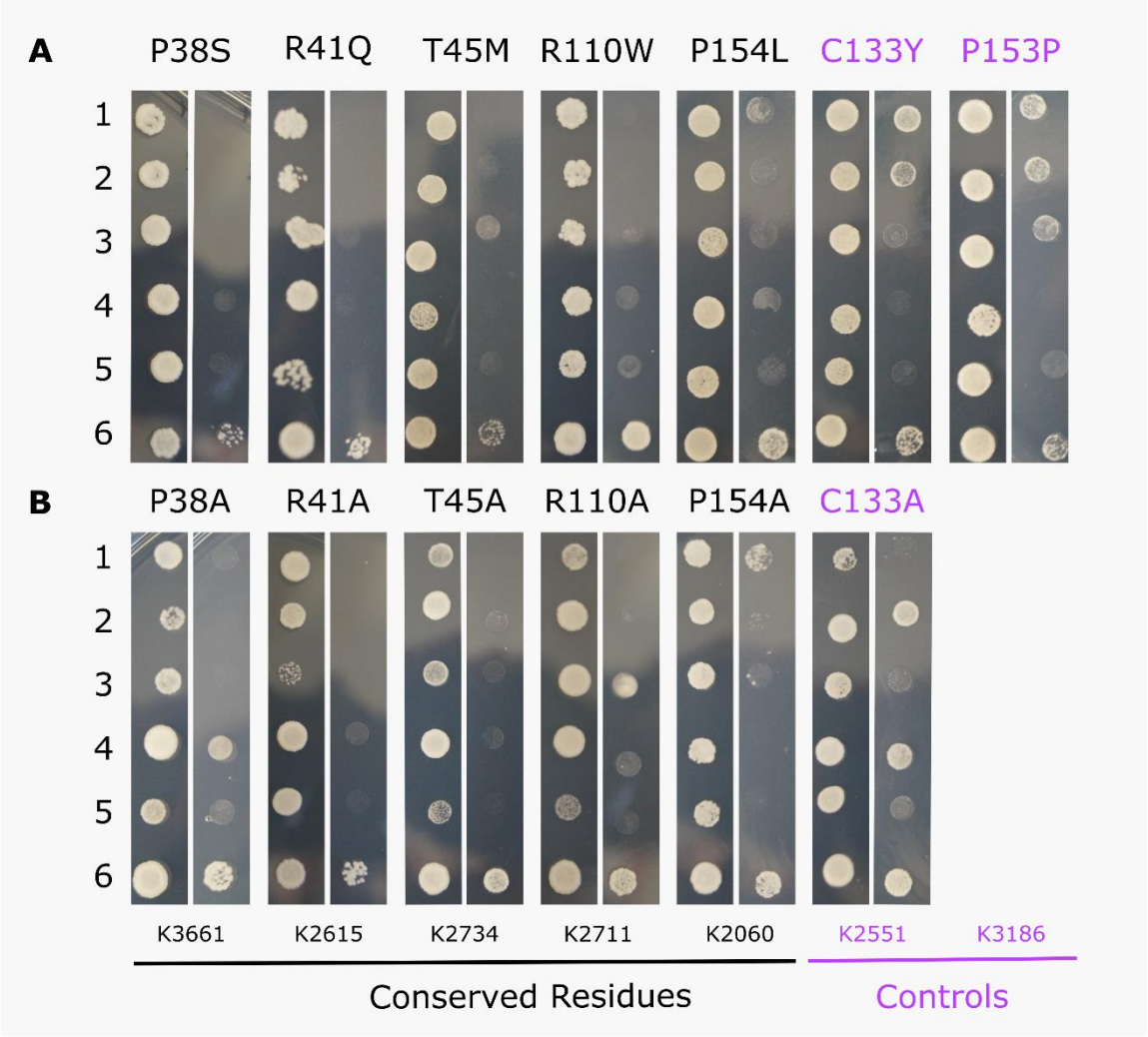

Figure 3-14: Mutations in conserved residues of NAM-A1 prevent protein-protein interactions. TILLING (A) and alanine (B) mutant alleles of NAM-A1 were tested on control media (left; SC-LT) and selective media (right; SC-LTH + 10mM 3AT); growth on the selective media is indicative of physical interactions between the expressed proteins. NAM-A1 alleles were co-transformed with the wild-type NAM-B1 allele in the complementary vector (1, NAM-A1 in pDEST22 and NAM-B1 in pDEST32; 2, NAM-A1 in pDEST32 and NAM-B1 in pDEST22). The ability of the mutant allele to interact with itself was also tested (3). Relevant controls were carried out for all alleles including pDEST32_NAM-A1 alone (4), pDEST22_NAM-A1 alone (5), and wild-type NAM-A1 in pDEST22 with wild-type NAM-B1 in pDEST32 (6). The corresponding TILLING line is given at the bottom of the figure.

Our initial screen focused on testing whether the transformed yeast were able to grow on selective media when the NAM-A1 allele and NAM-B1 were both present. If the yeast was able to grow it would indicate that the mutation does not prevent an interaction between the two proteins. In a

simple drop assay experiment, we found that the two control alleles retained the ability to bind NAM-B1 (Figure 3-14A). This was the case for the synonymous mutation (P153P) and for the missense mutation in a non-conserved residue (C133Y). In contrast, we found that the conserved residue missense mutations nearly completely abolished the ability of NAM-A1 to bind NAM-B1 (Figure 3-14A). Only P154A showed any sign of growth on the selective media, and this was at a very low level. In all cases, no or minimal growth was seen in the negative controls (4 and 5) and the positive control of NAM-$A1_{WT}$:pDEST22 x NAM-B1:pDEST32 showed consistent growth in all replicates (6).

We then screened the alanine mutant alleles and found very similar results (Figure 3-14B). Again, all conserved residue mutations led to a complete loss of NAM-A1/NAM-B1 interaction. The exception was again P154A, which showed a higher level of growth on the selective media than seen with the corresponding TILLING mutation. The only discrepancy was seen with the non-conserved residue control, C133A, where the only growth seen was also present in the corresponding control (2 and 4).

To investigate this further, we carried out a dilution assay on the same selective media which indicates the relative strengths of the protein interactions (Figure 3-15). Again, we found that the controls (purple) recapitulate the binding strength of wild-type NAM-A1. Of the conserved residue mutations (black), the P154L mutation shows a decrease in binding affinity, where the ability of the yeast to grow on the selective media is reduced compared to the controls. This decrease in affinity seems to be recovered completely in the corresponding alanine mutant which closely phenocopies the growth patterns seen in the controls. The remaining conserved residue mutations all show a complete loss of binding ability for the TILLING mutations. This loss is also seen in the corresponding alanine mutants except for P38S. This mutation leads to a reduction in binding affinity between the NAM-A1 and NAM-B1 proteins, but not a strong enough reduction to prevent any growth on the selective media at all.

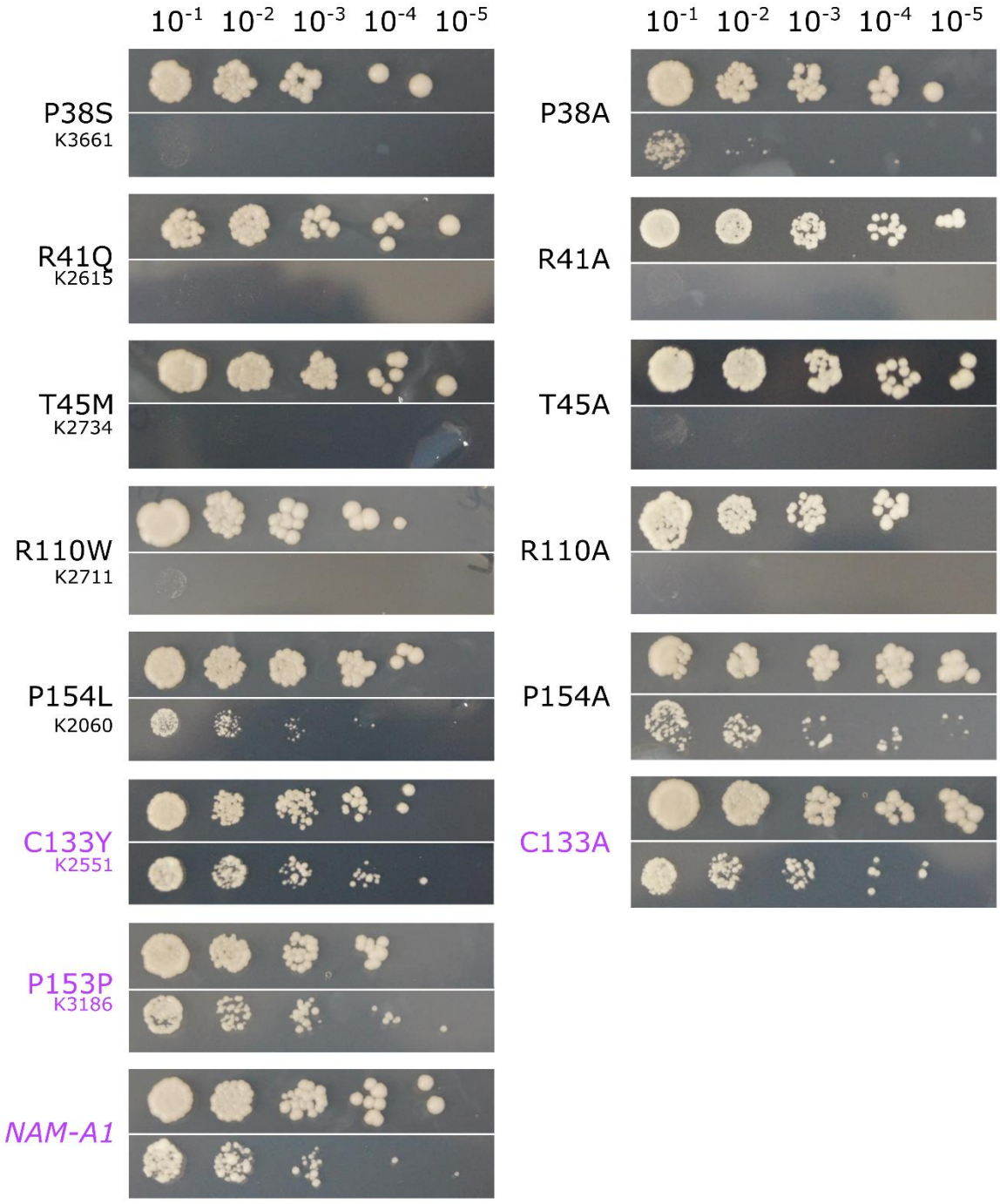

Figure 3-15: Mutations in the protein-binding domain of NAM-A1 inhibit protein binding. Protein-protein interactions between NAM-A1 alleles and wild-type NAM-B1 were tested on control (SC-LT; top row) and selective (SC-LTH + 10mM 3AT; bottom row) media at ten-fold OD600 dilutions, from 10^{-1} to 10^{-5}. Growth on selective media indicates interaction between the NAM-A1 allele and wild-type NAM-B1. No growth on selective media indicates no interaction. All controls showed equivalent growth on selective media as the wild-type NAM-A1 (purple). The left-hand panels correspond to the TILLING mutations, while the right-hand panels are alanine substitutions. The corresponding TILLING line is shown underneath the mutation on the left panel.

3.3.7 The missense mutations in NAM-A1 reduce the induction of cell death in N. benthamiana

Having shown that the conserved missense mutations may in part to prevent the action of *NAM-A1* by preventing or reducing its ability to form protein-protein interactions, we then investigated the effect that these mutations may have on the induction of cell death. Using the heterologous expression system in *N. benthamiana*, we were able to express the full-length sequences of the TILLING and alanine mutant alleles. We had found that expression of wild-type *NAM-A1* in *N. benthamiana* was able to induce a low, but consistent level of cell death. This was measured by visual signs of chlorosis and cellular auto-fluorescence, as previously described in Maqbool *et al*. 2015. We expressed the conserved residue mutations, and the corresponding controls, in *N. benthmiana* and then scored the resultant cell death phenotypes.

We found that three of the five conserved-region mutations were able to significantly reduce the visible cell death response (Figure 3-16). Chlorosis was significantly reduced in both the TILLING and alanine alleles of K2615, K2734, and K2711, all lines which had shown the strongest senescence phenotypes in the glasshouse (Figure 3-8). The TILLING mutation of K2060, P154L, led to a slight decrease in chlorosis while the alanine mutation did not, corresponding with the effect of the mutations on protein-protein interactions (Figure 3-15). Curiously, neither the TILLING nor the alanine allele of the K3661 mutation led to any decrease in cell death despite the mutations hindering protein binding and delaying senescence in the glasshouse. No significant differences in chlorosis levels were seen for the controls. A similar result was seen for the auto-fluorescence scores, where the majority of the conserved residue mutations led to a decrease in cell death (Figure 3-16B). Again, the principle discrepancy was the pair of K3661 alleles, which showed no decrease in cell death.

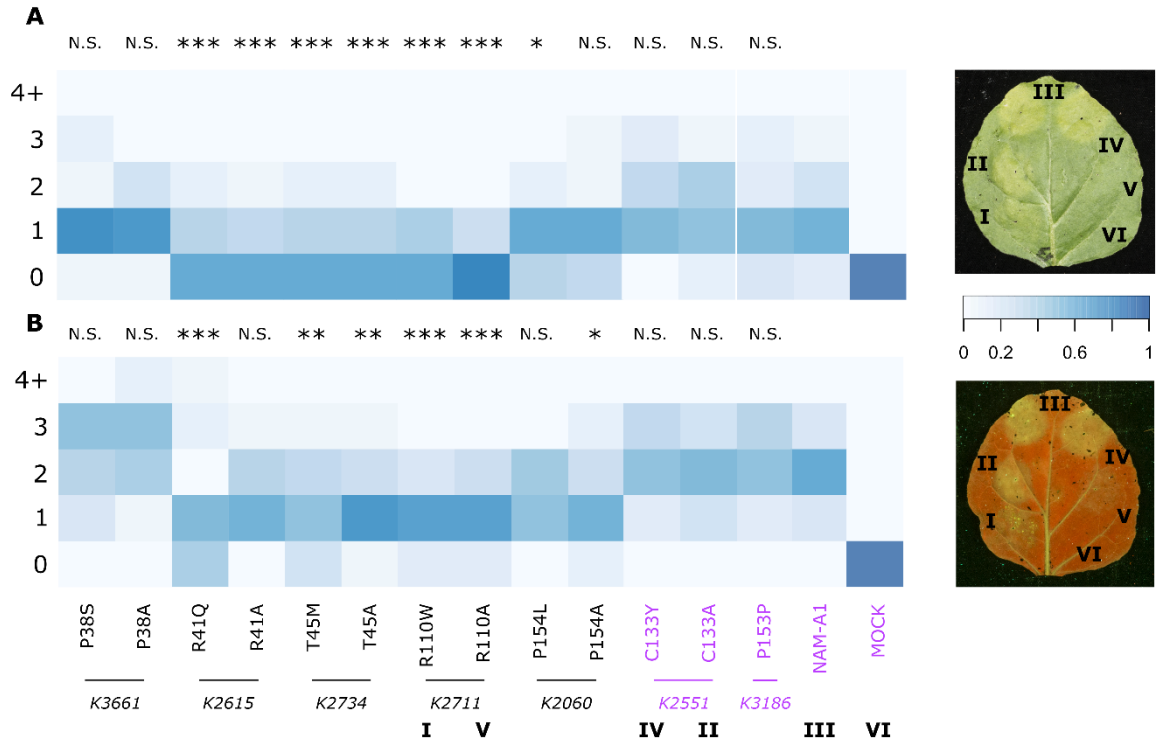

Figure 3-16: Mutant alleles of *NAM-A1* lead to a reduced cell death response. Heterologous expression of the mutant *NAM-A1* alleles in *N. benthamiana* shows a significant reduction in cell death response for both chlorosis (A) and cellular breakdown (B) compared to the wild-type allele. Cell death is scored on a scale derived from Maqbool *et al*. 2015, where 0 is no chlorosis or necrosis and 6 is a high level of chlorosis or necrosis. The heatmap depicts the proportion of replicates that were assigned a particular score, where a darker blue colour indicates a higher density of replicates having that score. Note that all scores greater than or equal to four were assigned into the same group ("4+") as few replicates were assigned scores greater than four; n varies from 20-35. All control alleles (purple) do not differ significantly from the wild-type. The corresponding TILLING line is shown below the TILLING/alanine mutant pairs. A typical image of a patch-infiltrated leaf is presented for both chlorosis scoring (top, adaxial) and UV scoring (bottom, abaxial); the Roman numeral labels correspond to those beneath the heatmap; M is mock treated. The UV image (bottom) has been flipped horizontally to allow comparison with the white-light image. Statistical comparisons shown are between the mutant allele and wild-type *NAM-A1* based on a linear model fit to the Poisson distribution. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

Given the lack of observed phenotype for the K3661 alleles, it was possible that the mutated proteins, containing a 10x-cMyc N-terminal tag, were not being expressed. To that end, we carried out a Western blot of the *N. benthamiana* leaves expressing each of the mutant alleles. We did observe highly variable levels of expression of the different alleles, with the R110W allele expressed the most strongly, while the C133Y allele was comparatively lowly expressed. However, these expression patterns had no correlation to the observed cell death phenotypes. Not only were the K3661 alleles (P38S and P38A) expressed at an intermediate level, but the proteins which were expressed the highest showed no cell death response, while those expressed at the lowest levels

showed the expected level of cell death response. This demonstrated that the variability in cell death was not due to variable levels of protein expression.

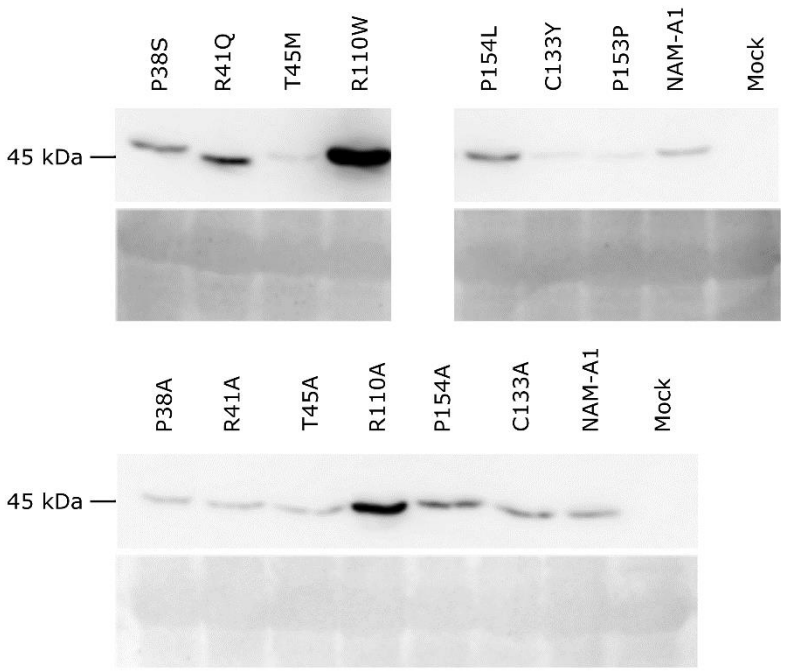

Figure 3-17: Expression levels of NAM-A1 mutant alleles does not correlate with observed hypersensitive response. All constructs showed expression of the FLAG-tagged proteins in *N. benthamiana* at 3 dpi. Expression levels were variable but did not correlate with observed cell death. The Ponceau stain of Rubisco levels (approximately 56 kDa) is shown below the Western blot.

3.4 Discussion

3.4.1 Highly conserved residues of the NAC domain are required for NAC transcription factor function across species

Previous work on the Arabidopsis NAC transcription factor ANAC019 had identified three residues that were thought to be required for protein-protein interactions and protein-DNA interactions, respectively. Two of these residues, corresponding to arginine 19 (R19) and glutamic acid 26 (E26) in ANAC019, were found to be essential for the formation of protein dimers (Olsen et al. 2005a). In this work, we have shown that mutating the corresponding arginine residue in the wheat NAC transcription factor *NAM-A1*, R41, also leads to a loss of protein dimerisation (Figure 3-15). Protein interactions were lost both when the arginine was mutated into the corresponding TILLING allele (glutamine) and into alanine, suggesting that the arginine residue itself is essential for the formation of NAC dimers. This finding also corresponds with predictions from Kang et al. (2018), which predicted that the R19 in ANAC019 would form a salt-bridge with the E26 residue present in the other monomer, stabilising the dimer formation (Figure 3-18A).

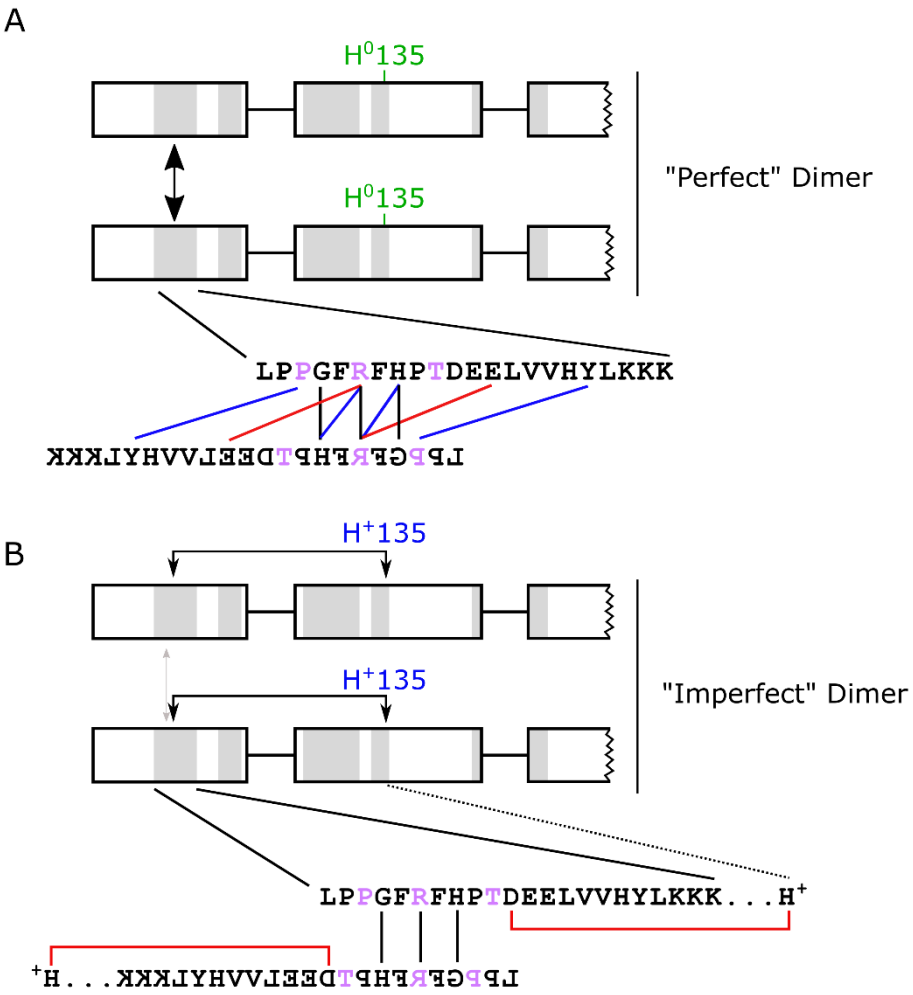

Figure 3-18: Deleterious mutations in NAM-A1 are located at key dimerisation residues. (A) The interactions formed in the "perfect" dimers state, as defined in Kang *et al*. 2018, include interactions between missense mutations tested in this thesis (purple) and other residues within subdomain i. (B) In the "imperfect" dimer state, many of these interactions are disrupted, and the T45 residue is located adjacent to the D residue which interacts with $H135^{+}$ from the same monomer. Interactions are colour coded as in Kang et al. (2018)—black, backbone-backbone; blue, backbone-side chain; red, salt bridge. The highlighted residues (purple) correspond to P38, R41, and T45 in NAM-A1. The orientation of the second monomer in the depiction of the inter-residue bonds is inverted to more closely match the known orientation of monomers in the NAC dimer.

We also tested the impact of mutating two other residues within the protein dimerisation interface, residues P38 and T45 in *NAM-A1*. Both of these residues are located in subdomain i and fall within the predicted dimerisation interface described in Kang et al. (2018). P38 itself acts to form stabilising hydrogen bonds between the two monomers, forming a backbone-sidechain interaction with a tyrosine residue in the other monomer (Figure 3-18A). T45 is located adjacent to D24 which forms a salt bridge with H135 in the protonated, imperfect dimer state (Figure 3-18B). This proximity suggests that mutations in T45 may lead to a slightly altered protein conformation that hinders the formation of these essential stabilising interactions across the dimerisation interface. In testing the effect of these mutations on protein dimerisation, we found that mutations of T45 (corresponding to T23 in ANAC019) lead to a complete loss of protein dimerisation in yeast

(Figure 3-15). Mutations in P38 (corresponding to P16 in ANAC019) also lead to disrupted protein dimerisation, though not to the extent seen with T45 or R41. While the TILLING allele, P38S, led to a complete loss of protein binding in a yeast two-hybrid assay, the corresponding alanine mutation showed only a decrease in binding efficiency (Figure 3-15). This suggests that while residues R41 and T45 are essential for protein dimerisation, P38 is not essential for dimer formation but may increase the stability of dimer formation through forming backbone-sidechain bonds.

Alongside the residues tested in subdomain i, we also tested the effect of mutations in residues predicted to affect protein-DNA binding, rather than protein-protein interactions. One of the residues selected, R110, was previously shown to be essential for DNA binding by ANAC019 (Olsen et al. 2005a, Welner et al. 2012). The second residue, P154, is located in the predicted DNA-binding region, but has not been tested for its effect on protein function. Based on their positions within the NAC domain, we hypothesized that mutating these residues would have no or minimal effect on protein dimerisation. This was found to be the case for P154, which had no substantial effect on protein binding in yeast when mutated to an alanine residue (Figure 3-15). However, we found that mutations in R110 led to a complete loss of binding, irrespective of whether the residue was converted to the TILLING allele, R110W, or to an alanine residue (Figure 3-15). This unexpected result suggests that the H135 residue in subdomain iii may not be the only residue in the putatively DNA-binding domain of the NAC transcription factor that plays a role in protein-protein interactions. It is possible that the R110 residue is required specifically for the formation of NAC protein dimers. Alternatively, it may be that this arginine residue is fundamentally important in the formation of overall protein structure. Our analysis of the conservation levels of each of the residues mutated found that the R110 residue was the most conserved of any, with 85% of the NAC transcription factors tested containing an arginine residue at this location (Figure 3-5). As further work is carried out on NAC transcription factors, increased analysis of the structural consequences of point mutations at conserved residues in the NAC domain would aid in distinguishing between the effect of specific mutations on protein structure as compared to on specific protein functions, such as DNA binding.

3.4.2 The Kronos TILLING population provides a tractable system for studying the impact of missense mutations on protein function

Beyond characterising the effects of mutations in the previously described residues on protein function *in vitro*, we were also able to test the impact of these mutations on protein function in wheat itself (*in planta*). Using the senescence regulator *NAM-A1*, we investigated the extent to which mutations in the specific residues recapitulated the highly delayed senescence phenotype characteristic of a loss of function mutant in *NAM-A1*. Of the five missense mutations tested, four led to severely delayed peduncle senescence phenotypes in the glasshouse, while the fifth, $K2060_{P154L}$, led to a slight delay that was only observed in one of the two glasshouse trials. This

outcome corresponds with the observed protein dimerisation phenotypes in yeast; the weakest mutation was that of P154L, which only led to a slight decrease in dimerisation ability, while the P154A mutation had a negligible impact. In contrast, all of the TILLING lines which showed consistent delayed senescence phenotypes also led to a complete loss of protein binding in yeast (Figure 3-15). We were thus able to demonstrate that not only were these specific residues essential for protein dimerisation, but also that the loss of this dimerisation ability leads to a clear mutant phenotype *in planta*.

Our use of the Kronos TILLING population in this manner was facilitated by the non-functional B-genome homoeolog of *NAM-A1*. As a result, single mutations in the A-genome homoeolog were sufficient to produce the complete mutant phenotype. However, in many cases mutations in a single homoeolog will be masked by the redundant, compensatory effect of the remaining functional homoeolog (see Figure 2-1, p.38). Nevertheless, the Kronos TILLING population still has the potential to allow similar studies to be carried out on many other genes. Approximately 60% of the high-confidence gene models with at least one mutation in the Kronos TILLING lines contain at least one premature truncation mutant (Krasileva et al. 2017). It is possible, therefore, that a premature truncation mutant could be identified in one homoeolog of the gene pair which is expected to completely disrupt protein function. A single cross could subsequently be made between this mutant and a missense mutant in the other homoeolog. This would artificially generate the same situation as found for *NAM-A1*: one non-functional homoeolog and a series of missense mutations in the remaining homoeolog (in the case of tetraploid wheat). These crosses could be repeated with multiple different missense alleles, creating a panel of lines ready for phenotyping (Figure 3-19). It is also worth noting that, following the sequencing of the Kronos genome, it will be possible to determine whether both homoeologs are expected to be functional. It may be the case that one or more of the homoeologs in Kronos contains a deleterious mutation that is not present in the reference cultivar, resulting in a similar situation to that of *NAM-A1*.

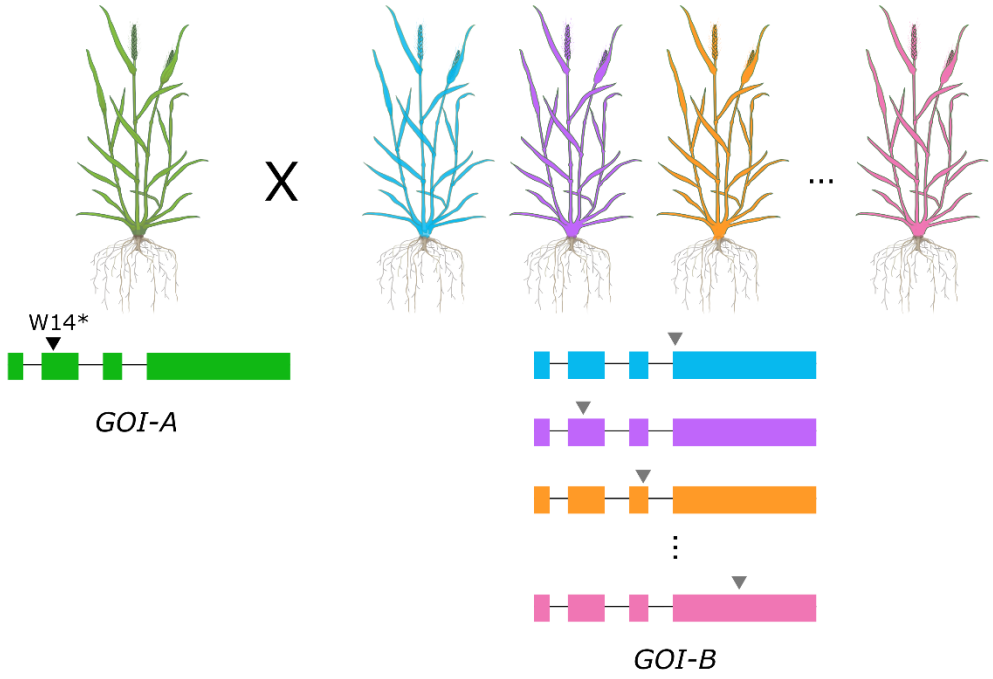

Figure 3-19: Crossing scheme to generate single missense mutations. The tetraploid Kronos TILLING population can be mined for premature stop codon variants (e.g. W14*, black) in one of the two homoeologs (here shown in green for *GOI-A*). This line can then be crossed to a panel of TILLING lines that contain missense mutations (grey) in desired residues of the other homoeolog (here *GOI-B*). These crosses would generate an artificial single-copy system for the gene of interest, consisting of one non-functional homoeolog and various missense mutations in the remaining homoeolog.

Having obtained an allelic series of mutations, it would then be possible to rapidly screen the progeny of these crosses for variable phenotypes caused by the different missense mutations in one homoeolog. However, while initial screening can often be carried out in the initial generations, one limitation when using any chemically-mutagenised population is the presence of unwanted background mutations. This difficulty is heightened in the wheat TILLING population due to the higher load of mutations that the polyploid wheat genome can accommodate compared to diploid species (Wang et al. 2012). As a result, any further analysis of the mutant lines should be accompanied by a series of backcrosses to reduce the background mutation burden. In our case, we screened the full set of TILLING mutants immediately in the M_5 generation, before any backcross had taken place. As the *NAM-A1* mutant phenotype of delayed senescence was highly visible and consistent across multiple independent TILLING lines, we were able to accurately score these phenotypes irrespective of the confounding effects of background mutations. Nevertheless, upon bringing a subset of the TILLING lines into field trials, we carried out an initial backcross to reduce the burden of background mutations by approximately 50%.

In future work, it could be highly beneficial to take the remaining TILLING mutants forward for further field trials to investigate the possibility of developing allelic series of agronomically-relevant genes that lead to a gradient of desired phenotypes. Within the set of TILLING mutations tested in the glasshouse, we observed delays in peduncle senescence ranging from only two or three

days to over 30 days. For those lines taken into the field, K2711$_{R110W}$ and K1107$_{SAV}$, we observed clear variation in the level of flag leaf senescence delay and, to a smaller extent, grain protein content reductions. Previous studies have highlighted the impact of allelic variation within populations on environmentally-dependent success. This is often seen in the case of flowering time variation, with allelic variation in crucial genes such as *PHOTOPERIOD1* (*PPD1*) closely associated with variable flowering time observed across latitudes (Worland 1996, Grogan et al. 2016, Whittal et al. 2018). This affect has already been shown in *NAM-A1* itself, where a natural missense variant of *NAM-A1* was associated with grain protein levels which fell between the wild-type allele and a third allele containing both the missense mutation and a frame-shift mutation in the C-terminal domain (Cormier et al. 2015). Interestingly, this missense mutation was located in a relatively non-conserved residue outside of subdomain iii, suggesting that potentially valuable variation can be mined from residues that cause a slight alteration in protein function, rather than complete abolition of protein function. In the future, this sort of insight could be combined with base-editing technology (Zong et al. 2017) to introduce highly discrete, specific variation into agronomically-important genes.

3.4.3 Mutating NAM-A1 leads to a stronger effect on peduncle senescence than on flag leaf senescence

While investigating the effects of the missense mutations of *NAM-A1* on wheat senescence, we noticed that all of the mutations had a stronger effect on peduncle senescence than on flag leaf senescence. In the glasshouse, four of the five missense mutations consistently led to delayed peduncle senescence in both experiments (Figure 3-7B, Figure 3-8B). In contrast, a delay in flag leaf senescence was only observed in one of the two glasshouse experiments (Figure 3-7A). This pattern was also observed in the field trials, where only K2711$_{R110W}$ consistently showed both delayed flag leaf and peduncle senescence, while K1107$_{SAV}$ only exhibited delayed flag leaf senescence in the final year (2018). These discrepancies suggest that *NAM-A1* is more important in regulating the onset of peduncle senescence than flag leaf senescence.

One caveat to this hypothesis is that peduncle and flag leaf senescence are measured based on different visual characteristics. Flag leaf senescence is typically measured in terms of senescence onset, scored when 25% of the flag leaf has turned yellow, typically progressing from the leaf tip to the base of the leaf (Figure 3-20A). In contrast, peduncle senescence is scored visually when the top 3 cm of the peduncle, directly below the spike, is completely yellow (as shown in Figure 3-20B). This is a measure more of final senescence, rather than onset. The reason the visual scoring techniques differ between the two phenotypes is a result of the different progressions of flag leaf and peduncle senescence. While flag leaf senescence occurs from the tip, making it easy to see when approximately 25% of the leaf has senesced, peduncle senescence typically occurs around the stem simultaneously. This pattern of senescence makes it difficult to determine the exact timing and location of peduncle senescence onset. We can account for the discrepancy in the visual scores

in part using SPAD measurements of flag leaf greenness throughout senescence. Comparing these measurements, we can see that although four of the five missense mutations exhibited clear delays in peduncle senescence in the glasshouse (Figure 3-8B), no corresponding significant delay in senescence was observed in any of the SPAD measurements taken (Figure 3-9). A similar pattern was observed for the field trials, where the strong delays in peduncle senescence in 2018 corresponded with slight, borderline significant delays in flag leaf senescence. Taken together, these results indicate that while *NAM-A1* does influence flag leaf senescence, it has a stronger effect on peduncle senescence.

Figure 3-20: Progression of leaf and peduncle senescence. (A) Stages of leaf senescence, from pre-senescent (far left) to fully senescent (far right). Flag leaves in the centre are at the approximately 25% senesced stage. (B) Stages of peduncle senescence, from pre-senescent (left) to fully senesced (right). Only the peduncle on the far right would be scored as senesced, while the central peduncle would still be considered non-senesced.

The particular role of *NAM-A1* in peduncle senescence has historically been under-studied compared to its role in flag leaf senescence. Previous studies have noted a delay in peduncle senescence, though this has typically been as an addition to a more in-depth study of flag leaf senescence (Uauy et al. 2006b, Avni et al. 2014, Pearce et al. 2014, Borrill et al. 2019a). However, given the known role of *NAM-A1* in nutrient remobilisation during senescence, it is perhaps not surprising that it would have an important role in peduncle senescence. During remobilisation, nutrients from the flag leaf must pass through the peduncle in order to reach the developing grain (see Figure 1-10A, p. 12). Not only have knock-downs of the *NAM* genes led to reduced micronutrient and protein content in the grains (Waters et al. 2009) but knock-out mutations of *NAM-A1* and *NAM-D1* in hexaploid wheat lead to higher nitrogen content in the peduncle at harvest (Avni et al. 2014). Initial analyses of the expression of the *NAM-1* genes indicated that they

were not expressed in the peduncle at 14 DAA, suggesting that their function lay primarily in the flag leaf, and that any observed peduncle phenotype was incidental to their principle function (Guttieri et al. 2013). However, recent RNA-Seq datasets have found high expression of *NAM-A1* and its homoeologs and paralogs in the peduncle and other stem tissues, as well as in the senescing flag leaf during grain filling (Borrill et al. 2016, Ramírez-González et al. 2018)(Table 3-17). By 15 DAA, the expression levels of these genes are higher in the peduncle than they are in the flag leaf. While this RNA-Seq dataset finishes at 15 DAA, a separate experiment found that *NAM* expression levels were higher in the peduncle than the grain from 7 DAA until 30 DAA, corroborating this finding (Borrill 2014).

Based on these results, it's tempting to speculate that the effect of mutating *NAM-A1* is to delay peduncle senescence, leading to sequestration of nutrients in the peduncle and thus hindering their remobilisation into the development grain. This would explain the reduced grain protein and micronutrient content in *nam-a1* mutants shown here and previously (Uauy et al. 2006a, Uauy et al. 2006b, Avni et al. 2014, Pearce et al. 2014). It also corresponds with earlier analyses that identified high levels of retained iron, zinc, nitrogen, and fructans in the peduncle of *nam-a1* mutants during senescence (Waters et al. 2009, Avni et al. 2014, Borrill et al. 2015b). While historically most focus has been on senescence of the flag leaf, these results suggest that, in the future, increased study of peduncle senescence may yield dividends in terms of our understanding of the senescence and remobilisation process. The development of a quantitative phenotyping framework for peduncle senescence would also aid in this goal. Currently, any quantitative scoring of chlorophyll content in the peduncle relies upon destructive sampling of the peduncle to chemically extract and quantify chlorophyll levels. The comparative ease of measuring leaf greenness using a SPAD meter facilitates the continued use of the flag leaf as the main indicator of monocarpic senescence in wheat. Despite these difficulties, our analysis here suggests that there is substantial insight to be gained by studying the specific regulating of peduncle senescence in wheat.

3.5 Summary

In this chapter, we have combined *in silico* predictions of critical residues in the NAC domain with biochemical and *in planta* studies of their role in normal NAC transcription factor function. We have focussed on the role of these residues in preventing the formation of protein dimers. We found a strong correlation between mutant phenotypes in the TILLING population, that is delayed (peduncle) senescence, with loss of protein dimerisation ability. We also utilised the heterologous expression system in *N. benthamiana* to quantify the impact of these missense mutations on the induction of cell death. Our characterisation of these mutations confirms previous work that was carried out in Arabidopsis and identifies residues critical for protein function which had not previously been identified. The method we present can be used for further characterisation of protein function in the Kronos TILLING population, enabling the use of wheat as a model species for testing the impact of missense mutation on protein function *in planta*.

Table 3-17: NAM genes are expressed in the peduncle and stem during reproduction. Expression data derived from the Wheat Expression Browser (Borrill et al. 2016) was obtained for the five functional NAM copies (A1, A2, B2, D1, and D2) from the developmental time course of Azhurnaya (Ramírez-González et al. 2018). Expression data is given in TPM ± SEM.

		NAM-A1	*NAM-D1*	*NAM-A2*	*NAM-B2*	*NAM-D2*
Timepoint	**Tissue**	TraesCS6A02G108300	TraesCS6D02G096300	TraesCS2A02G201800	TraesCS2B02G228900	TraesCS2D02G214100
Ear emergence	Internode #2	1.47 ± 0.66	1.5 ± 0.92	1.26 ± 0.48	1.43 ± 0.94	1.84 ± 0.9
	Peduncle	0.22 ± 0.06	0.49 ± 0.32	0.15 ± 0.17	0	0.01 ± 0.02
	Flag Leaf	0.46 ± 0.05	0.13 ± 0.07	0.11 ± 0.11	1.19 ± 0.58	1.38 ± 0.05
30% spike	Internode #2	0.61 ± 0.37	0.55 ± 0.41	0.28 ± 0.16	0.25 ± 0.19	0.87 ± 0.85
	Peduncle	0.24 ± 0.06	0	0.03 ± 0.03	0.01 ± 0.01	0
	Flag Leaf	0.52 ± 0.07	0.06 ± 0.01	0.1 ± 0.12	0.37 ± 0.28	0.59 ± 0.28
Milk grain	Internode #2	2.57 ± 0.08	3.05 ± 0.15	5.18 ± 0.96	5.42 ± 0.42	6.16 ± 0.74
	Peduncle	3.56 ± 0.52	10.01 ± 1.5	11.51 ± 1.66	9.79 ± 1.26	5.22 ± 1.52
	Flag Leaf	1.25 ± 0.46	0.3 ± 0.21	0.29 ± 0.37	4.17 ± 1.99	4.12 ± 1.26
Dough	Flag Leaf	2.92 ± 0.45	1.71 ± 0.67	1.55 ± 0.65	18.79 ± 7.47	13.61 ± 3.98
Ripening	Flag Leaf	9.59 ± 0.93	10.03 ± 1.65	8.04 ± 0.44	48.69 ± 10.66	30.25 ± 6.88

4 The NAC transcription factor *NAM-B1* is necessary but not sufficient to induce senescence.

Part of the work in this chapter (RNA extractions, qPCR, plant phenotyping) was carried out by Anna E. Backhaus, a PhD student at the John Innes Centre, during her rotation in the Uauy lab under the supervision of the author.

4.1 Introduction

The wheat NAC transcription factor *NAM-B1* is a positive regulator of senescence (Uauy et al. 2006b). In the previous chapter, we showed that the A-genome homoeolog, *NAM-A1*, is also a strong positive regulator of monocarpic senescence, particularly in the peduncle (Harrington et al. 2019c). This work corroborated earlier work which also found a delay in senescence in premature stop codon mutants of *NAM-A1* (Avni et al. 2014, Pearce et al. 2014). We also found, in the course of this study, that ectopic expression of *NAM-A1* in *Nicotiana benthamiana* leads to a low-level but consistent induction of cell death which is not seen when non-functional *NAM-A1* alleles are expressed. This raises the following question—would premature expression of *NAM-A1* or *NAM-B1* lead to premature senescence?

4.1.1 Ectopic expression of senescence regulators can lead to premature senescence

Work in other species has identified multiple positive regulators of senescence that, when over-expressed, lead to premature senescence (Guo and Gan 2006, Balazadeh et al. 2011, Li et al. 2018). One such example is the Arabidopsis mitogen-activated kinase *MKK9*. When overexpressed under the control of the constitutive 35s promoter the plants exhibited severe early senescence phenotypes, coupled with plant dwarfing and sterility, the severity of which correlated strongly with the expression levels of *MKK9* (Zhou et al. 2009). A high proportion of the positive transgenic lines (71/184) died within 30 days post germination, suggesting that the overexpression of *MKK9* was sufficient to induce premature senescence at very early developmental stages. This phenotype was confirmed using a glucocorticoid-inducible transgenic system, where premature senescence was observed in the transgenic lines within four days following dexamethasone (DEX, a synthetic glucocorticoid) treatment. This work not only indicated that *MKK9* was a positive regulator of senescence, but crucially also showed that overexpression of this gene was sufficient to induce premature senescence irrespective of the developmental stage of the plant.

In the case of *NAM-B1*, we know from previous work that its expression is necessary to induce timely monocarpic (whole-plant) senescence (Uauy et al. 2006b). Whether *NAM-B1* and its homoeologs are sufficient to induce premature senescence, however, is not clear. Earlier attempts to investigate this question utilised a constitutive expression transgenic system. However, when this was first attempted in 2003 no positive transgenic lines could be identified (Uauy, personal

communication). At the time, it was thought that this was due to overexpression of *NAM-B1* being embryo-lethal due to its promotion of senescence and cell death. However, a possible confounding factor in this analysis was the relatively low efficiencies of wheat transformation at the time. In 2003, the most up-to-date wheat transformation protocols would achieve transformation efficiencies of approximately 4.4% (Hu et al. 2003). By 2012, reviews of wheat transformation protocols stated that efficiencies of up to 30% could be achieved (Risacher et al. 2009, Harwood 2011), though in practice efficiencies of closer to 5% for *Agrobacterium*-mediated transformation are typically achieved (Adamski et al. 2018). It was possible, therefore, that the failure to recover positive transformants from the initial transgenic construct was not due to the overexpression of *NAM-B1* being embryo-lethal, but rather due to the low efficiencies of wheat transformation.

4.1.2 Inducible systems of gene expression

In order to revisit this question, we decided to develop a transgenic system for inducible expression of *NAM-B1*. This would allow us to recover positive transformants irrespective of whether *NAM-B1* was truly embryo-lethal, and then to test the effect of *NAM-B1* over-expression in the T_1 generation. Various inducible systems for gene expression have been developed, ranging from chemically-inducible methods, such as the glucocorticoid-inducible system described for *MKK9*, to environmentally-induced methods, such as heat-shock (Table 4-1). The majority of inducible promoters tested in monocots, however, have been environmentally-induced. A barley (*Hordeum vulgare*) heat shock promoter, *HvHSP17*, was previously shown to function well in wheat, driving the heat-shock induced expression of the reporter gene GUS (Freeman et al. 2011). Promoters for two genes, *OsWRKY71* and *TdCor39*, were used to drive the cold-inducible expression of the wheat *DREB3* gene in barley and rice (Kovalchuk et al. 2013). A wheat NAC transcription factor, *NAC69*, was successfully induced by drought stress under the control of a barley drought-responsive promoter, *HvDhn4s*, in transgenic wheat (Xue et al. 2011). These examples suggested to us that a stress-induced promoter would be the best option for studying the effect of *NAM-B1* expression in wheat.

Table 4-1: Inducible systems used in transgenic systems in plants. “GOI” is gene of interest, “TF” is transcription factor, “ER” is estrogen receptor.

System	Description	Benefits	Negatives	References
Heat-shock inducible	Heat-shock inducible promoter (i.e. *AtHSP18.2*) coupled to GOI	Rapid induction of gene (< 2 hours)	Can be detrimental to plant development (e.g. can cause sterility at anthesis)	Freeman et al. (2011) Shinmyo et al. (1998)
Cold inducible	Cold inducible promoter (i.e. *OsWRKY17*) coupled to GOI	Useful to test role of GOI in frost tolerance; avoids negative effects of heat stress.	Frost application can lead to seedling death	Kovalchuk et al. (2013)
Drought inducible	Drought-inducible promoter (i.e. *HvDhn4s*) coupled to GOI	Useful to test role of GOI in drought tolerance	Drought stress can lead to confounding phenotypes.	Xue et al. (2011)
Multi-stress inducible	Promoters like *Rd29A*, which are induced by multiple stresses such as cold, salt, and desiccation, coupled to the GOI	The *Rd29A* promoter is induced in nearly all tissues in response to a variety of stresses	Induction of the promoter by multiple different stresses can make it difficult to control	Yamaguchi-Shinozaki and Shinozaki (1993)
β-estradiol inducible-XVE:OlexA	The lexA repressor domain is fused to the VP16 activation domain and the human ER	No toxic effects have been seen in transgenic Arabidopsis, and has high levels of expression following induction without background expression	Phytoestrogens are not able to induce the system in Arabidopsis. High levels of background expression in plants with high estrogen levels (e.g. soy bean)	Zuo et al. (2000)
β-estradiol inducible-ER-C1	Fusion of human ER with maize C1 activator domain	No background expression of gene.		Bruce et al. (2000)
Ethanol inducible	The AlcR transcription factor is placed under an ethanol responsive promoter, binds the AlcA promoter upstream of GOI	Biodegradable inducer possibly useful for field applications	Highly volatile inducer; relies upon an ethanol responsive promoter which may be induced prematurely if biological ethanol levels increase	Caddick et al. (1998) and Roslan et al. (2001)

System	Description	Benefits	Negatives	References
Glucocorticoid inducible	A TF of interest is fused to the glucocorticoid receptor (GR) and sequestered in the cytosol; upon application of DEX the TF:GR fusion is released and translocates to nucleus	Only one construct required	Only applicable for studying transcription factors	Initially in yeast Schena et al. (1991); then shown in plants Kirch et al. (2003), Jasinski et al. (2005)
Glucocorticoid inducible-GVG/UAS	Fusion of GAL4 DNA-binding domain, the transactivating domain of VP16, and GR. Upon DEX application, translocates to nucleus.	Works for any gene (not just transcription factors)	Excessive GVG expression can lead to plant death and growth defects	Aoyama and Chua (1997)
Glucocorticoid inducible-pOp/LhGR	*lac* repressor fused to GAL4 activation domain and GR; DEX application allows translocation into nucleus.	No known side effects due to expression of the LhGR fusion protein.	DEX application can lead to the induction of defense-related genes, possibly confounding phenotypic analysis	Moore et al. (1998)
RH5992 (tebufenozide) inducible	Fusion of the GR and VP16 activating domain, the DNA-binding domain of GR, and the hormone regulatory domain of an ecdysone receptor.	Very responsive system that is based on a standard agrichemical which is not considered toxic in field conditions.	Reasonable background expression of the construct is observed before inducer application.	Martinez et al. (1999)
Methoxyfenozide inducible	Fusion of the ligand binding domain of the ecdysone receptor, the GAL4 and LexA DNA binding domains, and the VP16 activation domain	Non-toxic	Relatively slow induction (six hours for mRNA response)	Padidam et al. (2003)
Copper inducible	The Ace1 TF changes conformation upon perception of Cu^{2+}, binding the Ace1 recognition sequence	Copper is readily assimilated by the plant and is applied at non-toxic levels to induce expression	Copper can be phytotoxic at high concentrations, and as a natural molecule could lead to premature expression of the construct	Mett et al. (1993)

Previous studies of genes under the control of heat-shock inducible promoters, however, had found that mRNA levels drop rapidly within one hour after the end of the heat shock, with no mRNA detectable 16 hours after the end of heat shock (DeRocher et al. 1991). Our work studying the *NAM* genes has shown that they are expressed over a long timeframe during senescence, at least 10-15 days (Borrill et al. 2019a). It would be possible that expression of *NAM-B1* for only a few hours following stress induction would not be sufficient for any premature senescence phenotype to be exhibited. As such, we needed to develop a transgenic system that allowed not only induction of gene expression following application of a specific stress, but following this induction maintained the expression of *NAM-B1* permanently.

4.1.3 The Cre-loxP system for constitutive gene expression

To address this, we used the Cre-*lox*P system, initially developed for use in mice (Sauer 1998) but now regularly used in many species including plants (Vergunst et al. 2000). Cre recombinase is an enzyme derived from the bacteriophage P1 which recombines DNA between two specific DNA sequences, named the *lox*P sites (Sternberg and Hamilton 1981). Using this system, we could induce expression of the Cre recombinase following application of an environmental stress (Figure 4-1). Following this, the Cre recombinase would cut at the two *lox*P sites which flank a reporter gene, excising this gene and bringing the gene of interest, in this case *NAM-B1*, under the control of the constitutive promoter. This system would provide us with the ability to specifically interrogate the ability of *NAM-B1* to induce premature senescence at different developmental stages, ranging from seedling to reproductive stages.

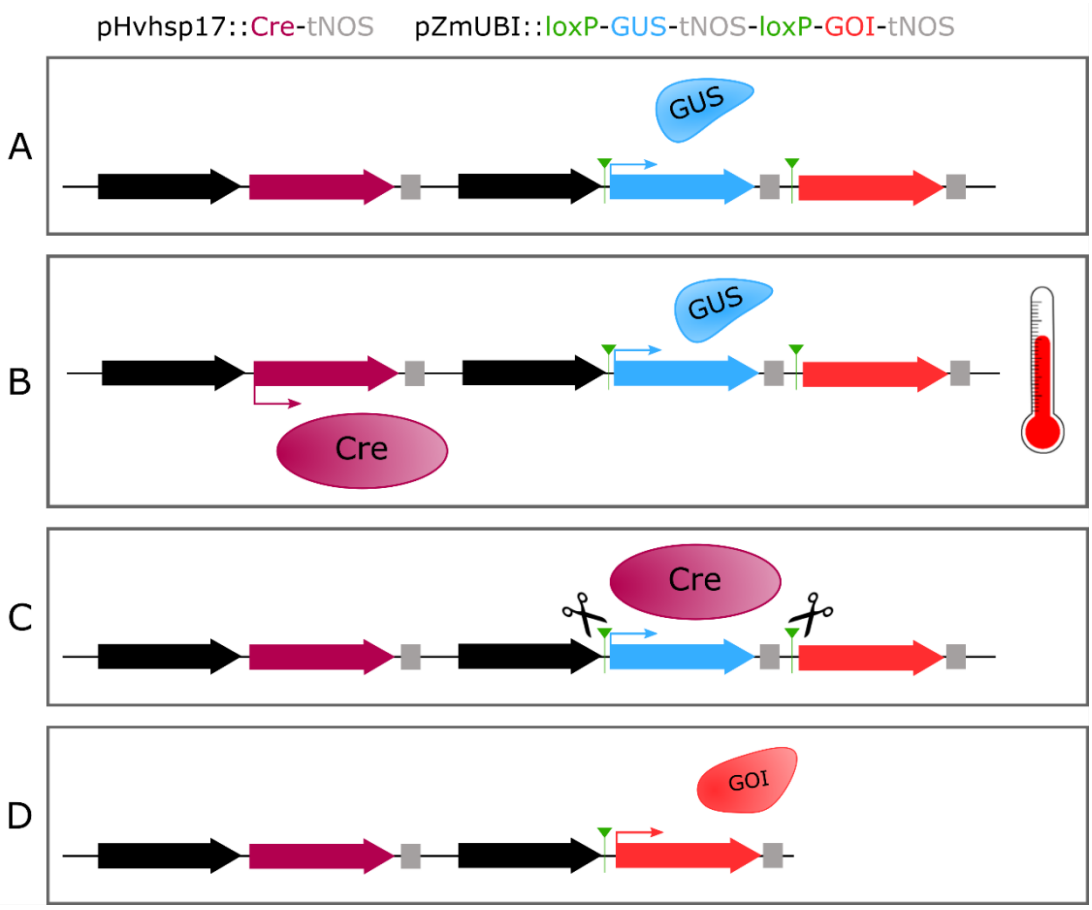

Figure 4-1: Schematic of the inducible Cre-*lox*P system. (A) The initial construct contains Cre recombinase, under the control of a stress inducible promoter, such as Hvhsp17, and a constitutive expression promoter, here the maize Ubiquitin promoter, driving expression of a reporter gene, in this case GUS (blue). The reporter gene is flanked by two *lox*P recombination sites (green) and is followed by a second gene of interest (GOI; red) which is not expressed in the initial construct. (B) Following application of stress, in this case a heat shock treatment, expression of the Cre recombinase gene is induced. (C) The Cre recombinase protein carries out recombination at the loxP sites flanking the GUS reporter gene. (D) Following excision of the reporter gene, the gene of interest is now in frame with the promoter sequence, and expression of the gene of interest is now under the control of the constitutive promoter.

4.1.4 Aims and hypotheses

In this chapter we hypothesized that we could use an inducible system to study the effect of ectopic *NAM-B1* expression on plant development and senescence onset. We transformed hexaploid wheat cv. Bobwhite with a construct that would constitutively express *NAM-B1* following application of a heat shock. We characterised the function of this construct and demonstrated that expression of *NAM-B1* was induced following application of a heat shock treatment. Preliminary phenotypic characterisation of the transgenic lines expressing *NAM-B1* found no evidence of premature senescence. We also developed and characterised a standard constitutive over-expression line of *NAM-B1*. Initial phenotyping also suggested that constitutive over-expression of *NAM-B1* does not induce premature senescence, though we did observe unexpected delays in peduncle senescence. Overall, we demonstrated that, while *NAM-B1* is necessary for monocarpic senescence, it is not sufficient.

4.2 Materials and Methods

4.2.1 Construct design

The functional *NAM-B1* coding sequence (from *Triticum turgidum*) was obtained from NCBI (GenBank DQ869673.1; Appendix 8.1.1) and 'domesticated' to remove the single *BbsI* site present in exon 3 (Appendix 8.1.2). Introns were approximated based on the gene sequence of the functional homoeolog *NAM-A1* in the Chinese Spring reference sequence (TGACv1) (Clavijo et al. 2017b). The domesticated gene, with introns, was synthesized by Genewiz. Golden-gate cloning (Engler et al. 2008) was used to develop the construct. The *NAM-B1* sequence, and the sequence of the ZmUbi+5'UTR (from pICSL12009; see all construct details in Table 4-2) were cloned directly into pUAP1 to generate universal level 0 constructs with the desired 4-base pair overhangs, hereafter pNAM-B1 and pZmUbi, respectively (See primers used in Table 4-3).

Table 4-2: Constructs used from outside sources.

Name	**Description**	**Origin**
pICH41421	NosT	The Sainsbury Lab (TSL) Synbio[1]
pICSL12009	ZmUbi promoter + ubiquitin 5' UTR (untranslated 1st exon + intron)	TSL Synbio[1]
EC15320	Heat-shock promoter terminator	ENSA
EC10161	LoxP Vector	ENSA
EC71173	Level 1 P2 HspHv17::Cre-U5-Cre	Sam Fox/Annis Richardson (Rico Coen)
pAGM8031	Level M Position 1	TSL Synbio[1]
pICH75111	GUS (β-glucuronidase gene) with 2 introns	TSL Synbio[1]
pICH47751	Level 1 Position 3	TSL Synbio[1]
pICH47802	OsAct promoter + Hygromycin coding sequence, Level 1 P1	TSL Synbio[1] and BRACT[2] (Rey et al. 2018)

[1] http://synbio.tsl.ac.uk/

[2] https://www.jic.ac.uk/research-impact/technology-platforms/genomic-services/crop-transformation/

Table 4-3**:** Primers used for generation of Level 0 constructs.

Primer	**Sequence**
NAMB1_Level0_Forward_AATG	ACGAAGACATCTCAAATGGGCAGCTCCGACTCATC
NAMB1_Level0_Reverse_GCTT	ACGAAGACATCTCGAAGCTCAGGGATTCCAGTTCACGC
MzUbi_Level0_Forward_GGAG	ATGAAGACATCTCAGGAGGTGCAGCGTGACCCGGTCGT
MzUbi_Level0_Reverse_TACT	ACGAAGACATCTCGAGTACCTGCAGAAGTAACACCAAA

To create the *lox*P-flanked GUS construct, the short *BsaI* protocol was carried out (synbio.tsl.ac.uk), using the *lox*P cassette developed by ENSA (EC10161), the GUS coding sequence from TSL Synbio (pICH75111), and the nosT terminator sequence (pICH41421). This created a "Level 0.5" vector, hereafter p*lox*P-GUS, with *lox*P sites flanking Gus-nosT, prepared to insert into a Level 1 vector (see Results Figure 4-2B).

Table 4-4: Reaction components for short *Bsa*I protocol.

Component	**1x (µL)**
Acceptor plasmid (100 ng/µL)	1
Donor plasmid (200 ng/µL)	1
*Bsa*I-HF (20 U/µL; NEB cat. no R3733)	1
CutSmart Buffer (10X; NEB cat. no. B7204S)	2
T4 DNA Ligase (400 U/µL; NEB cat. no. M0202)	1
ATP (10 mM)	2
dH_2O	2

Table 4-5: Reaction conditions for short *Bsa*I protocol.

Step	**Temperature**	**Duration**
Digestion-Ligation Reaction (x3)	37°C	10 min
	16°C	10 min
Final Digestion	37°C	10 min
Enzyme inactivation	65°C	20 min
Hold	16°C	-

The promoter construct (pZmUbi), the *lox*P-GUS construct (p*lox*P-GUS), the *NAM-B1* construct (pNAM-B1), and the nosT terminator (pICH41421) were then cloned into the Level 1 backbone, position 3 (pICH47751) using the long *BsaI* protocol. This construct is referred to as pNAMB1-LOX

Table 4-6: Reaction components for long *Bsa*I protocol.

Component	**1x (µL)**
Acceptor plasmid (100 ng/µL)	1
Donor plasmid (200 ng/µL)	1
*Bsa*I-HF (20 U/µL; NEB cat. no R3733)	0.25
T4 Ligase Buffer (10X; NEB cat. no. B0202)	1.5
T4 DNA Ligase (400 U/µL; NEB cat. no. M0202)	0.5
Bovine Serum Albumin (10X)	1.5
dH_2O	4.25

Table 4-7: Reaction conditions for long *Bsa*I and *Bbs*I protocol.

Step	**Temperature**	**Duration**
Initial Digestion	37°C	20 sec
Digestion-Ligation Reaction (x26)	37°C	3 min
	16°C	4 min
Final Digestion	50°C	5 min
Enzyme inactivation	80°C	5 min
Hold	16°C	-

Finally, the Level 1 construct, pNAMB1-LOX, was cloned into the Level M backbone (pAGM8031), alongside a selection cassette (OsAct::Hygromycin, pICH47802) and a HvHsp::CreRecombinase construct (EC71173, from the Coen group) (see Results Figure 4-2A). Cloning was carried out using the long *BbsI* protocol (synbio.tsl.ac.uk; Table 4-7 and Table 4-8). This construct will be referred to as HS_NAM-B1.

Table 4-8: Reaction components for the long *Bbs*I protocol.

Component	1x (µL)
Acceptor plasmid (100 ng/µL)	1
Donor plasmid (200 ng/µL)	1
*Bbs*I-HF (20 U/µL; NEB cat. no R3539)	0.25
T4 Ligase Buffer (10X; NEB cat. no. B0202)	1.5
T4 DNA Ligase (400 U/µL; NEB cat. no. M0202)	0.5
Bovine Serum Albumin (10X)	1.5
dH_2O	4.25

A second construct was developed which contained *NAM-B1* under the control of the pZmUbi promoter. The Level 1 vector was generated by cloning the promoter construct (pZmUBi), the *NAM-B1* construct (pNAM-B1), and the nosT terminator into the Level 1 backbone, position 2 (pICH47742) using the long *BsaI* protocol. This construct was then cloned into the Level M backbone (pAGM8031) alongside the selection cassette in position 1 (OsAct::Hygromycin, pICH47802) using the long *BbsI* protocol. This construct will be referred to as OE_NAM-B1. The sequences of all constructs at all stages were verified using Sanger sequencing as previously detailed in section 3.2.4.1 of Chapter 3.

4.2.2 Transformation and copy number determination

The constructs were transformed into *Triticum aestivum* cv. Fielder following previously established protocols as detailed in (Wu et al. 2003). Transformation was carried out by the BRACT platform at the John Innes Centre. Positive transformants were identified at the T_0 generation by iDNA Genetics using a Taqman probe against the Hygromycin resistance cassette. This process also assigned copy numbers of the construct to each independent event.

4.2.3 GUS staining

GUS staining of the transgenic lines was carried out to confirm the presence of the construct in the individual plants. Samples were taken at an early stage of development, either within five days of germination when grown on plates (seedling stage), or within three weeks of germination when grown in soil (2-3 leaf stage). Approximately 1 cm of the oldest leaf was sampled in each case, taken from the tip of the leaf. Samples were stored in a 96-well plate, incubated in 150 µL of GUS staining buffer (Table 4-9), containing 2 mM of X-Gluc.

Table 4-9: GUS Staining solution (100 mL) containing 2 mM X-Gluc.

Volume	**Reagents**
2 mL	10% Triton X-100
10 mL	0.5 M $NaPO_4$ pH 7.0
500 µL	100 mM Potassium ferrocyanide
500 µL	100 mM Potassium ferricyanide
87 mL	dH_2O
	Add 40 µL 50 mM X-Gluc stock to 1 mL staining buffer

After collecting the leaf samples into the GUS staining buffer, the plate was placed in a vacuum desiccator and the vacuum applied for two minutes. Vacuum application was repeated until all samples had sunk to the bottom of the wells. The plate was then incubated at 37°C overnight to allow the blue staining to develop. Finally, staining buffer was replaced with 70% EtOH to stop the staining reaction and clear the plant tissue. The presence of GUS activity in individual plants was visually assayed based on the presence of blue colour in the tissue.

4.2.4 qPCR

4.2.4.1 RNA extractions

Leaf tissue was sampled from individual plants and snap frozen in liquid N_2. The stage at which the tissue was sampled depended on the experiment in question. Initial validation of construct expression was carried out on seedling leaf samples, taken between five- and seven-days following germination from the first true leaf of the plant (i.e. second leaf stage). Tissue samples from plants grown on soil in controlled environment conditions were taken at the third leaf stage. In all cases, approximately 3 cm of tissue were sampled from the tip of the leaf.

The snap-frozen tissue was then ground to a fine powder using the TissueLyser II (QIAGEN, cat. no. 85300), with dry ice to ensure the tissue remained frozen during grinding. Following grinding, up to 100 mg of tissue was placed into 1 mL of TRIzol® Reagent (cat. no. 15596026, ThermoFisher) in 1.5 mL Eppendorf tubes. The tissue was incubated in the TRIzol® Reagent for five minutes, before 200 µL of chloroform was added and incubated for a further two minutes. The tubes were then spun for five minutes at 13,000 rpm in a table-top microcentrifuge (5415D, Eppendorf®). The top, aqueous layer (approximately 600 µL) was then recovered into 500 µL isopropanol in a fresh tube and incubated at room temperature for three minutes. The tubes were then spun for 10 minutes at 13,000 rpm to produce the RNA pellet. The supernatant was discarded and the pellet was washed in 300 µL RNAse-free 70% EtOH. The tubes were centrifuged again for 10 minutes at 13,000 rpm, and the supernatant was discarded. The pellet was then dried under air for approximately three minutes before being resuspended in 100 µL RNAse-free H_2O. All prepared RNA samples were quantified on a DeNovix Spectrophotometer (DS-11). The extracted RNA samples were all stored at -80°C.

4.2.4.2 cDNA synthesis

7 µL of extracted RNA (at approximately 200 ng/µL) was incubated with 1 µL of RQ1 RNase-Free DNase 10X reaction buffer and 2 µL of RQ1 RNase-Free DNase (1 U/µL) (Promega; M6101) at 37°C for 30 minutes. Following this, 1 µL of RQ1 DNase Stop Solution was added to the reaction, and it was incubated at 65°C for 10 minutes to inactivate the DNase. Following the DNA digestion reaction, the reverse transcription reaction was carried out using the Invitrogen M-MLV reverse transcriptase (ThermoFisher, cat. no. 28025021). The full DNase-treated reaction (11 µL) was added to 1 µL of 10 mM dNTPs and 1 µL of a primer mix containing a 1:2 dilution of oligo dT and a 1:20 dilution of random primers (ThermoFisher, cat. no. 48190011). This was heated at 65°C for five minutes to allow annealing of the primers to the RNA, then immediately chilled on ice. Following this, cDNA synthesis was carried out using the full 13 µL reaction from the previous step. This was added to 4 µL 5x First Strand Buffer, 1 µL M-MLV reverse transcriptase (both ThermoFisher, cat. no. 28025021), 2 µL 0.1M DTT, and 1 µL RNAse OUT (ThermoFisher, cat. no. 10777019). The mixture was then incubated for 10 minutes at 25°C, 50 minutes at 37°C, and 15 minutes at 70°C. Following synthesis, the cDNA was stored at -20°C.

4.2.4.3 RT-PCR

Validation of construct expression was carried out using RT-PCR on cDNA samples from lines carrying the heat-shock construct (HS_NAM-B1) and the overexpression construct (OE_NAM-B1). Touchdown PCR was carried out on the cDNA samples, as detailed in Table 4-10.

Table 4-10: Cycling conditions for touchdown RT-PCR.

Step	Temperature	Time
Initial Denaturation	95°C	5 min
Touchdown (x 10)	95°C	30 sec
	T_a + 10°C to T_a	1 min
	68°C	30 sec – 1 min
Amplification (x 25)	95°C	30 sec
	T_a	1 min
	68°C	30 sec – 1 min
Final Extension	68°C	5 min
Hold	10°C	-

Where T_a refers to the annealing temperature of the specific primer set used (Table 4-12). The annealing temperature decreases by a step of 1°C in each cycle to the final annealing temperature used for the 25 amplification cycles. The reaction set up is presented in Table 4-11.

Table 4-11: Reaction components for touchdown RT-PCR.

Reagent	**1X reaction (µL)**
10X Standard *Taq* Buffer (NEB, M0273)	2.5
dNTPs (10 mM)	0.5
Taq polymerase (NEB, M0273)	0.15
$MgCl_2$ (50 mM)	0.25
dH_20	18.6
Primer_F (2 µM)	0.5
Primer_R (2 µM)	0.5
cDNA	2

The primers used for the RT-PCR reactions are listed in Table 4-12.

Table 4-12: Primers used for qRT-PCR and RT-PCR. T_a (annealing temperature) is provided for those primers used for touchdown RT-PCR assays.

Primer	**Description**	**Sequence (5′ - 3′)**	**T_a**	**Reference**
PB36	Wild-type NAM-B1, forward	TTGTCCACTGCGCCAGC	54°C	Uauy et al. (2006b)
PB37	Wild-type NAM-B1, reverse	GGCTCCGACCAAACAGTTTC	54°C	Uauy et al. (2006b)
PB47	Actin, forward	ACCTTCAGTTGCCCAGCAAT	56°C	Uauy et al. (2006b)
PB48	Actin, reverse	CAGAGTCGAGCACAATACCAGTTG	56°C	Uauy et al. (2006b)
PB71	GAPDH, forward	TTAGACTTGCGAAGCCAGCA	N/A	From Peter Isaac, iDNA
PB72	GAPDH, reverse	AAATGCCCTTGAGGTTTCCC	N/A	From Peter Isaac, iDNA
SH049	Domesticated NAM-B1, forward	AGTTGAACGGGGTCGACGAT	56°C	This study
SH051	Domesticated NAM-B1, reverse	CTGCTGCCTCTCTCAGGTTG	56°C	This study
SH001	Cre recombinase, forward (1)	TTCACCGGCATCAACGTTTT	52°C	This study
SH002	Cre recombinase, forward (2)	ACCGGCATCAACGTTTTCTT	52°C	This study
SH033	Cre recombinase, reverse	AAATGCTCCTGTCCGTTTGC	52°C	This study

4.2.4.4 qRT-PCR

Expression levels of the constructs were quantified using qRT-PCR. The SH049/SH051 primer pair was used to quantify levels of the transgenic, domesticated *NAM-B1*. The actin and GAPDH primers were used to obtain baseline expression levels of the housekeeping genes. For each experiment, the reaction set up was as shown in Table 4-13, with cycling conditions as shown in Table 4-14.

Table 4-13: qPCR reaction conditions.

Reagent	**1x reaction (µL)**
cDNA (1:10 dilution)	2
dH_2O	0.5
LightCycler® 480 SYBR Green I Master 2x (04707516001, Roche)	5
Primer F (2 µM)	1.25
Primer R (2 µM)	1.25

Table 4-14: qPCR cycling conditions.

Programme	**Temperature**	**Duration**
Pre-incubation	95C	5 min
Amplification (45 x)	95C	10 sec
	60C	15 sec
	72C	30 sec
	78C	1 sec read
Cooling to	60C	-
Melt Curve to	95C	5 acquisitions per second

All qPCR experiments were carried out in 384-well plates (04729749001, Roche) on the LightCycler® 480 instrument (Roche). Following quantification, we identified the C_p, or crossing point, value for each sample, where the slope of the amplification curve reaches its maximum. The identification of the C_p point was carried out using the Fit Points method in the LightCycler® 480 software. At the same time, all melt curves were manually inspected to ensure that the amplification produced a single peak at the expected temperature, indicative of a single homoeolog amplification.

For each biological sample, three technical replicates for each primer were carried out, and the average C_p value of the three samples was taken. If the standard deviation in C_p values between technical replicates was higher than 0.5, that sample was disregarded. Expression levels were calculated for each of the samples, using the formula $2^{(0-C_p)}$. Here, "2" corresponds to a primer efficiency of 100%; primer efficiencies were calculated individually for all primer pairs using a dilution series of cDNA from the following equation:

$$-1 + 10^{\frac{-1}{m}}$$

where *m* is the slope of the regression line between the C_p values and cDNA concentration. All primer pairs had an efficiency of > 89%. Analysis of the qPCR results found no difference when specific efficiencies were used for each primer pair, rather than treating all primers as having the

same efficiency. For simplicity, therefore, all data presented in this chapter considers all primer pairs to have an efficiency of 100%.

We calibrated the expression levels of *NAM-B1* to the housekeeping gene, *Actin*, by dividing the *NAM-B1* expression level by the expression of *Actin*. This provided us with a value for relative expression with respect to *Actin* (RE_{Actin}) that could be compared between independent biological samples. The same method was used when calibrating the expression levels to a second housekeeping gene, *GAPDH*, which was used for analysis of the *NAM-B1* over-expression construct.

4.2.5 Plant growth and phenotyping

4.2.5.1 Seedling heat shock assays

Seeds were germinated at 4°C on damp filter paper for 48 hours in the dark before they were moved to room temperature and were grown under ambient light conditions. The seedlings were grown for five to seven days following germination before the heat shock treatment. Half of the seedlings were treated to a two-hour heat shock at 38°C (see below) in order to induce expression of the *Hvhsp17* promoter as previously described (Freeman et al. 2011). The remaining seedlings were left at ambient room temperature during this period. The seedlings were sampled for GUS, qRT-PCR, and PCR assays as described previously.

4.2.5.2 Controlled environment experiments

Seeds were germinated at 4°C in the dark for 48 hours on damp filter paper, and then allowed to grow for a further 48 hours under ambient light conditions. The germinated seedlings were then sown into P96 trays containing 85% fine peat with 15% horticultural grit. They were grown until the 2-3 leaf stage before transplanting into 1 L pots. The plants experienced cycles of 16 hours of light at 20°C and 8 hours of dark at 15°C. The same conditions were used for both the *NAM-B1* heat shock and overexpression experiments.

4.2.5.2.1 Heat shock treatment

Where required, a heat shock treatment was applied to plants grown in soil. In these cases, the heat shock was applied at the 2-3 leaf stage, two or three days before the plants were transplanted into the 1 L pots. One P96 tray was taken to controlled environment chamber (Sanyo MLR-351H) and incubated at 38°C for two hours, at 85% relative humidity and light level of 3 lumens, before being returned to the original environment. A second P96 tray contained the non-heat shock samples and was left in the original environment for the period of the heat shock.

4.2.6 Genotyping PCR

To determine whether the application of heat shock was able to induce excision of the *lox*P-flanked region by the Cre recombinase, we carried out a PCR on genomic DNA (gDNA). gDNA was extracted using the standard “quick and dirty” extraction protocol, detailed in section 1.7.1.1.

Primers which flanked the *lox*P sequences were designed (Table 4-15) and combined with the reaction components shown in Table 4-16. The PCR reaction was carried out as detailed in Table 4-17. Following amplification of the PCR fragments, the products were run on a 1.5% agarose gel and visualised using a UV transilluminator.

Table 4-15: Primers for *lox*P excision PCR.

Primer	**Sequence (5′ - 3′)**	**Reference**
Ub5UTR_End_F	CGATGCTCACCCTGTTGTTT	This study
PB65_NAM-B1_R2	GGTGATACCGCGTTGCTTTT	Philippa thesis

Table 4-16: Reaction components for *lox*P excision PCR.

Using Taq:	**1X reaction**
10X Standard *Taq* Buffer (NEB, M0273)	2.5
dNTPs (10 mM)	0.5
Taq polymerase (NEB, M0273)	0.125
$MgCl_2$ (25 mM)	0.5
dH_20	18.6
Primer F (10 μM)	0.5
Primer R (10 μM)	0.5
gDNA	2

Table 4-17: Reaction conditions for *lox*P excision PCR.

Step	**Temperature**	**Time**
Initial Denaturation	95°C	5 min
Touchdown (x 10)	95°C	30 sec
	T_a + 10°C to T_a	1 min
	68°C	30 sec – 1 min
Amplification (x 35)	95°C	30 sec
	T_a	1 min
	68°C	3 min
Final Extension	68°C	5 min
Hold	10°C	-

4.3 Results

4.3.1 Construct design

We designed two transgenic approaches to investigate whether *NAM-B1* can induce premature senescence when ectopically over-expressed early in development. First, we adapted an inducible system which allowed for constitutive induction of *NAM-B1* expression following heat-shock, referred to as HS_NAM-B1 (Figure 4-2A). Before heat shock, the construct expresses the GUS gene under the control of the constitutive *ZmUbi* promoter. The stop codon in the GUS gene prevents readthrough to the downstream *NAM-B1* gene. After a heat shock of 38°C, the barley heat shock promoter *Hvhsp17* is induced leading to expression of Cre recombinase. The Cre recombinase then cuts at two *lox*P sites which flank the reporter gene, GUS. Following *lox*P excision, recombination brings the *NAM-B1* coding sequence in frame with the *ZmUbi* promoter, leading to constitutive expression of the gene (see Figure 4-1 in section 4.1). This concept was initially developed in the lab of Enrico Coen, by Samantha Fox and Annis Richardson, who kindly shared their constructs and expertise (Table 4-2).

Our second approach was more straightforward, placing *NAM-B1* directly under the control of the *ZmUbi* promoter, referred to as OE_NAM-B1 (Figure 4-2B). In both cases, the constructs contained hygromycin selection, under the control of the *pOsAct* promoter in opposite orientation to the remainder of the cassette. All constructs were cloned using Golden Gate into the Level M backbone (pAGM8031). The *NAM-B1* sequence used in both constructs was 'domesticated' to remove the *BbsI* sites present, thus differentiating it from the wild-type *NAM-B1* sequence (Appendix 8.1.1 and 8.1.2).

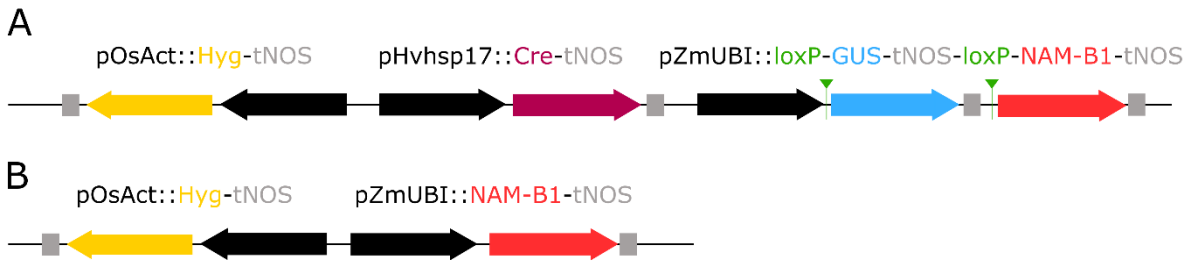

Figure 4-2: Two independent constructs were developed to test the effect of ectopic *NAM-B1* expression in wheat. (A) A heat shock inducible construct, HS_NAM-B1, utilised the barley *Hvhsp17* promoter to drive expression of Cre recombinase following heat shock, leading to irreversible excision of the reporter gene and induction of *NAM-B1*. (B) A constitutive overexpression construct, OE_NAM-B1, driving *NAM-B1* under the promoter *ZmUbi*, derived from maize.

4.3.2 Copy number analysis of T_0 plants

The cloned constructs were transformed into *Triticum aestivum* cv. Fielder, a standard variety used for *Agrobacterium*-mediated transformation (Adamski et al. 2018), by the BRACT platform at JIC. At the T_0 stage, the copy number of the positive transformants was obtained by iDNA Genetics

using a Taqman assay specific for the Hygromycin selection gene. We recovered five independent transgenic lines for the over-expression cassette, ZmUbi::NAM-B1, and 17 independent lines for the heat shock cassette (Table 4-18). We then obtained the seeds from the positive T_0 plants and carried forward the T_1 seeds for further screening.

Table 4-18: Copy number of positive T_0 transformants.

Construct	T_0 Plant	Copies of Hygromycin
OE_NAM-B1	2010.27.1	78
	2087.1.1	19
	2087.4.2	13
	2087.1.4	9
	2087.2.2	1 to 2
HS_ NAM-B1	2020-20-01	46
	2020-10-1	15
	2020-4-1	12
	2020-19-01	10
	2020-5-1	10
	2020-12-1	8
	2020-6-1	4
	2020-13-1	4
	2020-2-1	4
	2020-8-1	4
	2020-22-01	2
	2020-23-01	2
	2020-3-1	2
	2020-21-01	1
	2020-20-02	1
	2020-24-01	1
	2020-11-1	1

4.3.3 Validation of the heat shock Cre-loxP construct

4.3.3.1 GUS is expressed in the constructs

To screen individual T_1 seedlings for presence of the heat shock construct we carried out GUS staining on small samples of the seedling tissue. Seedlings containing the construct exhibited GUS staining compared to the complete lack of staining found in the 0-copy control (Figure 4-3). We found that all but two of the independent lines tested showed GUS staining in at least three of the five seedlings tested (Table 4-19; page 156).

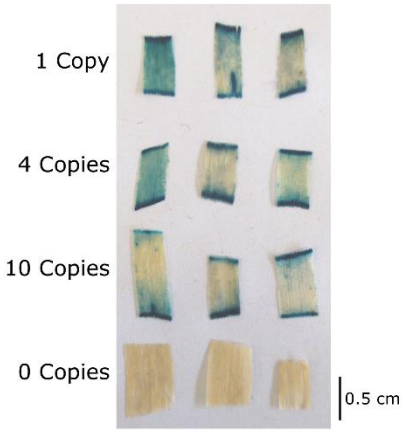

Figure 4-3: GUS staining is only present in plants containing the HS_NAM-B1 construct. Lines with one copy (2020-20-01), four copies (2020-2-1), and ten copies (2020-5-1) are presented as illustrative examples. The null transformant control (2010-54-01; 0 copies) is shown at the bottom.

4.3.3.2 Expression of Cre recombinase

After identifying positive transformants, we next tested whether we were able to induce expression of Cre recombinase following heat shock. We carried out Reverse Transcriptase PCR (RT-PCR) on cDNA extracted from four seedlings which had undergone heat shock, and four controls which contained the construct but which had not experienced heat shock (Figure 4-4). The heat shock treatment was applied for two hours at 38°C. Tissue was sampled for RNA extraction two hours following the end of the heat shock. We observed expression of Cre recombinase in all of the samples which had undergone heat shock, but none in the untreated controls or the water control. The PCR product was sequenced using Sanger sequencing and was found to match the expected Cre recombinase sequence.

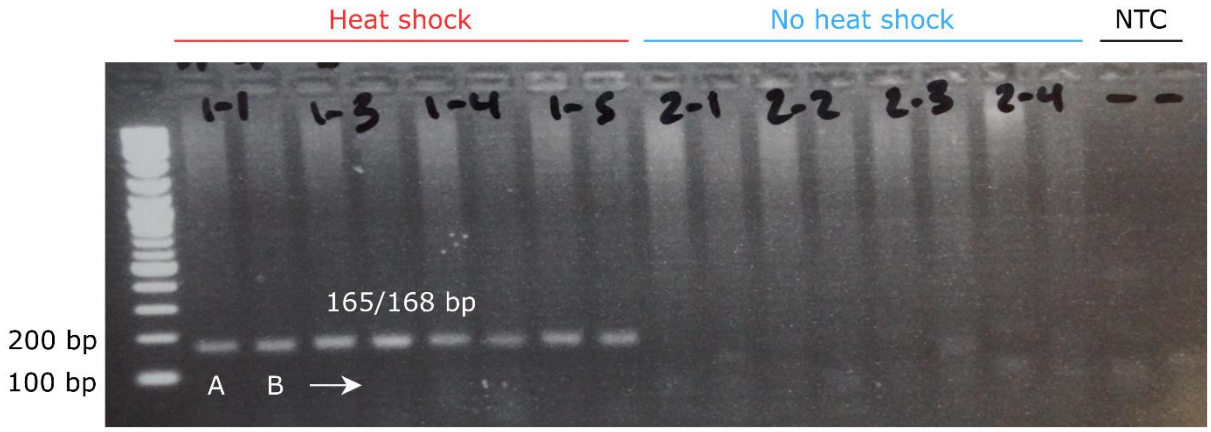

Figure 4-4: Cre recombinase expression in HS_NAM-B1 transgenic lines. Here RT-PCR was carried out on each sample using two primer pairs- SH033 and SH001 (A) and SH033 and SH002 (B); see primer sequences in Table 4-12. The expected band size was 165 bp for primer pair A, and 168 bp for primer pair B. NTC is no transcript control (water).

4.3.3.3 Heat shock drives excision of the reporter gene

After demonstrating that the Cre recombinase is expressed following heat shock, we then investigated whether the Cre recombinase is able to excise the reporter gene. To do this, we carried out a PCR on genomic DNA extracted from seedlings before and after heat shock. Tissue was sampled from the same seedling one hour before and one day after heat shock. We identified successful excision of the reporter construct, illustrated by the loss of the large 2490 bp band and gain of the small 151 bp band depicted in Figure 4-5A, in the vast majority of lines tested (Table 4-19). A representative example for three independent lines (2020-3-1, 2020-24-01, and 2020-21-01) is shown in Figure 4.

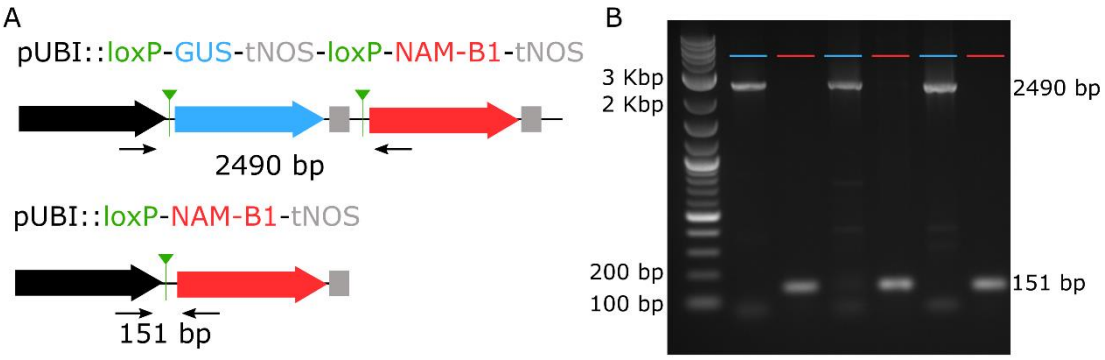

Figure 4-5: Diagram of *lox*P-flanked construct. (A) Excision of the GUS gene and expression of NAM-B1 occurs following heat-shock and expression of Cre Recombinase. The approximate location of the primers used to validate the GUS excision are shown in black, alongside the expected band size. (B) Gel image shows the resultant DNA bands before (blue) and after (red) heat shock.

Overall, we found that 12 of the 14 independent T_0 lines tested expressed the reporter gene GUS and successfully carried out excision of the reporter gene following heat shock (Table 4-19). Two lines, 2020-11-01 and 2020-22-01, did not exhibit GUS staining. The large PCR band, indicative of the presence of the reporter gene within the construct, was present for the 2020-22-01 both before and after heat shock. This suggests that the construct is indeed present in the 2020-22-01 line, but that the heat shock was unable to induce excision of the *lox*P sites. As there was no GUS expression in this line either, we hypothesized that the construct had been inserted in a location which limited or eliminated construct expression. The large band was not present in the second line, 2020-11-01, either before or after heat shock. This suggested that either the parent T_0 line did not pass on the construct to its progeny, with the construct only being transformed into somatic rather than germ line cells, or that the five seedlings tested happened to all segregate without the construct. As many independent lines did have GUS expression and successful GUS excision, we did not investigate this further.

Table 4-19: 12 of 14 independent T_0 lines successfully expressed GUS and carried out heat-shock-induced *lox*P recombination. The percentage of T_1 seedlings that germinated, showed GUS staining before heat shock, and demonstrated successful excision of the reporter gene following heat shock are indicated in the respective columns. N = 5 for all except those indicated with *, where N = 4.

T_0 Line	**Hygromicin Copy Number**	**Germination**	**GUS Staining**	**Excision PCR**
2020-20-02	1	100%	100%*	100%
2020-21-01	1	80%	100%*	100%*
2020-11-01	1	100%	0%	0%
2020-24-01	1	100%	80%	80%
2020-3-01	2	100%	60%	60%
2020-23-01	2	100%	40%	20%
2020-22-01	2	100%	0%	0%
2020-2-1	4	100%	100%	100%
2020-13-1	4	80%	100%*	50%*
2020-8-1	4	100%	100%	80%
2020-5-1	10	100%	100%	80%
2020-19-1	10	100%	60%	60%
2020-4-1	12	100%	100%	100%
2020-20-01	46	100%	100%	100%
2020-54-01	0	100%	0%	0%

4.3.4 NAM-B1 is expressed following heat shock

After characterising the heat-shock-driven excision of GUS by the Cre recombinase, we then turned to validating that *NAM-B1* itself is indeed expressed following heat shock. RNA was extracted from seedlings 24 hours after they experienced heat shock. Controls which had not experienced heat shock were also sampled. We then carried out RT-PCR using primers specific for the domesticated version of *NAM-B1* included in the construct (Table 4-12). In the initial test of five seedlings from the 2020-2-1 line, we found that four seedlings induced expression of *NAM-B1* within 24 hours following heat shock (Figure 4-6). No expression of *NAM-B1* was visible in any of the negative controls. The amplified PCR product was then verified using Sanger sequencing and was found to match the expected domesticated *NAM-B1* sequence. This demonstrates that the construct can be successfully used for induction of *NAM-B1*.

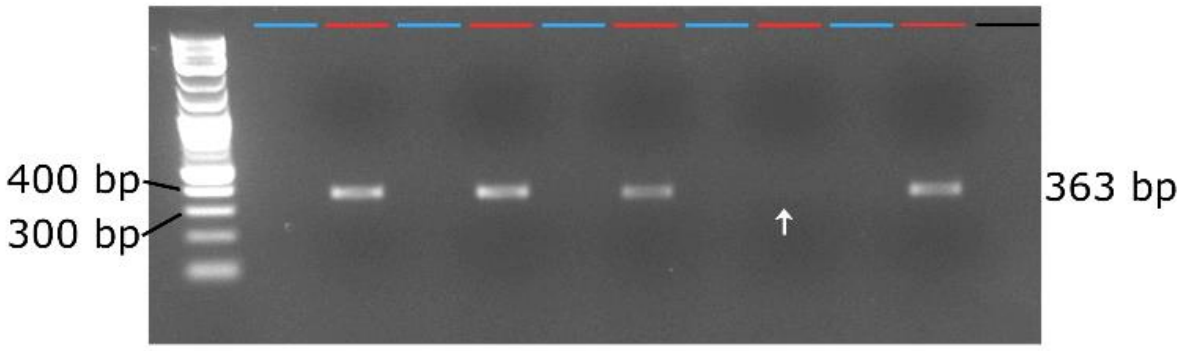

Figure 4-6: *NAM-B1* is only expressed in heat-shocked transgenic lines. Expression of *NAM-B1* was tested in heat-shocked (red) and non-heat-shocked (blue) transgenic lines using the primer pair SH049 and SH051, which is specific for the domesticated *NAM-B1* cDNA. Note that one of the treated samples shows no *NAM-B1* expression (white arrow). The no template control (NTC) is clear (black).

We then carried out qRT-PCR on these samples, discarding the single outlier sample which had no expression of *NAM-B1* following heat shock. We found significant upregulation of *NAM-B1* expression following heat shock compared to seedlings from the same line which had not experienced heat shock (Figure 4-7). Amplification of *NAM-B1* occurred within 20 to 22 cycles in the heat shock treated samples, compared to the 26-29 cycles required for amplification of *Actin* indicating that *NAM-B1* could be over 100-fold more expressed than *Actin*. Amplification by the *NAM-B1* primers was only observed after 31-34 cycles for the untreated controls, over 500-fold less than that observed in the treated samples.

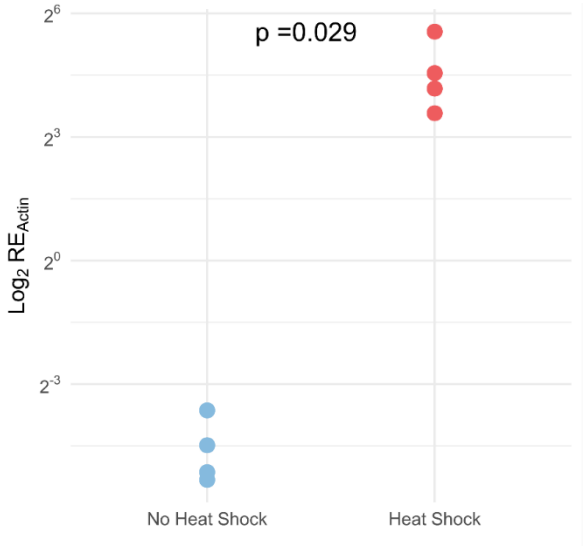

Figure 4-7: NAM-B1 expression is significantly induced within 24 hours following heat shock. The statistical comparison was carried out using the Wilcoxon rank-sum test. The expression values have been altered with a log_2 transformation. N = 4 for each treatment.

4.3.5 Heat shock treatment of the transgenic lines does not induce premature senescence in seedlings

After confirming that the HS_NAM-B1 lines could induce *NAM-B1* expression following heat shock, we then tested whether induction of *NAM-B1* expression led to premature cell death. This was first tested in seedlings grown on plates which were treated with a two-hour heat shock at 38°C. The seedlings were then visually scored for symptoms of premature senescence, such as necrosis or chlorosis, for ten days post heat shock (DPHS). Eight independent transgenic lines were tested, including six lines which expressed the GUS reporter gene and two lines which did not (Table 4-20). The zero-copy control, 2010-54-01, was also included in the experiment. These seedlings were tested for GUS expression before heat shock, and successful excision of the reporter gene following heat shock, as detailed in Table 4-19. We found that none of the heat shocked seedlings showed any symptoms of premature senescence, as compared to the seedlings which had not experienced heat shock. An image of the 2020-20-02 and 2020-21-01 lines two days after heat shock is shown in Figure 4-8, where no symptoms of visual senescence can be observed in either the treated or untreated seedlings.

Table 4-20: Heat shock treatment of transgenic *NAM-B1* lines as seedlings does not induce premature senescence. DPHS is days post heat shock.

Line	Copy Number	GUS expression	1 DPHS	2 DPHS	3 DPHS	4 DPHS	7 DPHS	10 DPHS
2020-20-02	1	Yes	No	No	No	No	No	No
2020-21-01	1	Yes	No	No	No	No	No	No
2020-24-1	1	Yes	No	No	No	No	No	No
2020-3-01	2	Yes	No	No	No	No	No	No
2020-23-01	2	Yes	No	No	No	No	No	No
2020-8-01	4	Yes	No	No	No	No	No	No
2020-11-1	1	No	No	No	No	No	No	No
2020-22-01	2	No	No	No	No	No	No	No
2010-54-01	0	No	No	No	No	No	No	No

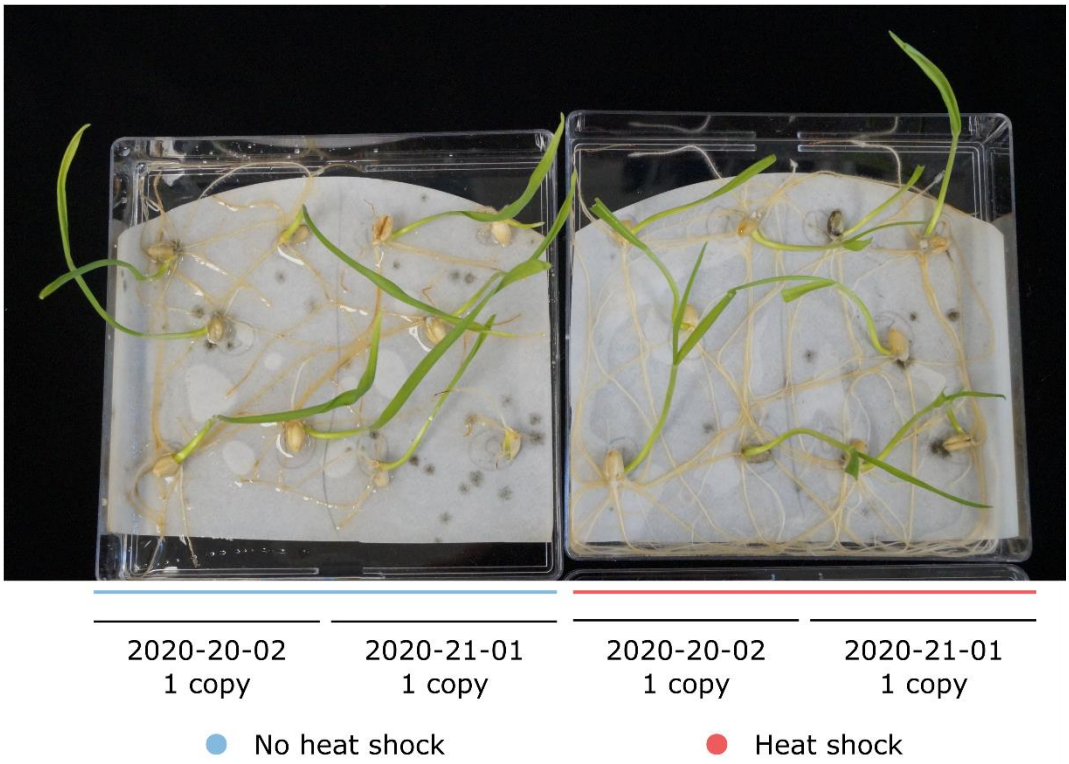

Figure 4-8: HS_NAM-B1 seedlings show no signs of premature senescence following heat shock. Two representative lines, 2020-20-02 and 2020-21-01, were photographed two days after a two hour, 38°C heat shock was applied to one half of the plants (red). The other half were kept at room temperature during the heat shock (blue).

This experiment suggested that the expression of *NAM-B1* alone was not sufficient to induce premature cell death in seedling plants. However, it was possible that *NAM-B1* expression does induce premature senescence, but that it takes longer than ten days for visual symptoms to be observed. Additionally, while the successful excision of the reporter gene was confirmed for these seedlings, the corresponding expression of *NAM-B1* was not confirmed. Due to the limitations of growing wheat seedlings on damp filter paper, this experiment could not realistically be continued beyond ten DPHS. To investigate this question further, therefore, we carried out the heat shock experiment on plants grown in soil.

4.3.6 NAM-B1 expression does not lead to premature senescence under controlled environment conditions

24 seedlings from each of three independent lines (2020-20-02, 2020-2-1, and 2020-5-1), along with 24 seedlings from the zero-copy control 2010-54-01, were sown into two P96 trays and grown under controlled environment conditions. Half of the seedlings for each line were sown in each tray, allowing us to subject only one tray to heat shock at the three-leaf stage of growth. This was essential due to space constraints, allowing all plants to undergo heat shock at the same time. All plants were tested for GUS expression before heat shock, and we excluded plants with no visible blue staining from further analysis.

Following heat shock, we began monitoring visual symptoms of senescence. No evidence of senescence was visible in the early developmental stages of the plant, corresponding with the results from the seedling assay (Table 4-20). At heading, we began measuring the greenness of the leaf as a proxy for chlorophyll content using the SPAD meter and continued our visual scoring of senescence. We found no significant difference in leaf and peduncle senescence timing in the three transgenic lines as a result of heat shock (Figure 4-9). Curiously, we did see a significant delay in both flag leaf and peduncle senescence in the zero-copy line as a result of heat shock. The SPAD measurements recapitulated the visual scoring, showing no substantial difference between the heat shock treated and untreated plants.

Table 4-21: Summary statistics for visual senescence in the heat shock transgenic lines.

Line	Copy Number	Heat Shock Treatment	N	Days from Heading to Flag Leaf Senescence		Days from Heading to Peduncle Senescence	
				Mean	SD	Mean	SD
2010-54-01	0	No	12	38	4.0	42	3.2
2010-54-01	0	Yes	11	44	5.3	47	4.9
2020-20-02	1	No	11	43	4.4	46	4.6
2020-20-02	1	Yes	11	44	5.7	46	5.9
2020-2-1	4	No	12	39	5.8	44	2.5
2020-2-1	4	Yes	12	44	4.4	46	4.7
2020-5-1	10	No	12	42	8.5	45	4.8
2020-5-1	10	Yes	12	41	4.6	46	5.8

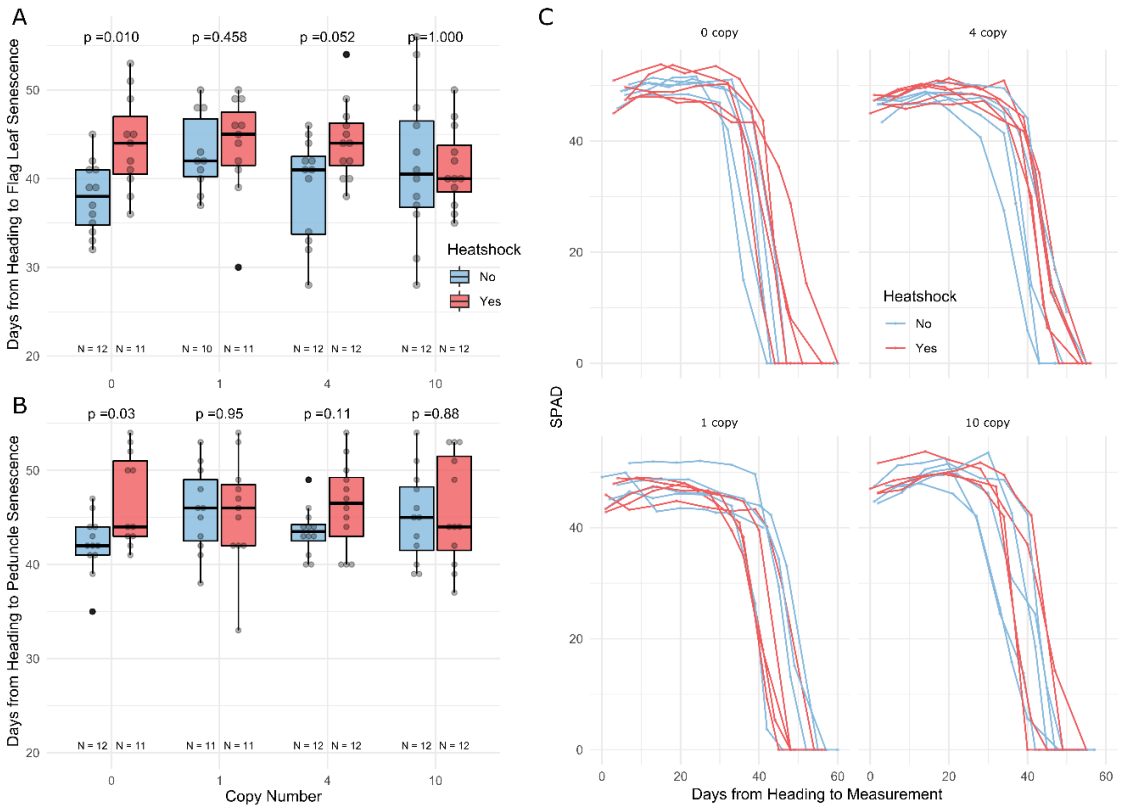

Figure 4-9: Over-expression of *NAM-B1* following heat shock does not influence senescence onset or progression. Visual senescence was scored for the flag leaf (A) and peduncle (B). Flag leaf greenness was also scored using a SPAD meter as a proxy for chlorophyll content (C). SPAD measurements were taken at intervals following heading for five individual plants from each line and condition, except the 10-copy line after heat shock for which N = 4. Statistical comparisons between the treated and untreated plants lines were carried out using the Wilcoxon rank-sum test.

To be sure that the heat shock induced *NAM-B1* expression, we carried out the gDNA excision PCR as described in Figure 4-5 on all samples before and after heat shock treatment. We found that all treated samples which expressed GUS before heat shock also showed the expected loss of the large PCR product following treatment. Those samples which did not express GUS did not recover the expected PCR products either before or after heat shock, supporting the conclusion that these plants were transgene free (results not shown).

Having verified that the *lox*P excision had taken place following heat shock treatment, we then sampled a subset of the plants to test *NAM-B1* expression. qRT-PCR was carried out as previously on RNA extracted from leaf tissue sampled one day before and one day after heat shock at the three-leaf stage (Zadok 13). We found that *NAM-B1* expression was significantly upregulated in the transgenic lines ("1+ copies") following heat shock treatment ($p < 0.01$, Wilcoxon rank-sum test; Figure 4-10). This corresponded to a difference of approximately 7 cycles before amplification, or a 128-fold increase in expression.

There was no significant change in expression in the 0-copy control following treatment ($p = 0.23$, Wilcoxon rank-sum test; Figure 4-10). We also tested the positive transgenic lines before heat

shock treatment and found no difference in *NAM-B1* expression levels between the plants that would not be treated, and those that would (p = 0.61, Wilcoxon rank-sum test; Figure 4-10). This demonstrated that the difference between the two groups was due to the application of a heat shock treatment.

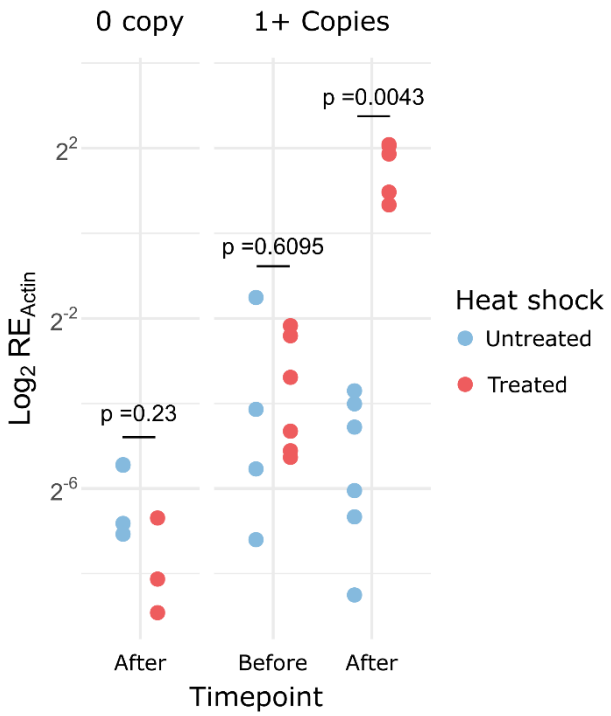

Figure 4-10: *NAM-B1* expression is significantly induced following heat-shock. Expression levels of *NAM-B1* were obtained using qRT-PCR and normalised using the level of *Actin* expression. To better display the data, expression levels were adjusted by a log_2 transformation.

These results demonstrated that, following heat shock treatment, the expression of *NAM-B1* is increased in the positive transgenic lines. However, there did appear to be a small level of *NAM-B1* gene expression in the untreated plants (Figure 4-10). We found no significant difference in *NAM-B1* expression levels between the zero-copy control and the untreated transgenic lines (p = 0.47, Wilcoxon rank-sum test). This comparison included all transgenic lines sampled before heat shock and the untreated transgenic plants tested following the application of heat shock. Nevertheless, individual plants appeared to show higher levels of *NAM-B1* expression than would be expected, approaching 0.35 X the level of actin. This is far higher than the average background amplification seen in the 0-copy controls, of approximately 0.006 X. This may suggest that there is a slight background level of *NAM-B1* expression in the untreated heat shock controls. However, we found that the untreated lines with the highest levels of background *NAM-B1* expression also had the weakest amplification of *Actin* (at greater than 30 cycles), suggesting that this apparent high level of expression may be due in part to poor amplification of the *Actin* control.

Given the unexpected delay in senescence in the zero-copy control, we repeated the experiment in the same controlled environment conditions using T_1 seedlings which were siblings to those used in

the first experiment (Figure 4-9). After again scoring the plants for visual senescence following heat shock at the three-leaf stage, we observed no variation in flag leaf or peduncle senescence timing for any of the transgenic lines as a result of heat shock (Figure 4-11, Table 4-22). Here we also saw no variation in the zero-copy control, as we would expect.

Table 4-22: Summary statistics for the repeat experiment with the heat shock lines.

Line	Copy Number	Heat Shock Treatment	N	Days from Heading to Flag Leaf Senescence		Days from Heading to Peduncle Senescence	
				Mean	SD	Mean	SD
2010-54-01	0	No	12	45	2.7	52	3.5
2010-54-01	0	Yes	12	47	4.2	51	3.1
2020-20-02	1	No	9	45	3.5	53	5.3
2020-20-02	1	Yes	11	46	3.8	54	5.2
2020-2-2	4	No	12	44	3.4	54	4.7
2020-2-2	4	Yes	12	45	3.0	53	4.1
2020-5-1	10	No	12	45	3.0	53	4.3
2020-5-1	10	Yes	10	45	5.3	52	6.8

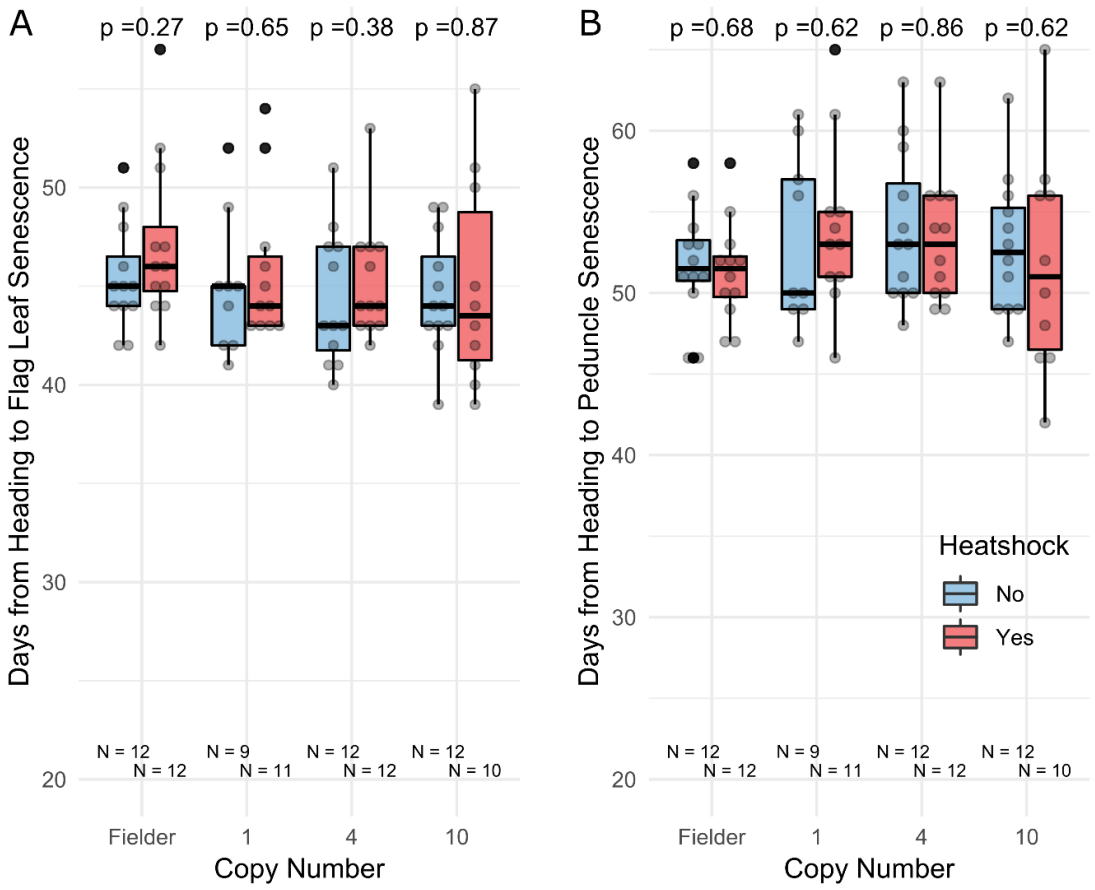

Figure 4-11: Expression of *NAM-B1* consistently does not affect plant senescence in controlled environment conditions. As previously, visual senescence was scored for the flag leaf (A) and peduncle (B). Statistical comparisons between the treated and untreated plants were carried out using the Wilcoxon rank-sum test.

4.3.7 *Constitutive over-expression of NAM-B1 does not lead to premature senescence*

Alongside the heat-shock inducible construct discussed in detail in 4.3.1, we also developed transformants which constitutively over-expressed *NAM-B1* under the control of the *Zea mays* Ubiquitin promoter. A diagram of this construct is provided in Figure 4-2B. Following assessment of the construct copy number in the T_0 plants (Table 4-18), we took forward the T_1 seeds for further characterisation. First, we confirmed that *NAM-B1* was being expressed in the T_1 generation. Using RT-PCR, we were able to show that the domesticated version of *NAM-B1* specific to the construct was being expressed in seedling tissue (Figure 4-12). Using Sanger sequencing, we confirmed that the amplified PCR products corresponded to the expected cDNA sequence.

The level of *NAM-B1* amplification varied for each individual T_1 plant and between lines. Line 2087-1-1 had consistently low expression of *NAM-B1*, with no band visible in the second sample and only faint band visible in the other two samples. Line 2087-1-4 also had low amplification of *NAM-B1* in two of the three samples, producing only faint bands. The remaining three lines all showed strong amplification of the *NAM-B1* sequence. Some of this variation will be due to variation in cDNA concentration, though samples were normalised to approximately 250 ng/µL before cDNA synthesis from the extracted RNA (500 ng in total). The variation may also be due to differences in *NAM-B1* expression levels both within and between the independent lines.

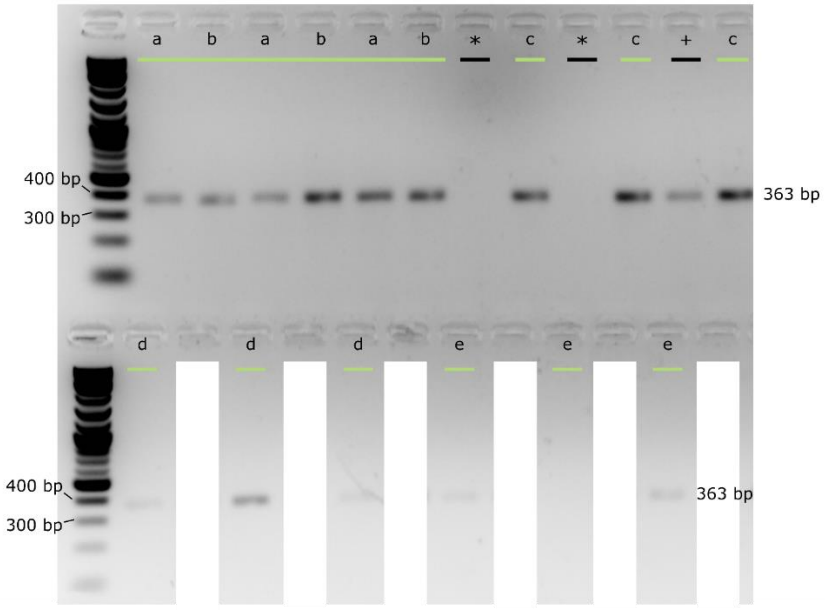

Figure 4-12: NAM-B1 is expressed in the OE construct. RT-PCR was carried out using the primers SH049 and SH051 which are specific to the domesticated NAM-B1 cDNA (Table 4-12). Three individual T_1 plants were sampled from each of the five independent T_0 lines- 2087-4-2 (a; 13 copies), 2010-27-1 (b; 78 copies), 2087-2-2 (c; 1 to 2 copies), 2087-1-4 (d; 9 copies), and 2087-1-1 (e; 19 copies). Two individual T_1 plants from the 0-copy control lines, 2010-54-1, were also tested as negative controls (*), and a single heat-shock treated sample from the 2020-2-1 line was included as a positive control (+). The expected product size of the primer pair is 363 bp. Lanes containing PCR reactions unrelated to the figure were hidden with white boxes.

In order to investigate this variation further, we quantified the levels of *NAM-B1* expression in the same plants. Using primers specific for the domesticated version of *NAM-B1* present in the construct, we carried out qRT-PCR on the three individual T_1 plants for each of the five independent transgenic lines, relative to the housekeeping gene *GAPDH* (Figure 4-13). We found that all transgenic plants expressed the domesticated *NAM-B1* while no expression was found in the 0-copy control. There was substantial variation in *NAM-B1* expression levels between the five lines, as shown in Figure 4-13A. One line, 2087-4-2 which contains 13 copies of the transgene, expressed *NAM-B1* between 1 X and 12 X the level of *GAPDH* expression. The remaining four lines were expressed at levels at or below those of *GAPDH* (Figure 4-13B). All seedlings tested expressed *NAM-B1* above the negative baseline of the 0-copy control, corresponding with the observed bands in RT-PCR (Figure 4-12). The lowest level of expression observed was 0.04 X that of GAPDH, with a median expression level of 0.56 X GAPDH and an average expression level of 1.51 X GAPDH across all of the T_1 plants.

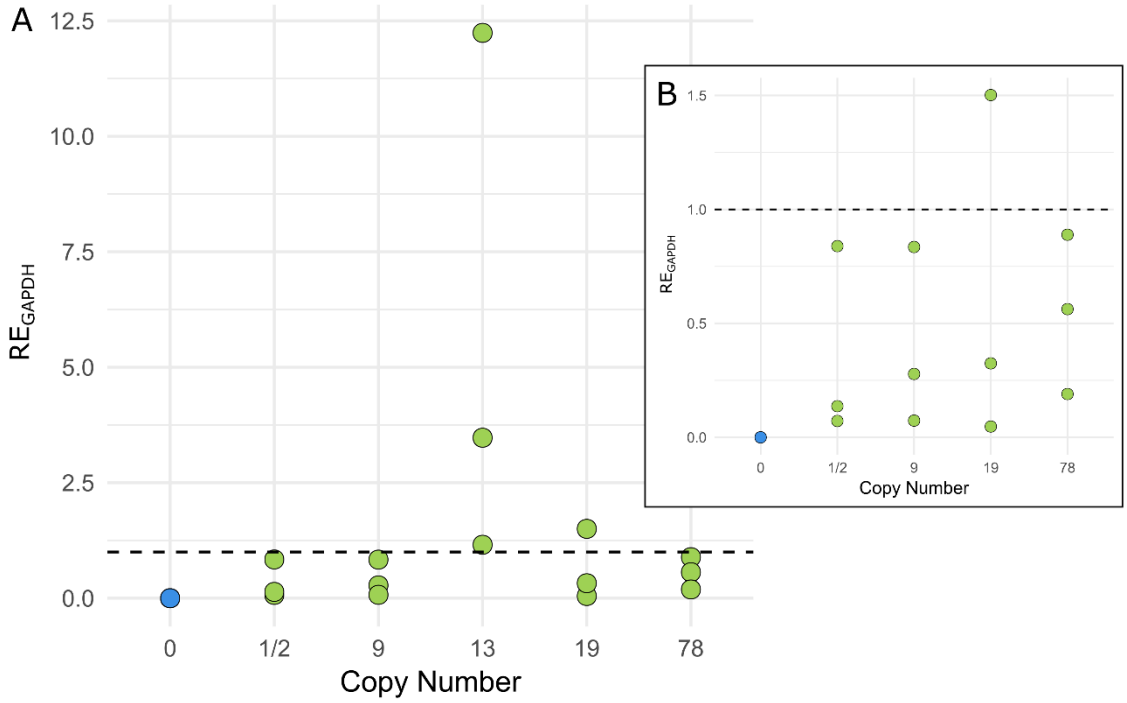

Figure 4-13: *NAM-B1* is expressed in the five independent OE_NAM-B1 transgenic lines. qRT-PCR was carried out on the samples using the domesticated NAM-B1 specific primer pair SH049/SH051, relative to the housekeeping gene *GAPDH* (Table 4-12). The *GAPDH* expression level is indicated by the dashed line. No expression is seen in the null-transformant control (0 copy; blue). All transgenic lines are shown in panel (A), while a subset of lines is shown in panel (B) to highlight the variation in expression level between 0 and 1 X GAPDH.

To investigate whether the expression of *NAM-B1* in these transgenic lines has any effect on senescence timing, we grew the transgenic lines under controlled environment conditions. These were the same conditions under which the heat shock lines were grown. The plants were scored for visual signs of senescence and we found that, as with the heat shock lines, there was no variation in flag leaf senescence timing in the lines over-expressing *NAM-B1*. We did however observe a

striking and unexpected pattern of delayed peduncle senescence in these lines, compared to the zero-copy control. Only line 2010-27-1, containing 78 copies of the transgene, showed no significant delay in senescence. The length of the delay in senescence was also substantial, with peduncle senescence taking over 90 days from heading for two of the transgenic lines, compared to only 55 for the control (Table 4-23). However, much greater variation in senescence timing was observed in the transgenic lines compared to the control with the standard deviation of peduncle senescence reaching 28 days for line 2087-2-2 (1/2 copies).

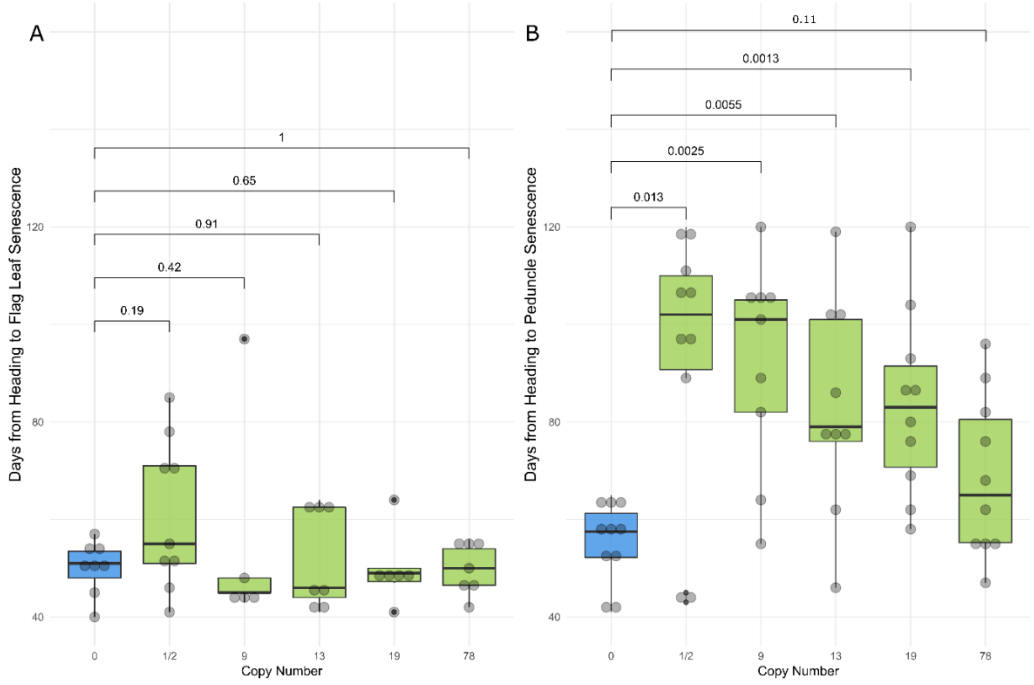

Figure 4-14: Constitutive over-expression of *NAM-B1* does not affect leaf senescence but leads to delayed peduncle senescence. Individual plants were scored for visual flag leaf (A) and peduncle (B) senescence, as previously described. All statistical comparisons between the transgenic lines and the zero-copy control (blue) were carried out using the pairwise Wilcoxon rank-sum test.

Table 4-23: Summary statistics for the OE_NAM-B1 transgenic lines.

Line	Copy Number	N	Days from Heading to Flag Leaf Senescence		Days from Heading to Peduncle Senescence	
			Mean	Standard Deviation	Mean	Standard Deviation
2010-54-01	0	10	47	8.4	55	8.2
2087-2-2	1/2	10	58	16.8	93	27.6
2087-1-4	9	9	45	21.3	92	21.4
2087-4-2	13	9	49	11.4	83	22.2
2087-1-1	19	10	44	10.1	84	19.0
2010-27-1	78	10	45	8.6	69	16.6

We complemented the visual senescence scores with measurements of flag leaf greenness using the SPAD meter, and similarly found high levels of variation between individual plants within the same line (Figure 4-15). The SPAD profiles for individual plants in the control line match the trend

of the LOESS smoothed line, with variation in exact senescence timing amongst the individual plants but with most senescing completely within 60 days of heading. This is also seen in three of the five transgenic lines, with the individual plants reasonably closely approximating the locally-estimated scatterplot smoothing (LOESS) line. In contrast, individual plants in the 1/2 and 13 copy lines showed almost no signatures of senescence far beyond 60 days post-heading. Notably, the visual flag leaf senescence scores of these lines appear to separate into an early and a late senescence group, with visual flag leaf senescence occurring before or after 60 days post-heading (Figure 4-14A).

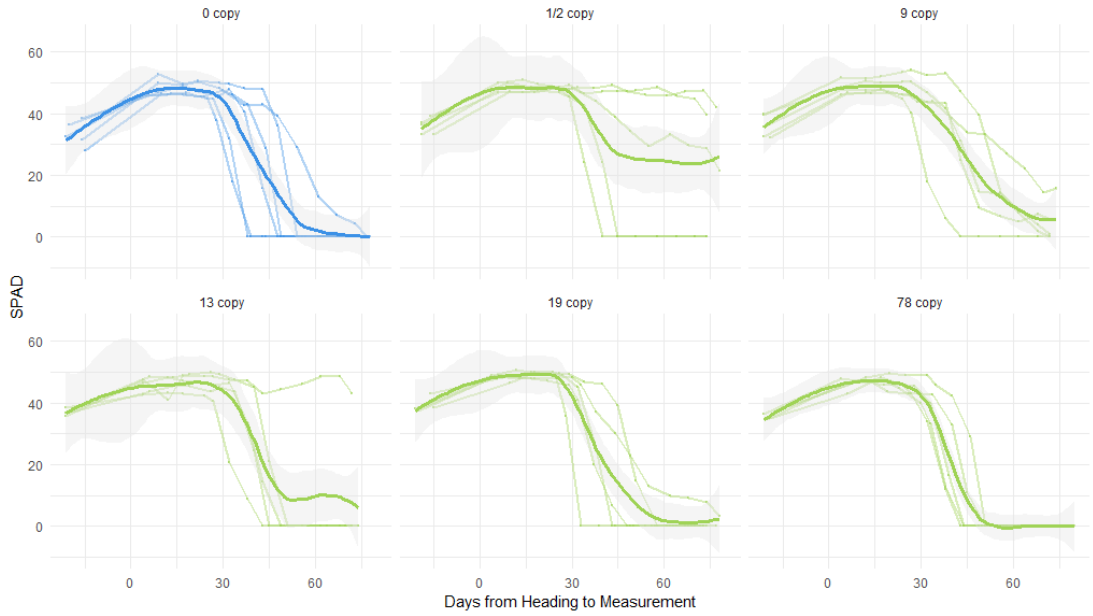

Figure 4-15: Individual plants over-expressing *NAM-B1* show variable leaf senescence profiles. SPAD measurements were taken for five individual plants of each independent line and the zero-copy control and intervals following heading. A single pre-heading timepoint was taken for each plant between 20 and 15 days prior to heading. The senescence profile for each line is depicted using a LOESS smoothed fit line, shown in a darker colour with the 95% confidence interval in light grey.

4.4 Discussion

4.4.1 The heat shock inducible construct works successfully in wheat

Here we have shown that we can use the barley heat shock promoter to drive an inducible Cre-*lox*P system in wheat. It was previously shown that this heat shock promoter could successfully drive the expression of the reporter gene GUS in transgenic wheat following heat shock (Freeman et al. 2011). Combining this promoter with the Cre-*lox*P system, as established in the lab of Enrico Coen and demonstrated here, opens up the possibilities of testing many different wheat genes in a temporally sensitive manner. This can be particularly useful in cases like *NAM-B1*, where the developmental stage at which the gene is able to induce a change in form or function in the plant is unknown.

The construct could also be used in an inverted manner, where the gene of interest is located between the *lox*P sites and a reporter gene is located downstream of the final *lox*P site. This could allow ectopic expression of an exotic gene up to a desired stage of development, after which the transgene would be excised following heat shock. Additionally, previous work with this construct in barley has shown that shorter periods of heat shock are able to induce a mosaic effect, whereby Cre-recombinase activity is induced in single cells (Richardson, personal communication). As a result, the gene of interest only becomes expressed in a small subset of the cells. This result can be used to track the division and growth of individual cells, particularly when the induced gene is a reporter such as GFP. Such a system would be particularly appropriate for detailed studies of development in, for instance, wheat meristems.

Yet the way our Cre-*lox*P system is regulated leads to possible difficulties. Notably, heat-shock inducible constructs, while providing well-controlled gene expression, suffer from the associated biological responses follow a heat shock treatment. Heat shock can, for example, lead to infertility (Draeger and Moore 2017) and meiotic defects (Dowrick 1957) in wheat if experienced near anthesis. Application of any stress treatment is also likely to lead to a substantial shift in gene expression patterns (Qin et al. 2008), at least temporarily, possibly leading to confounding results following treatment. One alternative could be to use a chemically inducible system, as reviewed earlier in the chapter (Table 4-1). However, many of these inducers come with their own drawbacks, including toxicity in some cases, and are also likely to induce global changes in gene expression following application.

In other species, particularly in mice where use of the Cre-*lox*P technology is more mature, tissue-specific promoters have also been used to govern the temporal and spatial expression of Cre-recombinase (Jimeno et al. 2006). As gene expression profiles are better characterised in wheat, it will be possible to characterise specific wheat promoters that have desired tissue expression profiles. These promoters could then be used to drive Cre-expression without the need for an exogenous stimulus. The expression pattern of the Cre recombinase could be coupled with constitutive promoters that would then lead to constant expression of the gene of interest. Alternatively, the system could use a promoter which would regulate a specific pattern of expression for the gene of interest. This would allow the decoupling of the initial Cre-recombinase expression from the pattern of expression for the gene of interest. More complex regulatory systems could place the Cre recombinase under the control of two different promoters, through a Split-Cre system where the N and C terminal domains of the protein are expressed separately (Hirrlinger et al. 2009). Only co-expression of the two components would be sufficient for its recombinase activity, facilitating the establishment of highly cell-type specific Cre expression.

However, our work here has also highlighted possible limitations of the system. While the results are not yet conclusive, it may be the case that low-levels of the gene of interest are expressed before application of the heat shock treatment (Figure 4-10). It is indisputable that, following heat

shock application, levels of transgenic *NAM-B1* expression in the treated plants are significantly higher than in the treated plants (e.g. see Figure 4-7 and Figure 4-10). Nevertheless, further work should be carried out to determine whether the HS_NAM-B1 construct is truly "leaky," or if signals of low-level expression in untreated plants are due to either low levels of amplification of the non-transgenic copies of *NAM* or contamination during analysis. It is important to note that we saw no significant increases in *NAM-B1* expression compared to that observed in the 0-copy control (Figure 4-10). This suggests it is most likely that the expression observed in the untreated controls is due to non-specific amplification.

4.4.2 Ectopic expression of NAM-B1 does not lead to premature senescence.

Having demonstrated that the heat shock-inducible construct does express *NAM-B1* following heat shock treatment, we then investigated whether ectopic expression of *NAM-B1* led to premature senescence. As *NAM-B1* is a positive regulator of senescence, we originally hypothesized that ectopic expression of *NAM-B1* early in development would lead to seedling death. Earlier attempts to recover positive transformants that constitutively over-expressed *NAM-B1* consistently failed (Uauy, personal communication) suggesting that ectopic expression of the gene was embryo lethal. To avoid this outcome, we developed the heat shock-inducible construct which would not express *NAM-B1* until plants were treated with a heat shock.

Using this construct, we then investigated whether induction of *NAM-B1* early in development led to premature senescence. Following application of heat shock, no significant earliness in senescence was observed in the plants expressing *NAM-B1* compared to those not expressing *NAM-B1* (Figure 4-9). This result was repeatable (Figure 4-11) and was complemented by a second transgenic line, constitutively expressing *NAM-B1*, that also showed no premature senescence (Figure 4-14). The total absence of any senescence phenotype was surprising at first, given that introgression of *NAM-B1* leads to premature senescence of approximately 3 days (Uauy et al. 2006a). Even if ectopic expression of *NAM-B1* were not sufficient to induce extremely early senescence immediately following heat shock, it would be expected to, at a minimum, recapitulate the phenotype observed when wild-type *NAM-B1* is introduced into a modern background.

The fact that we didn't see this premature senescence may indicate that the transgenic copy of *NAM-B1* is actually non-functional. We were unable to demonstrate that the NAM-B1 protein was being produced ectopically, as the transgenic construct does not introduce a tag to the protein. As a result, we could not carry out Western blot assays to demonstrate protein presence. It is therefore possible that, while the transgene was being expressed, the protein itself was not being translated or was being degraded post-translationally.

Alternatively, it's possible that under the controlled environment conditions used to phenotype the transgenic plants the subtle earliness phenotype observed in the field was not visible. We have found that, with other experiments, senescence phenotypes that were observed under glasshouse

conditions were not repeatable under controlled environment conditions (see Chapter 5). Previous work on recombinant inbred lines (RILs) containing a functional copy of *NAM-B1* showed no variation in senescence timing when grown in the glasshouse (Uauy, personal communication). The early senescence of approximately three days described above was only observed when studied in the field (Uauy et al. 2006a). Unfortunately, due to UK regulations on growth of genetically modified crops in the field we were unable to test these plants under field conditions during the span of this PhD.

While it is possible that the NAM-B1 protein was non-functional or not present, our results support an alternative hypothesis that, while expression of *NAM-B1* is necessary for the appropriate onset and progression of senescence, it is not sufficient to induce the developmental process. This suggests the possibility that one or more trans-factors may interact with NAM-B1 to co-regulate the onset of senescence. Given the tightly regulated nature of monocarpic senescence, and that its successful completion is essential for adequate grain filling, it is not surprising that other genes besides *NAM-B1* may have a role in ensuring the correct timing of senescence.

Based on our knowledge about *NAM-B1*, we can speculate about likely candidate genes which could act to co-regulate senescence. NAC transcription factors act as dimers when binding DNA, suggesting that perhaps *NAM-B1* may interact with other NAC transcription factors. Alternatively, *NAM-B1* may interact with other proteins from entirely different families—SlNAC1 in tomato (*Solanum lycopersicum*) is known to interact with a replication accessory protein of tomato leaf curl virus (Selth et al. 2005). Indeed, the C-terminal domains of NAC transcription factors, with their highly flexible, intrinsically disordered nature, have the potential to interact with various structurally-diverse partners (Jensen and Skriver 2014). The Arabidopsis NAC transcription factor ANAC019 demonstrates that these different forms of protein interactions need not be mutually exclusive. The NAC domain of ANAC019 physically interacts with various RING-H2 proteins (Greve et al. 2003), while ANAC019 itself interacts with the C-terminal domain of a different NAC transcription factor, CPL1 (Bang et al. 2008). In the following chapter, we will investigate potential candidate genes which may physically interact with *NAM-B1* and correspondingly regulate the onset and progression of monocarpic senescence.

4.4.3 The constitutive expression construct is associated with delayed senescence

Alongside the heat shock-inducible system discussed in detail above, we also developed a constitutive over-expression construct, expressing *NAM-B1* under the control of *pZmUbi*. As with the heat shock-inducible system, we found that expression of *NAM-B1* from this construct did not lead to any premature senescence. However, we found that transgenic plants constitutively over-expressing *NAM-B1* had significantly delayed peduncle senescence. This result contradicts the known function of *NAM-B1* as a positive regulator of senescence. Previous work has shown that

RNAi knock-down of *NAM-B1* and its homoeologs leads to significantly delayed senescence, while in the previous chapter we demonstrated that single mutations in *NAM-A1* lead to delayed senescence, particularly peduncle senescence.

One possible explanation for this result is that constitutive over-expression of *NAM-B1* led to silencing of both the transgenic construct and the related homoeologs (*NAM-A1* and *NAM-D1*) and orthologs (*NAM-A2*, *NAM-B2* and *NAM-D2*, on chromosome 2). This would explain the delayed peduncle senescence observed in the transgenic lines. However, we saw no evidence of consistently delayed flag leaf senescence, which was previously observed in lines silenced using RNAi (Uauy et al. 2006b). Additionally, we observed high levels of expression of the transgenic *NAM-B1* in the transgenic lines, which fails to support this hypothesis.

What, then, could be the cause of the extremely delayed peduncle senescence? It's important to note that, while on average flag leaf senescence was not delayed in the transgenic lines, there were individual lines which showed delayed flag leaf senescence. This was particularly clear in the SPAD data, where some individual plants in two of the transgenic lines showed nearly no loss of leaf greenness up to 60 days post heading (Figure 4-15). One possibility is that silencing of the *NAM* genes alongside the transgenic *NAM-B1* has only occurred in a subset of the individual T_1 plants. As we did not test expression levels of *NAM-B1* in all individual plants, we cannot say with certainty that this is not the case. We did carry out a preliminary correlation between the expression levels of *NAM-B1* and the days from heading to flag leaf and peduncle senescence and found no correlation between expression levels and senescence timing (r^2 of 0.03 and 0.18, respectively).

Due to time and space constraints, we were not able to repeat the experiment with the constitutive over-expression lines, instead choosing to focus on the heat-shock inducible lines. We are currently repeating the experiment with the constitutive over-expression lines and will investigate whether this delayed peduncle senescence response is reproducible. We will also sample individual plants both early in development and following flag leaf emergence to ascertain whether the native *NAM* genes and the *NAM-B1* transgene are being expressed as expected.

4.5 Summary

In this chapter, we have demonstrated the successful use of a heat shock-inducible construct that leads to constitutive activation of a gene of interest, in this case *NAM-B1*. This sets the stage for further use of this construct in wheat to support the fine-grained study of increasingly complex phenotypes. We then utilised this construct to investigate whether over-expression of *NAM-B1* at early stages of development would lead to premature senescence. Finding that it did not, we therefore hypothesize that *NAM-B1* is not sufficient to induce senescence on its own, and that it may require one or more interacting partners. In the next chapter we will explore candidate interacting partners of NAM-B1, and their role in regulating monocarpic senescence in wheat.

5 The NAC transcription factor *NAC-3* interacts with *NAM-B1* and is a positive regulator of senescence.

5.1 Introduction

In the previous chapter, we used an inducible transgenic system to show that over-expression of *NAM-B1* is not sufficient to induce premature senescence. This suggests that NAM-B1 may act in partnership with other proteins to control the onset of senescence in wheat. The next question, therefore, is what proteins might NAM-B1 interact with?

5.1.1 *NAC transcription factors form protein-protein interactions*

As we demonstrated in Chapter 3, NAM-B1 and its A-genome homoeolog NAM-A1 can physically interact with each other in yeast two-hybrid and co-immunoprecipitation assays. This corresponds with the known ability of NAC transcription factors to form both homo- and hetero-dimers with other NAC transcription factors (Olsen et al. 2005a). Initial research identified primarily instances of homodimerization, such as with the Arabidopsis NAC1 and ANAC019 transcription factors (Xie et al. 2000, Ernst et al. 2004). At a similar time, in *Brassica napus* the NAC transcription factor BnNAC14 was found to heterodimerise with multiple different NAC transcription factors in a yeast two-hybrid assay (Hegedus et al. 2003). Interestingly, this NAC transcription factor did not seem to form homodimers with itself suggesting that this is not a universal feature of NAC transcription factors. Two Arabidopsis NAC transcription factors, ANAC019 and ANAC092, were later shown to form a heterodimer using a competitive EMSA assay (Olsen et al. 2005a). Notably, this showed that ANAC019 could both homodimerize and form heterodimers, suggesting that the two forms of binding were not mutually exclusive. More recently, heterodimerisation between NAC transcription factors in rice was found to be the catalyst for movement of one partner from the endoplasmic reticulum into the nucleus (Mathew et al. 2016).

It is entirely possible, therefore, that NAM-B1 and the other NAM genes may also form heterodimers with unrelated NAC transcription factors that then allow the regulation of critical downstream senescence-associated genes. But it is not only heterodimer formation that may govern NAM-B1 action—NAC transcription factors are able to interact with non-NAC proteins through their C-terminal domain. This region, which is highly variable between NAC transcription factors, is intrinsically disordered, a property that makes it likely to form interactions with many different types of proteins. It has been suggested that the short motifs within the C-terminal domain (Ooka et al. 2003, Shen et al. 2009, Borrill et al. 2017) may act as molecular recognition features (MoRFs) that facilitate interactions with other proteins (Vacic et al. 2007, Jensen and Skriver 2014). These MoRFs are short sections of intrinsically disordered regions that are involved in protein recognition and binding. A particular feature of intrinsically-disordered regions is the ability to both bind many different proteins and to show highly plastic 3-dimensional conformations, present many different binding sites specific for different partners (Sun et al. 2013). The ability of the NAC C-terminal

domain to form protein interactions has been shown with the barley NAC HvNAC013, which can interact with the RCD1 protein via a specific motif, regulating plant senescence (Kjaersgaard et al. 2011, Kragelund et al. 2012).

5.1.1.1 The yeast two-hybrid technique allows identification of novel protein-protein interactions

Multiple techniques exist to study protein-protein interactions. Some allow the identification of interactions *in planta*, such as immunoprecipitation of a protein of interest, followed by mass spectrometry of the bound proteins (Domon and Aebersold 2006, Free et al. 2009). Techniques such as this have the benefit of identifying interactions that occur in the plant itself, as well as being able to identify multi-protein interaction complexes. However, they require either the introduction of a transgenic copy of the gene of interest, with a fused epitope tag to facilitate immunoprecipitation, or the use of an antibody raised against the protein of interest. As discussed in the previous chapter, wheat transformation is not trivial (Harwood 2011, Adamski et al. 2018), while the raising of native antibodies to wheat proteins is made difficult and costly in many cases due to the use of grass- and grain-eating animal hosts (Uauy, personal communication). At the same time, while these *in planta* techniques identify true protein interactions, they are only able to identify proteins which are interacting with the bait protein at the given moment of protein extraction. As a result, they provide a limited snapshot of the bait protein's interactome.

In contrast, *in vivo* techniques such as yeast two-hybrid allow the identification of a broad range of possible interacting proteins (Fields and Sternglanz 1994, Brückner et al. 2009). While the protein-protein interactions found through this system may not occur *in planta* (e.g. if the two genes in question are never simultaneously expressed), it provides a more global picture of possible interactions. An additional benefit to the use of this system is the high-throughput manner in which a library screen can be conducted (Brückner et al. 2009). After the development of a cDNA library from the tissues and timepoints of interest, the bait protein is screened against protein fragments encompassing a large portion of the sampled proteome. Putative protein interactions can then be validated using targeted yeast two-hybrid assays and co-immunoprecipitation experiments. One caveat of this system is that it can only identify pairwise protein interactions in the first instance and would need to be combined with other methods such as co-immunoprecipitation to study larger protein complexes (Brückner et al. 2009). However, in the case of NAC transcription factors which are known to form protein dimers, the yeast two-hybrid system has the potential to be highly informative.

5.1.2 *Advances in wheat genomics aid in the identification of candidate genes*

These examples demonstrate that there are many possible ways that NAM-B1 may be interacting with one or more proteins which are also required for regulating the onset of senescence. Following

identification of possible candidate genes, we can now use the new genomic and transcriptomic resources available for wheat to develop stronger hypotheses regarding gene function. Publicly-available RNA-Seq datasets have been integrated into the online expVIP wheat expression database (Borrill et al. 2016, Ramírez-González et al. 2018), with shared metadata terms allowing comparison between different samples from different experiments. It is now possible to take any gene of interest and rapidly identify patterns of gene expression across a wide variety of tissue, development, and stress conditions.

5.1.2.1 Gene co-expression networks

One such dataset is a developmental timecourse generated using the spring wheat variety Azhurnaya (Ramírez-González et al. 2018). This dataset consists of 209 samples covering a wide range of wheat development, from seedling stage (Zadoks 7) through whole plant senescence and grain ripening (Zadoks 90) (Zadoks et al. 1974). Across these developmental stages, many different tissues were sampled, ranging from spikelet components to the senescing flag leaf. The dataset provides fine-grained resolution of the gene expression patterns throughout a representative sample of wheat developmental stages and tissues. This, along with other publicly-available sets of RNA-Seq expression data, has been used to generate gene co-expression networks and transcription factor-target predictions that can further guide hypotheses surrounding gene function (Ramírez-González et al. 2018). Gene co-expression networks can be used to identify clusters of genes associated with specific tissues, developmental stages, or stresses. The metadata assigned to the independent RNA-Seq studies allowed the identification of gene expression clusters relevant to particular tissues—roots, leaves and shoots, grain, and spike—as well as for abiotic and biotic stresses.

5.1.2.2 Prediction of transcription factor targets

Of particularly importance for work on transcription factors is the GENIE3 network. This network was developed from the same RNA-Seq datasets used to form the gene co-expression networks. Using a Random Forest approach, the network predicts targets of transcription factors based on similarity in expression patterns (Huynh-Thu et al. 2010). The network produces a ranked list of all possible pairwise interactions between the set of transcription factors and the given gene universe. Initial tests of the GENIE3 algorithm found that it was able to predict previously-characterised regulatory networks from the simple *Escherichia coli* system (Huynh-Thu et al. 2010). More recently, the GENIE3 algorithm has been used in crop species such as wheat (Ramírez-González et al. 2018), maize (Huang et al. 2018), and sunflower (Marchand et al. 2014) to predict transcription factor targets in a variety of tissues and environmental conditions. For transcription factors included in the network, the predicted targets may provide an indication as to the role of the transcription factor. A transcription factor that predominantly targets genes involved in the hypersensitive response, for example, is likely to itself have a role in defence responses. However, as few genes have been characterised in wheat, most often little or no information on gene function

is associated with the predicted targets. Gene Ontology enrichment analyses can overcome these difficulties, suggesting processes that transcription factors may be involved in based on their predicted targets (Harrington et al. 2019a).

5.1.2.3 Bespoke gene expression datasets facilitate the study of wheat senescence

Alongside these large-scale analyses, specific experiments designed to investigate a single or small set of processes can also be informative. Recently, we published a senescence-specific RNA-Seq dataset that captured variation in gene expression from three to 26 days after anthesis (Borrill et al. 2019a). This dataset encompassed the progression of senescence until the beginning of visual senescence signs appear, providing insight into the genetic regulation of nutrient remobilisation and the induction of programmed cell death. During this research, we identified a set of candidate senescence regulatory genes, one of which was confirmed in wheat. We can now use these different sets of resources, from the large-scale developmental datasets to the senescence-specific expression data, to help us characterise the possible functions of new candidate senescence regulators. These functions can then be validated using germplasm such as the TILLING population, initially introduced in chapters 2 and 3 (Uauy et al. 2009, Krasileva et al. 2017), in which deleterious mutations in a gene of interest can be rapidly identified using the online *in-silico* TILLING platform on EnsemblPlants (Kersey et al. 2017).

5.1.3 *Aims and hypotheses*

Given the inability of *NAM-B1* overexpression to induce premature senescence, we hypothesized that NAM-B1 may physically interact with one or more proteins to regulate the onset and progression of senescence. In this chapter we used a yeast two-hybrid library screen to identify a set of NAC transcription factors which physically interact with NAM-B1. The role of these NAC transcription factors in regulating senescence was tested using mutants derived from the Kronos TILLING population, identifying one pair of homoeologs, the NAC3-1 genes, which are positive regulators of senescence. We then used the gene expression datasets to characterise their role in the onset and progression of senescence and utilised the GENIE3 network to predict targets of the NAC3-1 genes. Finally, we present the hypothesis that the NAM-1 proteins may interact with the NAC3-1 proteins to co-regulate monocarpic senescence in wheat.

5.2 Materials and Methods

5.2.1 *Y2H screen*

An ULTImate™ yeast two-hybrid (Y2H) screen was carried out by Hybrigenics Ltd. A custom cDNA library was prepared by Hybrigenics, using RNA extracted from wheat flag leaf blade, flag leaf sheath, peduncle, and grains sampled at 8, 14, and 21 days after anthesis (library reference TALPG). The cDNA of *NAM-B1* was used as the bait fragment (amino acids 1-208) with a single synonymous substitution in the third exon ([G/T] at position 1,011). This fragment encompassed

the entirety of the NAC domain and three additional amino acids. The complete cDNA sequence of *NAM-B1* is provided in Appendix 8.1.1.

Three independent ULTImate™ library screens were carried out, using three different proprietary bait vectors: pB27 (N-LexA-bait-C fusion), pB66 (N-GAL4_BD-bait-C fusion), and pB35 (N-GAL4_BD-bait-C fusion, inducible) where the bait is the *NAM-B1*(1-208) sequence. The ULTImate™ technology utilises a yeast mating process to facilitate screening of many interactions at once (Figure 5-1). In brief, bait vectors (those containing *NAM-B1*(1-208)) are transformed into yeast cells of sex type a. Concurrently, vectors containing the cDNA library fragments (prey) are transformed into yeast cells of sex type α. All prey peptides are expressed from the pP6 vector (N-GAL4_AD-prey-C fusion) and fused to the activation domain of GAL4. Following mating of the bait and library cells, only those yeast expressing opposite sex types (i.e. a and α) can progress to the diploid stage, co-expressing the bait and prey fusion proteins.

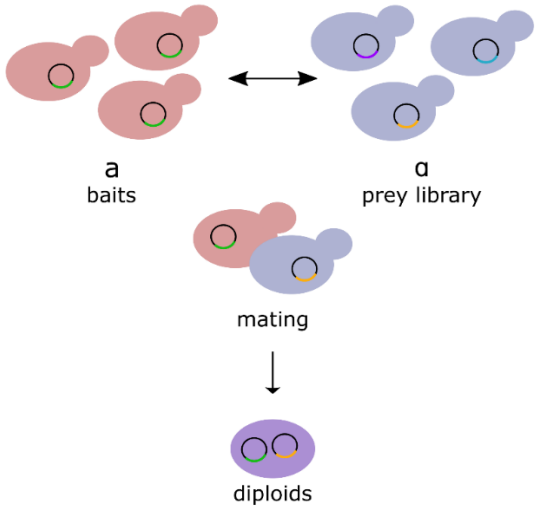

Figure 5-1: Outline of the Hybrigenics ULTImate™ Y2H screen. Adapted from Hybrigenics.

Following mating, interaction of the bait protein (NAM-B1) with the prey peptide restores yeast growth by bringing the LexA or GAL4 binding domain (BD), fused to the bait protein, into contact with the GAL4 activation domain (AD) fused to the prey peptide. This interaction then induces expression of the *HIS3* gene, under control of the *GAL4* promoter, recovering growth on Histidine-depleted media.

Identified interactions from the library screen were scored on a scale from A to D, where A refers to very high confidence interactions. These ranks are derived from the Predicted Biological Score (PBS), provided by Hybrigenics for each putative interaction. The PBS is based on the number of independent prey fragments that support the interaction between the bait and the prey, compared to the chance of identifying these interactions through random chance.

5.2.2 *Cloning of the Y2H candidate genes*

Following identification of candidate interacting partners of NAM-B1 (Table 5-1), the full-length sequences of each gene were identified in the TGACv1 gene models (Clavijo et al. 2017b), and later converted to the most recent RefSeqv1.1 gene models (IWGSC et al. 2018). All gene sequences are available in the Appendix 8.1.1 and 8.1.3. The gene sequences were synthesized by either Genewiz into the pUC57 vector, or by Life Technologies into the pMS-RQ vector.

Table 5-1: Candidate interacting partners of NAM-B1.

RefSeq ID	**Putative Role**
TraesCS2A02G101100	NAC TF
TraesCS2B02G118200	NAC TF
TraesCS2B02G118500	NAC TF
TraesCS5A02G049100	Salicylic Acid-induced NAC TF
TraesCS5B02G054200	Salicylic Acid-induced NAC TF
TraesCS6A02G108300	NAC TF (*NAM-A1*)
TraesCS3B02G371200	NAC TF

The sequences for *NAM-B1* and the identified candidate interacting proteins were cloned into pCR8 using TOPO® cloning, as explained in detail in Chapter 3, section 3.2.4. Briefly, the genes were amplified using the primers provided in Table 5-2 with the proofreading polymerase Q5, and then cloned into the Gateway vector pCR8. Both the full-length coding sequences and truncated versions of the genes were cloned, which contained only the NAC domain. Figure 5-4 highlights the premature truncation of the genes with a purple asterisk, and the short, NAC-domain sequences are also provided in Appendix 8.1.1 and 8.1.3.

Table 5-2: Primers used for the amplification of the full-length and NAC domain coding sequence of the NAM-B1 candidate interacting partners.

	Gene	**Forward**	**Reverse**
Full-length	TraesCS2A02G101100	ATGTCGGACGTGACGGC	TCAGATCTTCCACATGTCGGAGT
	TraesCS2B02G118200	ATGTCGGACGGTGCCG	TCAGATCTTCCACATGTTGGAGT
	TraesCS2B02G118500	ATGTCGGACGTGACGGC	TCAGATCTTCCACATGTTGGAGT
	TraesCS5A02G049100	ATGTCGATGAGCTTCCTGAGC	CTAGAAGTGGTTCCAAGTAGAGCTC
	TraesCS5B02G054200	ATGTCGATGAGCTTCCTGAGC	CTAGAAGTGGTTCCAAGTAGAGCTC
	TraesCS6A02G108300	ATGAGGTCCATGGGCAG	TCAGGGATTCCAGTTCACG
	TraesCS3B02G371200	ATGATGACGGCAATGGTGA	TCAGAATGGTGGCAAGATTGT
NAC domain	TraesCS2A02G101100	ATGTCGGACGTGACGGC	TCAGAAGGCGAGCGGGT
	TraesCS2B02G118200	ATGTCGGACGGTGCCG	TCACGCCGCCAAGTCTTTGTT
	TraesCS2B02G118500	ATGTCGGACGTGACGGC	TCAGGGCGGCGCATTCTT
	TraesCS5A02G049100	ATGTCGATGAGCTTCCTGAGC	CTAAATGTTGCTGAGGTTCCTTGG
	TraesCS5B02G054200	ATGTCGATGAGCTTCCTGAGC	CTAAATGTTGGTGAGGTTCCTTGG
	TraesCS6A02G108300	ATGAGGTCCATGGGCAG	TCACTGCTGATCTCCGG
	TraesCS3B02G371200	ATGATGACGGCAATGGTGA	TCACTCCACCTTGACCTTCT

5.2.3 *Y2H validation*

Validation of the protein-protein interactions identified in the Hybrigenics library screen was carried out using the ProQuest™ Two-Hybrid System, from Invitrogen. In this system, the low copy-number vectors pDEST22 and pDEST32 (Figure 5-2) are used for expression of the gene of interest with either the GAL4 DNA Binding Domain (BD; bait) or GAL4 Activation Domain (AD; prey) in a yeast strain that contains a set of reporter genes. In this case we used the yeast strain MaV203 which is recommended for use with the ProQuest™ system (Invitrogen, Cat. No. 11445-012). The genotype of the MaV203 strain is as follows:

MATa; leu2-3,112; trp1-901; his3Δ200; ade2-101; cyh2R; can1R; gal4Δ; gal80Δ; GAL1::lacZ; $HIS3_{UASGAL1}$::HIS3@LYS2; $SPAL10_{UASGAL1}$::URA3.

In brief, the MaV203 strains are auxotrophic for leucine and tryptophan, allowing selection for colonies transformed with pDEST22 (Trp) and pDEST32 (Leu). MaV203 contains three GAL4-inducible reporter genes including $HIS3_{UASGAL1}$*::HIS3* which induces the production of histidine upon interaction of the bait and prey proteins.

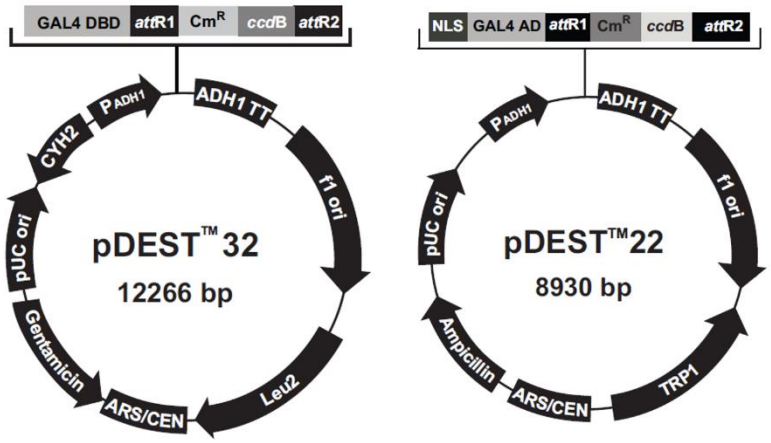

Figure 5-2: Diagram of the pDEST™32 and pDEST™22 vectors. pDEST™32 contains an GAL4 DNA Binding Domain, while pDEST™22 contains the GAL4 Activation Domain.

5.2.3.1 Cloning into pDEST™22 and pDEST™32

An LR reaction was used to transfer the gene sequences from the Gateway donor vector pCR8 (ENTR™) into both the pDEST™32 and pDEST™22 vectors (DEST™) (Figure 5-2), as detailed in Table 5-3.

Table 5-3: LR reaction conditions

Reagent	**1x (µL)**
ENTR™ vector (25 ng/µL)	1
DEST™ vector (50 ng/µL)	1
1X TE Buffer	2
LR Clonase™ II	1

The LR reaction (5 µL in total) was mixed gently by tapping then left to incubate at room temperature overnight. 0.5 µL of Proteinase K was then added to the reaction, mixed by tapping, and incubated at 37°C for 10 minutes to terminate the reaction. The completed reaction was then transformed into library competent DH5α as described in section 3.2.4 in Chapter 3. pDEST™22 reactions were plated on LB plates with Carbenicillin resistance (100 µg/mL), pDEST™32 reactions were plated on LB plates with Gentamicin resistance (10 µg/mL). Colonies were selected the next day and grown in cultures overnight with the appropriate antibiotics. (QIAprep Spin Miniprep Kit, ID: 27104). Miniprep concentration was quantified using a DeNovix Spectrophotometer (DS-11).

Correct insertion of the genes into pDEST™22 and pDEST™32 was validated by restriction digest, using the enzyme *Eco*RI which digests on either side of the gene insertion site in the *attB1* flanking regions. As a result, digestion with *Eco*RI produces a characteristic banding pattern for each vector (Figure 5-3). The digest was carried out as shown in Table 5-4.

Table 5-4: *Eco*RI Restriction Digest Reaction Conditions.

Reagent	**1x (µL)**
CutSmart Buffer (10X)	3
*Eco*RI (NEB HF)	1
DNA (500 ng/µL)	1
dH_2O	25

Following mixing, the reactions were incubated at 37°C for one hour and then run on a 1% agarose gel to visualise the resulting digest patterns (Figure 5-3).

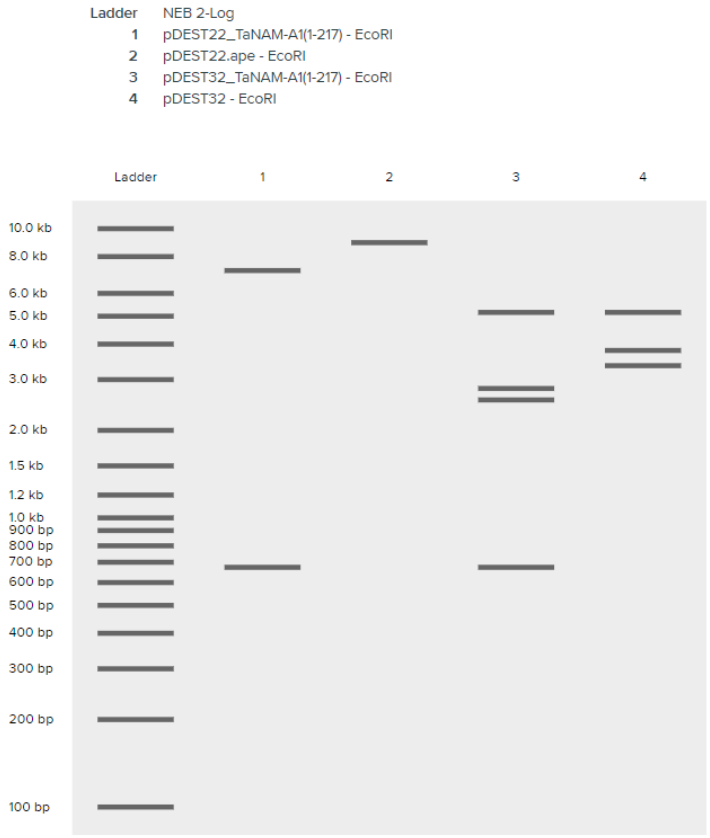

Figure 5-3: Restriction Fragments of pDEST™22 and pDEST™32. in silico digests of empty pDEST™22 and pDEST™32 (2 and 4) show distinct banding patterns compared to a digest of the same vectors with the NAM-A1(1-217) insert (1 and 3). Image was obtained from Benchling.com

5.2.3.2 Yeast transformation

Four transformations were carried out for each candidate gene (gene of interest; GOI) against the bait protein (i.e. NAM-B1):

1. pDEST™32 x pDEST™22_GOI
2. pDEST™32_NB1 x pDEST™22_GOI
3. pDEST™32_GOI x pDEST™22_NB1
4. pDEST™32_GOI x pDEST™22

The first and fourth transformations are controls to account for autoactivation by the GOI (1) and for baseline non-specific growth (4). The second transformation retests the original library screen, using NAM-B1 as the bait and the GOI as the prey. The third transformation then reverses this orientation, using the GOI as the bait and NAM-B1 as the prey.

Relevant control transformations were also produced:

1. None x None
2. pDEST™32 x pDEST™22
3. pDEST™32_NB1 x pDEST™22
4. pDEST™32 x pDEST™22_NB1

The yeast transformation protocol is as follows, modified from the ProQuest™ Two-Hybrid Small Scale Yeast Transformation Protocol:

Prepare competent MaV203 cells

1. Prepare an overnight culture of MaV203 in 10 mL of YPAD broth; shake at 30°C overnight.
2. The following day, dilute the culture to an OD_{600} of 1.0 in 50 mL of YPAD and grow for four hours at 30°C, shaking.
3. Transfer the 50 mL culture to sterile falcon tubes and spin down at 2500 rpm for 10 minutes.
4. Pour off the supernatant and resuspend the pellet in 40 mL 1X TE.
5. Pellet again at 2500 rpm for 10 minutes.
6. Pour of the supernatant and resuspend in 2 mL 1X LiAc/0.5X TE
7. Incubate resuspended cells at room temperate for 10 minutes. Transform immediately.

Transform competent MaV203

1. For each transformation, take 500 ng of the relevant pDEST™22 and pDEST™32 plasmids and combine in a sterile 1.5 mL Eppendorf tube.
2. Denature 100 µg of UltraPure™ Salmon Sperm DNA (Invitrogen, Cat. No. 15632011) per transformation (10 µL of 10mg/mL) at 95°C for 10 minutes. Add denatured DNA to each transformation.
3. Add 100 µL of competent MaV203 to each transformation; mix gently by tapping the side of the tube.
4. To each transformation, add 700 µL of 1X LiAc/40% PEG-3350/1X TE and mix by inverting tubes.
5. Incubate tubes at 30°C for 30 minutes.
6. Add 88 µL of fresh DMSO to each transformation; mix gently by inverting tubes.
7. Heat shock transformants at 42°C for 15 minutes.
8. Centrifuge briefly (~10 seconds at 8000 rpm) to pellet the cells; remove supernatant.
9. Resuspend cells in 1 mL 1X TE, and pellet again as above; remove supernatant.
10. Resuspend cells in 100 µL 1X TE, and plate 10 µL and 90 µL on two SC-LT plates.
11. Grow at 30°C for three nights (over the weekend).

5.2.3.2.1 *Recipes*

5.2.3.2.1.1 SC -Leu -Trp Plates

Synthetic Complete (SC) medium consists of a nitrogen base, a carbon source, and a "dropout" solution containing essential amino acids, nucleic acids, trace elements and vitamins. For selection purposes, certain amino acids are omitted or "dropped out" (e.g., leucine, tryptophan) from the dropout solution. SC plates for yeast growth and selection are made in the following way:

1. Prepare the following solutions:

    40% glucose

    100 M uracil

    100 mM histidine-HCl

2. Autoclave the 40% glucose and filter sterilize the amino acids. Store the amino acids in the dark or wrapped in foil.
3. The liquid medium and the agar are autoclaved in two separate 2 L flasks. For 2 L of medium, to one flask add 13.4 g of Yeast Nitrogen Base without Amino Acids (Formedium, CYN0401) and 2.8 g of SC Quadruple Dropout, -His, -Leu, -Trp, -Ura (Formedium, DSCK282)
4. Add a clean stir bar, suspend in 1 L distilled water and adjust the pH to 5.9 with NaOH. Keep the stir bar in the flask to stir the medium after autoclaving.
5. To the second flask add 40 g of agar in 900 ml distilled water. The agar will be solubilized during autoclaving.
6. Autoclave both flasks for 20 minutes to sterilise.
7. After autoclaving, pour the contents of the flask containing agar into the flask containing medium. Cool in a 50°C water bath for about 1 hour. Add 100 ml autoclaved 40% glucose.
8. For SC -Leu, -Trp plates, add 16 mL of 100 mM Histidine-HCl and 16 mL of 20 mM uracil to the 2 L of mixture.

5.2.3.2.1.2 10X TE

100 mM Tris-HCl, 10mM EDTA, pH 7.5; sterilise by autoclaving.

5.2.3.2.1.3 10X LiAc

1M Lithium Acetate; filter sterilise.

5.2.3.2.1.4 1X LiAc/0.5X TE

Combine 10mL 10X LiAc, 5 mL 10X TE, and 85 mL sterile dH_2O. Filter sterilise.

5.2.3.2.1.5 1X LiAc/40% PEG-3350/1X TE

Combine 10 mL 10X LiAc, 5 mL 10X TE and 40g PEG-3350; make up to 100 mL with sterile dH_2O. Filter sterilise.

5.2.4 *Co-Immunoprecipitation validation*

To test the ability of proteins to interact *in planta*, we carried out Co-Immunoprecipitations (Co-IPs) using plant tissue from *Nicotiana benthamiana* which was over-expressing the desired genes of interest.

5.2.4.1 Cloning into pGWB12 and pGWB21

An LR reaction was used to clone the full-length gene sequences into the desired Gateway Binary Vectors, as detailed in 5.2.3.1. The full-length sequence of *NAM-B1* was cloned into pGWB12, containing an N-terminal FLAG tag under the control of the 35s promoter. The full-length sequences of the candidate genes (Table 5-1) were cloned into pGWB21, containing an N-terminal 10x cMyc tag under the control of the 35s promoter. The sequences of both binary vectors are available from Nakagawa et al. 2007.

5.2.4.2 Protein expression in *Nicotiana benthamiana*

5.2.4.2.1 *Transformation into Agrobacterium tumefaciens.*

The cloned genes in pGWB12 and pGWB21 were transformed into *A. tumefaciens* strain LBA4404 following a standard protocol for electrically-competent cells. Briefly, for each transformation 1 μL of miniprepped DNA was added to a single aliquot (40 μL) of *A. tumefaciens*. After incubation on ice, the cells were transferred to an electroporation cuvette. Electroporation was carried out at 2.5 kV, 25 μFD, and 400 Ω. Cells were immediately resuspended in 1 mL of SOC medium and allowed to recover at 28°C for no less than three hours. Following recovery, cells were pelleted and resuspended in 100 μL of dH_2O before plating on selective plates, LB with Rifampicin (50 μg/mL) and Kanamycin (100 μg/mL).

5.2.4.2.2 *Infiltration into* N. benthamiana

Transformed *A. tumefaciens* colonies were picked and cultured in LB broth with Rifampicin (50 μg/mL) and Kanamycin (100 μg/mL). After reaching turbidity (approximately 24-48 hrs), the OD_{600} of the cultures was measured. Cultures were then pelleted and resuspended in 5 mL of MMA (Table 5-5) and incubated at room temperature in the dark for at least one hour. Following incubation, the desired combinations of constructs were combined and diluted to reach a final OD_{600} of 0.2 for each construct in the co-infiltration mixture. For the co-immunoprecipitation experiment, each co-infiltration mixture included the candidate gene, *NAM-B1*, and the silencing inhibitor P19.

Table 5-5: Preparation of MMA solution for *N. benthamiana* infiltration.

MMA solution	Final Concentration	Stock solution	100 mL
MgCl2	10 mM	1 M	1 mL
MES/KOH pH 5.6	10 mM	1 M	1 mL
Acetosyringone	150 µM	150 mM	100 µL
dH_2O	X	X	97.9 mL

The prepared cultures were then infiltrated into *N. benthamiana* leaves using the standard needle-less syringe method as detailed in (Santi et al. 2008) with the aim of infiltrating the entire leaf. Approximately three-week-old plants were used for infiltration, and plants were left for three days before harvesting. Leaves were snap-frozen in liquid N_2 in preparation for protein extraction.

5.2.4.2.3 *Protein Extraction*

Snap-frozen leaves were ground in liquid N_2 using a mortar and pestle. Ice-cold extraction buffer (100 mM Tris-HCl, pH 7.5; 150 mM NaCl; 0.1% (v/v) Triton x-100; 1% (w/v) PVPP; 5 mM EDTA; 10% glycerol; 2 mM DTT; 1% protease inhibitors (Sigma)) was added to the ground tissue, at a ratio of 5 mL per 2 g of tissue. The mixture was centrifuged at 5000 rpm for 15 minutes, at 4°C, and the supernatant containing the extracted proteins was retained.

5.2.4.2.4 *Co-immunoprecipitation and western blot*

1.5 mL of extracted protein ("input") was added to 20 µL of Anti-FLAG® M2 Magnetic Beads (Sigma, cat. no. M8823) in a LoBind Eppendorf tube (cat. no. 10031282). The beads were previously washed two times in tris-buffered saline (TBS). The protein extract and the Anti-FLAG® beads were incubated at 4°C, shaking for one hour. The samples were then briefly centrifuged before being placed on a magnetic rack to separate the beads from the protein extract. A sample of the protein extract ("unbound") was taken, and the remainder discarded. The beads were then washed three times in TBS before being resuspended in 50 µL of 1X SDS sample loading buffer (Table 5-6). The appropriate amount of 5X SDS sample loading buffer was added to the unbound sample. Both samples were boiled for ten minutes.

Table 5-6: Components of 9.5 mL 5X SDS loading buffer.

Reagent	**Volume (mL)**
1M Tris/HCl (pH 6.8)	0.6
50% glycerol	5
10% SDS	2
1% bromophenol blue	1
dH_2O	0.9

Samples were then loaded onto 12% acrylamide gels, with 5% acrylamide stacking gels. Two identical gels were loaded for each experiment. They were run in 1X Running Buffer (3 g Tris, 14.4 g Glycine, 1 g SDS in 1 L dH_2O at pH 8.3) for 20 minutes at 70 V, followed by a further hour at 100 V. The Blue Protein Standard, Broad Range (NEB, cat. no. P7786) was loaded alongside to visualise the size of the protein bands. The gels were then transferred into a wet transfer cassette and transferred onto nitrocellulose membrane in 1X transfer buffer (1.56 mM Tris, 12mM Glycine) at 95mA for one hour and fifteen minutes. The membranes containing the transferred protein were then blocked in 5% milk and TBS for 30 minutes to one hour. One membrane was then incubated in an anti-FLAG HRP-conjugated antibody (Abcam, cat. no. AB49763) and the other in an anti-cMyc HRP-conjugated antibody (Abcam, cat. no. AB62928) overnight at 4°C, shaking.

Following antibody incubation, the membranes were washed once with TBS, twice with TBS + 0.05% Tween 20, and once more with TBS alone. They were then developed using the Pierce™ ECL Western Blotting Substrate (Thermo Scientific, cat. no. 32109) according to the manufacturer's instructions. Membranes were exposed using an ImageQuant LAS 500 chemiluminescence CCD camera (GE Healthcare Life Sciences, cat. no. 29005063).

5.2.5 *Cell death screening in* Nicotiana benthamiana

To test whether the yeast two-hybrid candidate genes and the TraesCS2A02G101400 and TraesCS2B02G118500 alleles could induce cell death, we utilised the transient expression system in *Nicotiana benthamiana*. This system was described in detail in section 3.2.7 of Chapter 3. In brief, the full-length wild-type and mutant alleles of the genes of interest were cloned into pGWB21, containing an N-terminal 10X cMyc tag. The wild-type allele of *NAM-B1* was cloned into pGWB12, with an N-terminal FLAG tag. The constructs were transformed into *A. tumefaciens*, and then infiltrated into *N. benthamiana* following the patch infiltration method depicted in Figure 3-2, p.103. The level of cell death induced by the genes and the mutated alleles was scored on a scale of 0-6, where 6 corresponds to high cell death and 0 to no cell death. This follows from the scale established previously in Maqbool et al. 2015 and reproduced as Figure 3-3, p.103.

5.2.6 *Plant materials and KASP*

All wheat lines used in this study were obtained from the Kronos TILLING collection, as first described in Krasileva et al. Seeds were obtained at the M_4 or M_5 generation and genotyped using KASP markers (Table 5-7) as detailed in section 2.2.6, Chapter 2.

Table 5-7: KASP markers used for yeast two-hybrid candidate gene TILLING mutants.

Primer	Gene	Sequence (5′-3′)
K2132_WT	TraesCS5B02G054200	GAAGGTCGGAGTCAACGGATTgaaccggaaccccggcg
K2132_Mut	TraesCS5B02G054200	GAAGGTGACCAAGTTCATGCTgaaccggaaccccggca
K2132_Com	TraesCS5B02G054200	ctctctcgtcgcatcag
K2825_WT	TraesCS5A02G049100	GAAGGTCGGAGTCAACGGATTaaccgagctactgtgtcgg
K2825_Mut	TraesCS5A02G049100	GAAGGTGACCAAGTTCATGCTaaccgagctactgtgtcga
K2825_Com	TraesCS5A02G049100	caccctatccttcccggtt
K3964_WT	TraesCS5A02G049100	GAAGGTCGGAGTCAACGGATTgtgggtggcaaggagtgg
K3964_Mut	TraesCS5A02G049100	GAAGGTGACCAAGTTCATGCTgtgggtggcaaggagtga
K3964_Com	TraesCS5A02G049100	gcatatttccgatcctttaggcta
K3842_WT	TraesCS2A02G101400	GAAGGTCGGAGTCAACGGATTacccctttctccgtggcgc
K3842_Mut	TraesCS2A02G101400	GAAGGTGACCAAGTTCATGCTacccctttctccgtggcgt
K3842_Com	TraesCS2A02G101400	aaggatcgcaagtacccgac
K2236_WT	TraesCS2A02G101400	GAAGGTCGGAGTCAACGGATTCgagccgtgggatctccc
K2236_Mut	TraesCS2A02G101400	GAAGGTGACCAAGTTCATGCTCgagccgtgggatctcct
K2236_Com	TraesCS2A02G101400	aacggacgctcctctcgc
K4485_WT	TraesCS2B02G118500	GAAGGTCGGAGTCAACGGATTcttccggttccaccccac
K4485_Mut	TraesCS2B02G118500	GAAGGTGACCAAGTTCATGCTcttccggttccaccccat
K4485_Com	TraesCS2B02G118500	ttgggggtgaggtagtgc
K4681_WT	TraesCS2B02G118500	GAAGGTCGGAGTCAACGGATTagtagtactcaccggggagatcc
K4681_Mut	TraesCS2B02G118500	GAAGGTGACCAAGTTCATGCTagtagtactcaccggggagatct
K4681_Com	TraesCS2B02G118500	cgtcgacctcaacaagacg
K0996_WT	TraesCS2A02G101100	GAAGGTCGGAGTCAACGGATTaagggtgccgccgcgatg
K0996_Mut	TraesCS2A02G101100	GAAGGTGACCAAGTTCATGCTaagggtgccgccgcgata
K0996_Com	TraesCS2A02G101100	gaagcgccagatccatcac
K1308_WT	TraesCS2A02G101100	GAAGGTCGGAGTCAACGGATTctcaacaagaacgagccgtg
K1308_Mut	TraesCS2A02G101100	GAAGGTGACCAAGTTCATGCTctcaacaagaacgagccgta
K1308_Com	TraesCS2A02G101100	cctgcaaatccaatcagatcagac
K2169_WT	TraesCS2B02G118200	GAAGGTCGGAGTCAACGGATTccttccccgtcgccttcc
K2169_Mut	TraesCS2B02G118200	GAAGGTGACCAAGTTCATGCTccttccccgtcgccttct
K2169_Com	TraesCS2B02G118200	aggcctcgcgaagatggg

5.2.7 *Plant growth and phenotyping*

5.2.7.1 Plant growth conditions

In this chapter, all plants were grown under either standard glasshouse conditions or controlled environment conditions. Greenhouse-grown plants were grown under lights for 16 hours of the day, with the remainder determined by the current photoperiod. Plants grown in controlled environment conditions experienced cycles of 16 hours of light at 20°C and 8 hours of dark at 15°C. Seedlings were germinated in the dark at 4°C on damp filter paper for 48-72 hours. They were then recovered at room temperature for 24 hours before sowing into 85% fine peat and 15% horticultural grit in P96 trays. At the three-leaf stage, seedlings were transplanted into 1 L pots containing Petersfield Cereal Mix (Petersfield, Leicester, UK).

5.2.7.2 Visual senescence scoring

Flag leaf and peduncle senescence was scored visually on the primary tiller of each plant, tagged upon heading. Flag leaves were scored for senescence onset when 25% of the leaf appeared yellow and necrotic. Peduncle senescence was scored when the top 3 cm of the peduncle were completely yellow. In all cases, visual senescence scores are reported as days from heading, thus accounting for delayed senescence due solely to delayed development overall.

5.2.7.3 SPAD scoring

SPAD measurements were also taken on the main flag leaves as a measurement of leaf greenness, and thus as a proxy for leaf chlorophyll content. Measurements were typically begun around heading and carried out at intervals thereafter until complete senescence of the flag leaf had occurred. The SPAD-502 chlorophyll meter (Konica-Minolta) was used to take eight measurements of each flag leaf, four on either side of the mid-vein, which were then averaged to produce a single SPAD value for each flag leaf.

5.2.8 *Expression analysis of candidate genes*

5.2.8.1 Developmental expression datasets

Public datasets were used for all expression analysis of candidate genes. The developmental time course was derived from the spring wheat cultivar Azhurnaya and includes 209 samples covering developmental timepoints from early seedling stage through late senescence (Ramírez-González et al. 2018). A second dataset captures senescence-specific changes in expression in the mature flag leaf (Borrill et al. 2019a). This dataset was derived from the spring wheat cultivar Bobwhite. The developmental data set was included in the GENIE3 network, while the flag-leaf senescence dataset was not (Ramírez-González et al. 2018).

5.2.9 *Phylogenetic analysis of the NAC-3 Clade*

5.2.9.1 Cladogram of NAC3 proteins

Protein sequences of the *T. aestivum* genes within the NAC3 clade were aligned using the online Clustal Omega tool (Sievers et al. 2011) (Figure 5-22). A sequence-based phylogeny was constructed from the multiple sequence alignment in Jalview (v2.10.5). The PAM250 matrix (Schwartz and Dayhoff 1978) was used to calculate sequence similarity between the 11 protein sequences, and a phylogenetic tree was calculated using the Neighbour Joining method (Saitou and Nei 1987).

5.2.9.2 Phylogeny of NAC3 clade in the Triticeae

The EnsemblPlants Compara Gene Tree (Vilella et al. 2009) was extracted for one of the wheat NAC3 genes, TraesCS2B02G118500, including the sequenced species within the Triticeae- *Triticum aestivum*, *Triticum dicoccoides*, *Triticum urartu*, *Aegilops tauschii*, *Hordeum vulgare*, and *Brachypodium distachyon*. The sequenced members of the *Oryza* genus were also included along with the *Oryza* outgroup, the wild grass *Leersia perrieri*. The phylogeny was then visualised in FigTree (v 1.4.2) using a radial tree layout.

5.2.9.3 Synteny analysis of the NAC3 clade

Positional information for the NAC3 clade and its orthologs in the *Triticinae* was extracted from EnsemblPlants (Kersey et al. 2017) for *T. aestivum*, *T. dicoccoides*, *A. tauschii*, and *H. vulgare*. We also manually curated the genomes for other NAC transcription factors located near to the orthologs identified in the phylogeny, identifying a single additional NAC3 ortholog in the *A. tauschii* genome. We also manually curated the list of NAC3 orthologs in *T. urartu*, using the more complete genome released in 2018 that had not yet been integrated into EnsemblPlants (Ling et al. 2018), identifying five members of the NAC3 clade.

5.2.10 *Analysis of the GENIE3 network*

The wheat GENIE3 network, as described initially in Ramírez-González et al. (2018), was interrogated as detailed in Harrington et al. (2019a). Briefly, the network was generated using the GENIE3 R package on 850 independent RNA-Seq samples, predicting targets of the 3,384 annotated transcription factors in the RefSeqv1.0 gene annotation (IWGSC et al. 2018). All further analyses considered only the top 1 million connections, consistent with previous GENIE3 analyses (Huang et al. 2018). We used a series of bespoke R scripts to access and analyse the GENIE3 network, all of which are available at https://github.com/Uauy-Lab/GENIE3_scripts/ and in Appendix 8.2.2.

5.2.10.1 Comparison of the GENIE3 network with the nam-a1 mutants

5.2.10.1.1 *Pseudoalignment and differentially-expressed gene identification*

As previously described in Harrington et al. (2019a), publicly-available reads from Pearce et al. (2014) were accessed from https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE60635. We

pseudoaligned reads from the wild-type and *nam-a1* mutants at 12 and 22 days after anthesis (DAA) against the RefSeqv1.1 transcriptome (IWGSC et al. 2018), using kallisto (v 0.43.1) (Bray et al. 2016) under standard settings for single-read datasets (--single -b 30 -l 200 -s 20). Reads were only pseudoaligned to the A and B genomes as they are derived from the tetraploid cultivar Kronos. The R package Sleuth (Pimentel et al. 2017) was used to determine differentially expressed genes (DEGs) between the wild-type and *nam-a1* genotypes at each timepoint, using default settings for the Walk test (sleuth_wt; v 0.30.0). The Benjamini-Hochberg adjusted p-value, or *q-value*, threshold was set at < 0.05 when identifying DEGs.

5.2.10.1.2 *Calculation of Shared Ratios*

To provide a comparable metric for the level of overlap between any two given gene lists, we used the "shared ratio" approach, which accounts for the size of each gene list. The shared ratio was calculated as follows:

$$\frac{|A \cap B|}{|A|}$$

Where A and B are sets of genes, and |A| < |B|.

For example, given two sets of genes A and B, where A contains 5 genes and B contains 10 genes, if they share two genes in common the shared ratio is 2/5, or 0.4. This calculation was implemented in various R scripts available in https://github.com/Uauy-Lab/GENIE3_scripts/ and as detailed in Appendix 8.2.2.

5.2.10.1.3 *Comparison between DEGs and GENIE3 targets*

We calculated a distribution of shared ratios for the predicted targets of 1000 randomly-selected transcription factors present in the GENIE3 network and the list of DEGs at each timepoint (12 and 22 DAA). We then calculated the shared ratio between the DEGs at each timepoint in the RNA-Seq experiment (12 and 22 DAA) and the predicted targets of *NAM-A1* (see Figure 5-24). This analysis was implemented in https://github.com/Uauy-Lab/GENIE3_scripts/Genie3_Statistics_SharedRatios_RNASeqDEGs_Figure1.Rmd and is reproduced in Appendix 8.2.2.

5.2.10.2 Identification of novel senescence candidate genes

We also leveraged the GENIE3 network to identify a set of candidate senescence regulatory genes which had higher shared ratios between their predicted targets and the *nam-a1* DEGs than *NAM-A1* itself (Figure 5-27). Expression analysis of the candidates was carried out as detailed in section 5.2.8.

5.2.10.3 Gene Ontology (GO) term analysis

GO term enrichment analysis was carried out for the GENIE3-predicted targets of the *NAC3-1* triad. The analysis was carried out as previously described in Borrill *et al*. (2019), using the GOSeq (v 1.34.1) package in R (Young et al. 2010).

5.2.11 *Data visualisation, manipulation, and statistical analyses.*

All graphs were made using R, principally with the ggplot2 (v 3.1.1) (Wickham 2016) and ggpubr (v 0.2) (Kassambara 2019) packages as well as the "aheatmap" function of the NMF package (v 0.21.0) (Gaujoux and Seoighe 2010). Networks in Figure 5-25 and Figure 5-28 were visualised using Cytoscape (v 3.7.1) (Shannon et al. 2003). Data manipulation and statistical analysis was also carried out in R, using the packages dplyr (v 0.8.0.1) (Wickham et al. 2019) and tidyr (v 0.8.3) (Wickham and Henry 2018).

5.3 Results

5.3.1 *7 NAC Transcription Factors Bind NAM-B1 in vitro*

5.3.1.1 A yeast two-hybrid library screen identified candidate interacting proteins with NAM-B1

To test whether NAM-B1 was able to interact with other proteins, we carried out a yeast two-hybrid (Y2H) screen. A cDNA library was constructed by Hybrigenics from RNA extracted from leaf, peduncle, and grain tissue. An initial attempt at the Y2H screen found that use of the full NAM-B1 sequence led to auto-activation, that is, growth of the yeast on selective media without the presence of the pDEST^{TM}32 construct containing the GAL4 activation domain. Following this, we tested the Y2H screen using solely the NAC domain of NAM-B1, amino acids 1-215, which showed no signs of auto-activation. This region consists of the NAC domain to the canonical "CRVFKK" motif at the end of subdomain v, along with 10 further amino acids. Much of the C-terminal domain is removed, preventing its action as a transcriptional activator (Jensen et al. 2010).

The Y2H screen identified a set of 13 genes with varying levels of confidence in their interaction (Table 5-8). Only two interactions were initially identified as high quality, with a Predicted Biological Score (PBS) score of "A", based on the number of independent prey fragments (derived from the cDNA library), as detailed in section 5.2.1. A further 9 genes were given a PBS score of "D", indicating that they were supported by few interactions and therefore may be either true binding partners or false positives. Two of the putative interactions were not given a PBS score ("N/A"), due to the clone sequence only comprising the 5′ untranslated region (UTR) of a gene (TraesCS6A02G221700) or where the 5′ sequence is missing (TraesCS2D02G241300).

Table 5-8: Potential interaction partners of NAM-B1, as identified through a yeast two-hybrid library screen from Hybrigenics. NAC transcription factors are highlighted in light blue. The RefSeqv1.1 gene ID (IWGSC et al. 2018) was derived from the best BLAST hit of the identified cDNA fragments against the v1.1 transcriptome. The number of supporting interactions refers to the number of independent cDNA clones which contained a fragment of the gene sequence and bound to NAM-B1. The putative gene role was derived from EnsemblPlants (Kersey et al. 2017) and Uniprot (Consortium 2018), while the Arabidopsis Orthologue and Gene ID were also obtained from EnsemblPlants (Kersey et al. 2017).

RefSeq ID	**Global PBS Score**	**Number of supporting interactions**	**Putative Role**	**Arabidopsis Orthologue**	**Arabidopsis Gene ID**
TraesCS2A02G101100	D	2	NAC TF	ANAC046/ANAC087	AT3G04060/AT5G18270
TraesCS2B02G118200	D	2	NAC TF	ANAC046/ANAC087	AT3G04060/AT5G18270
TraesCS2B02G118500	D	6	NAC TF	ANAC046/ANAC087	AT3G04060/AT5G18270
TraesCS2B02G233300	A	12	photosystem II reaction center PSB29 protein	N/A	N/A
TraesCS2B02G620600	A	13	unknown, has DUF IPR022203	N/A	AT4G33480
TraesCS2D02G241300	N/A	1	nonsense-mediated mRNA decay	UPF1	AT5G47010
TraesCS3B02G371200	D	1	NAC TF	N/A	N/A
TraesCS4B02G087400	D	2	F-box protein	N/A	N/A
TraesCS5A02G049100	D	2	Salicylic Acid-induced NAC TF	ANAC021	AT1G56010/AT3G12977
TraesCS5B02G054200	D	2	Salicylic Acid-induced NAC TF	ANAC021	AT1G56010/AT3G12977
TraesCS6A02G049800	D	1	Alpha/beta-gliadin A-I (prolamin)	N/A	N/A
TraesCS6A02G108300	D	1	NAC TF (*NAM-A1*)	NAC018/NAC025/NAC056	AT1G52880/AT1G61110/AT3G15510
TraesCS6A02G221700	N/A	1	nonsense mediated mRNA decay	UPF2	AT2G39260

The report from the Y2H screen included the corresponding cDNA sequences for each of the putative interacting proteins. These sequences were then compared using a BLAST analysis (Altschul et al. 1990) originally against the IWGSC CSS genome (IWGSC 2014), and later against the RefSeqv1.0 genome (IWGSC et al. 2018). The best BLAST hits for each cDNA fragment are reported in Table 5-8. We identified seven NAC transcription factors as potential interactors with NAM-B1, including its A genome homoeolog NAM-A1 (TraesCS6A02G108300). Three of the NAC transcription factors are orthologous to the Arabidopsis NAC transcription factors ANAC046 and ANAC087, which are known to be positive regulators of senescence and cell death (Oda-Yamamizo et al. 2016, Huysmans et al. 2018). Two other NAC transcription factors are orthologous to ANAC021, a senescence-associated transcription factor which is upregulated in senescence (Balazadeh et al. 2008). The *Zea mays* ortholog of these genes is NAC108 (Zm00001d041472) which was recently found to be a positive regulator of senescence, with its downregulation leading to a stay-green phenotype and increased grain yield in maize (Zhang et al. 2019). Of the remaining candidate interactors, the two "A" candidates included a protein involved in the photosystem II reaction centre (PSB29) (Kashino et al. 2002) and a protein containing a domain of unknown function. Based on previous work which has demonstrated that NAC transcription factors act as dimers to bind DNA and can form heterodimers with other NAC transcription factors (Olsen et al. 2005a), we focussed further work on the NAC transcription factors present in the list of candidates.

The NAC transcription factors all shared a full, canonical NAC domain with the five highly conserved subdomains previously introduced in Chapter 3. Unrelated NAC transcription factors typically have highly divergent C-terminal domain (CTD) regions, which can act as transcriptional activators or repressors (Xie et al. 2000, Tran et al. 2004, Olsen et al. 2005b, Jensen et al. 2010, Yamaguchi et al. 2010). By aligning the NAC transcription factors, we can see that the two NACs on chromosome 5, TraesCS5A02G049100 and TraesCS5B02G054200, have very similar CTDs (Figure 5-4). Overall, these two proteins share 98.04% identity suggesting that they are homoeologous genes. The three NACs on chromosome 2, TraesCS2A02G101100, TraesCS2B02G118200, and TraesCS2B02G118500, also all show high levels of similarity in protein sequence. Across the full protein sequence, they share up to 86% pairwise identity. This is substantially higher than that between two unrelated NACs, such as *NAM-A1* (TraesCS6A02G108300) and TraesCS3B02G371200 which share only 31% sequence identity, but less than would be expected between true homoeologs which typically share between 92% and 100% protein sequence identity (Ramírez-González et al. 2018). This suggests that, while the three NACs on chromosome 2 are closely related, they are likely not true homoeologs.

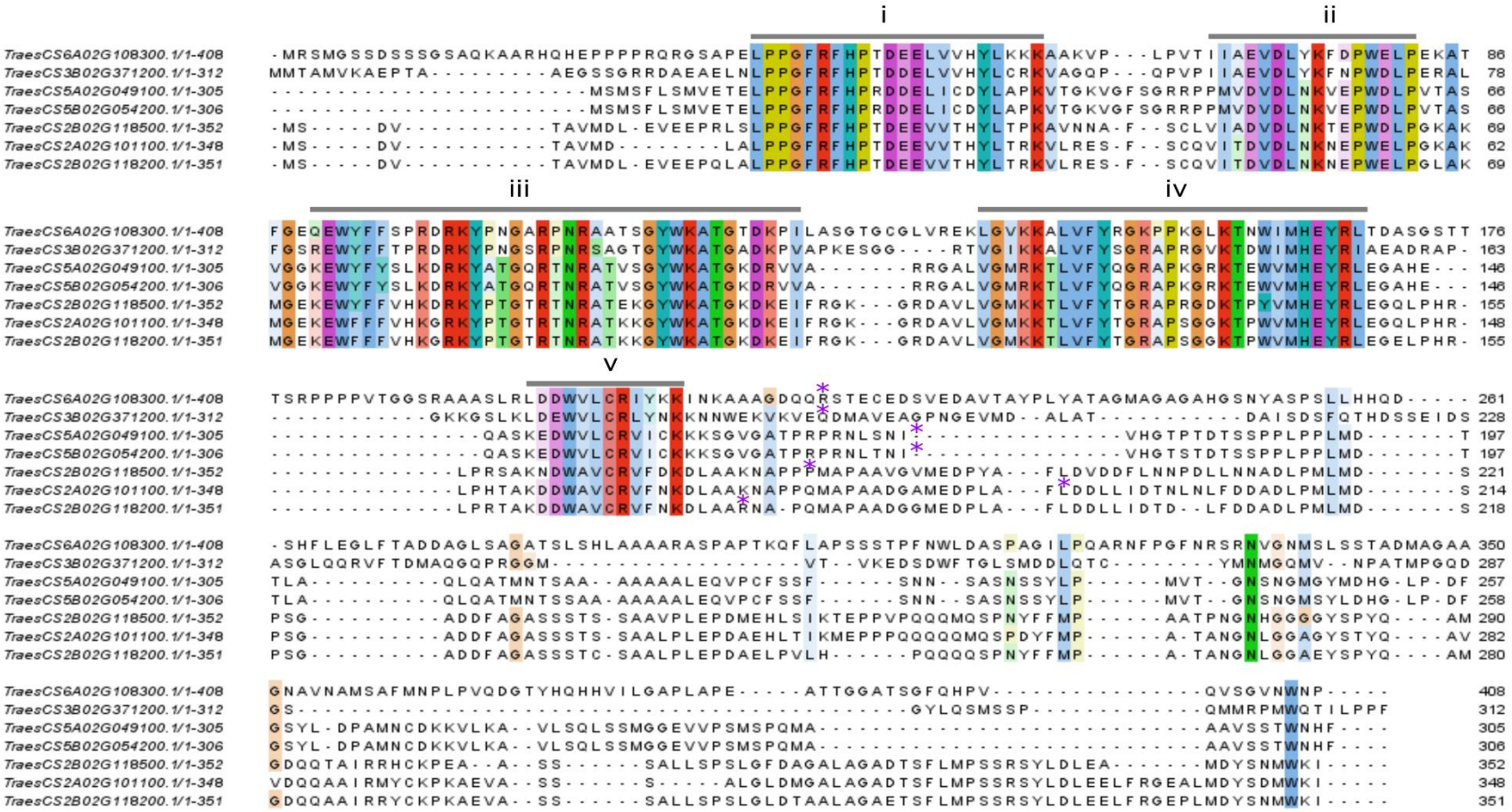

Figure 5-4: Protein alignment of the candidate NAC transcription factors. Residues are coloured using the CLUSTAL colour scheme, with colours only shown at residues with a conservation score ≥ 40. Purple asterisks indicate the premature termination of the NAC domain used in the Y2H experiments. NAC subdomains are numbered from i to v and indicated with a solid grey line.

5.3.1.2 Validation of interactions using Y2H

To determine which putative interacting proteins were true positive interactions, we had the full cDNA sequence of the corresponding genes synthesized. These cDNAs were then cloned into the Y2H vectors pDEST22 and pDEST32 which contain the GAL4 DNA binding domain and activation domain, respectively (Figure 5-2). Initial tests with the candidate NAC transcription factors found that all full-length cDNAs led to auto-activation of the GAL4 promoter, indicating that the C-terminal domain had transcriptional activation activity (data not shown). We therefore used only the NAC domain for all further Y2H validation. The amino acid at which the NAC domain was truncated for each transcription factor is indicated with a purple asterisk in Figure 5-4.

We were able to repeat and validate the interactions between NAM-B1 and five of the candidate genes, all of which were NAC transcription factors (Figure 5-5). The three genes orthologous to *ANAC046*/*ANAC087* supported growth on the selective media when fused to the GAL4 activation domain and co-expressed with NAM-B1 fused to the GAL4 binding domain. One of the genes orthologous to *NAC021*, TraesCS5B02G054200, showed growth on the selective media when fused to either the binding or activation domain of GAL4 and co-expressed with NAM-B1. However, the other orthologue, TraesCS5A02G049100, did not appear to bind NAM-B1 in the Y2H system. Auto-activation was observed for TraesCS3B02G371200, despite the removal of the C-terminal domain. Finally, NAM-A1 (TraesCS6A02G108300) was able to bind NAM-B1 when fused to both domains, though the interaction was stronger when NAM-A1 was fused to the binding domain, and NAM-B1 was fused with the activation domain.

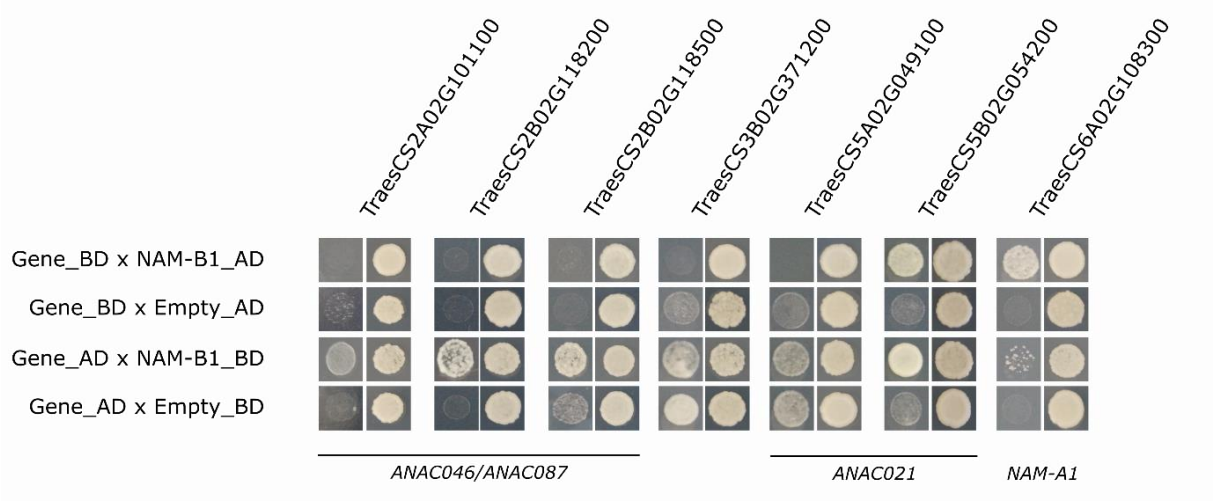

Figure 5-5: NAM-B1 physically interacts with five NAC transcription factors in vivo. Yeast two-hybrid protein interactions were validated on selective media (SC-LTH + 10mM 3AT, left) and control media (SC-LT; right). Growth on selective media when the gene of interest and NAM-B1 are co-expressed (rows 1 and 3) indicates that the two proteins can interact. Growth on selective media where one of the vectors is the empty vector (rows 2 and 4) indicates that any growth seen in the row above is likely due to auto-activation rather than a true interaction between the protein of interest and NAM-B1. The Arabidopsis orthologs (ANAC0XX) and wheat gene names are provided for each candidate gene where possible.

5.3.1.3 Validation of interactions using Co-Immunoprecipitation

While the Y2H comparison provided an indication of potential interaction partners for NAM-B1, it did not validate the ability of NAM-B1 to interact with the NAC transcription factors *in planta*. To validate these interactions in a separate system, we carried out a Co-Immunoprecipitation (Co-IP) experiment. The full coding sequence of *NAM-B1* was cloned into the Gateway binary vector pGWB12, containing an N-terminal FLAG tag. The full coding sequence of the candidate interactors were cloned into pGWB21, containing an N-terminal Myc tag. In both vectors, the expression of the gene was driven by a 35s promoter. For each interaction, the NAM-B1 construct and the candidate gene construct were transformed into *Agrobacterium tumefaciens* and co-infiltrated into *Nicotiana benthamiana*. Following protein extraction, the FLAG-tagged NAM-B1 was immunoprecipitated using magnetic FLAG beads and run on an SDS-PAGE gel. After incubation with the c-Myc antibody, we were able to test whether NAM-B1 had bound the candidate gene.

Using this method, we validated all seven of the interactions identified in the Y2H screen (Figure 5-6). In most cases, the interaction between NAM-B1 and the candidate protein was weak compared to the homo-dimer formation. While we had difficulty obtaining high levels of protein expression, both for NAM-B1 and the candidate interactors, we were able to show that NAM-B1 has the capacity to bind all of the candidate genes *in planta*. However, as the levels of protein in each sample are likely to vary between candidates, and indeed between replicates of each test, we cannot draw any conclusions from these experiments as to the relative strengths of the interactions with NAM-B1. Curiously, two of the candidates which had only shown auto-activation in the Y2H system, TraesCS5A02G049100 and Traes6A02G108300, did bind to NAM-B1 in the Co-IP experiments. It may be the case that the small fragments of the C-terminal domain which remained attached to these genes in the Y2H (see Figure 5-4) induced auto-activation of the system and obscured true binding with NAM-B1 (Figure 5-5). As we were able to validate the interactions of all of the candidate genes with NAM-B1, we proceeded to further characterise all seven of the NAC transcription factors.

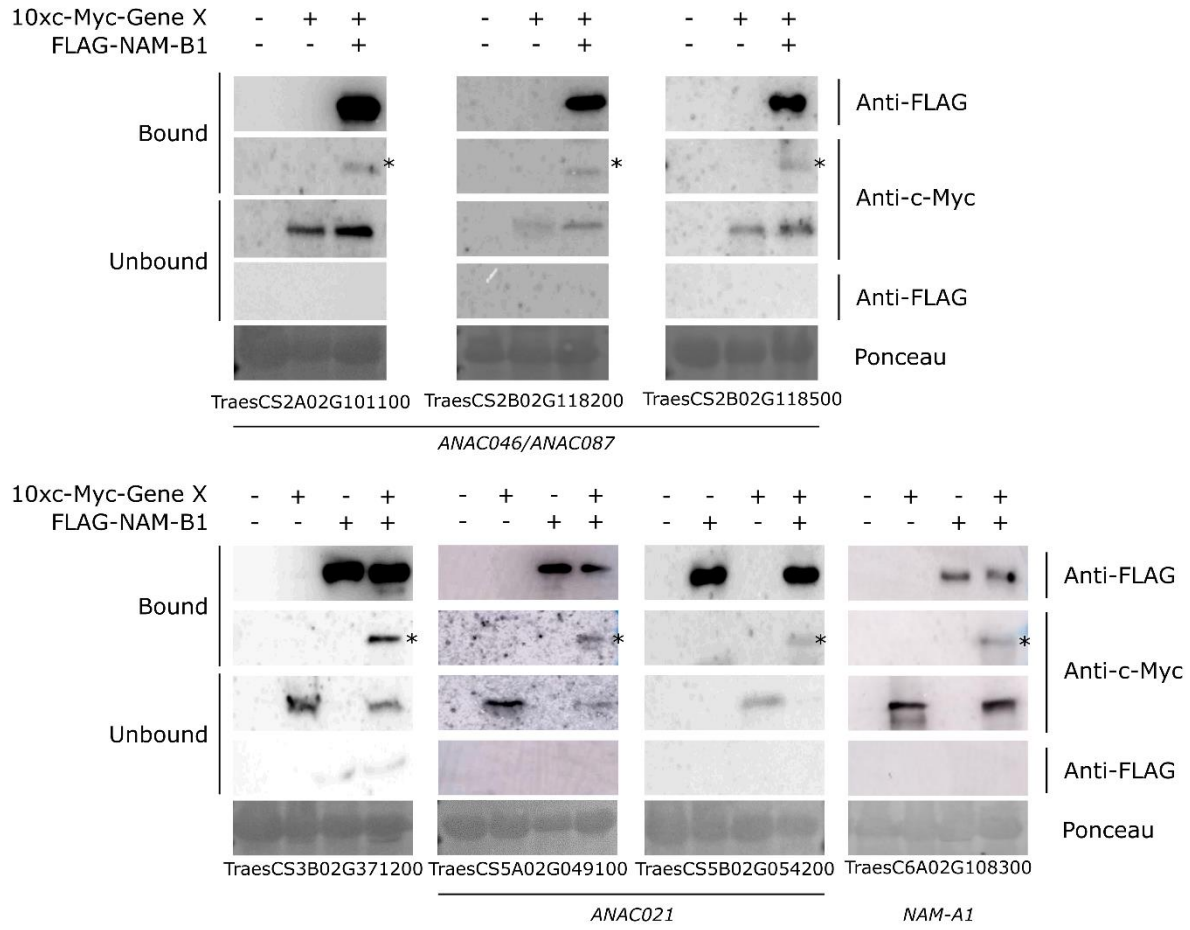

Figure 5-6: NAM-B1 binds the candidate NAC transcription factors in planta. The 10x c-Myc tagged candidate proteins were pulled down alongside Flag-tagged NAM-B1 using anti-Flag magnetic beads ("Bound"). The "Unbound" fraction was extracted following incubation with the anti-Flag beads and reflects the input protein without the proteins bound to the beads. The Flag-tagged bands for NAM-B1 are approximately 46 kDa; the 10x-Myc bands vary for the candidate genes, ranging from approximately 50 kDa to 60 kDa. Images were adjusted for contrast and brightness in order to emphasize the protein bands; all images taken from the same blot were adjusted uniformly. Arabidopsis orthologs (ANAC0XX) and wheat gene names are provided where available.

5.3.2 *Expression profiles of the candidate genes across development and flag leaf senescence*

The initial step in characterising the candidate genes was to investigate their expression across a time-course of wheat development. To do this, we used the developmental time course from the spring wheat cultivar Azhurnaya (Ramírez-González et al. 2018) alongside a senescence-specific timecourse we previously produced (Borrill et al. 2019a). We found that three of the seven candidate genes were upregulated more than two-fold in senescence compared to the rest of the Azhurnaya time course (Figure 5-7). These genes included TraesCS2A02G101100, TraesCS2B02G118500, and TraesCS6A02G108300 (*NAM-A1*). Of these genes, *NAM-A1* is principally expressed during senescence and grain filling, reaching a maximum average TPM of almost 50 in the flag leaf blade during grain ripening. Outside of these two developmental stages, the expression levels of *NAM-A1* are consistently low, at or below 1 transcript per million (TPM). In contrast, TraesCS2B02G118500 also peaked in expression in the flag leaf blade during grain ripening, reaching almost 40 TPM, but unlike *NAM-A1* it is also expressed at moderate levels during earlier stages of development (Figure 5-7). It reaches approximately 10 TPM in the roots during the 3 leaf to 5 leaf stage and is also present in the seedling radicle and roots at 8 TPM. The root expression corresponds with the known function of its Arabidopsis orthologs, *ANAC087* and *ANAC046*, in regulating programmed cell death in the root cap columella (Huysmans et al. 2018).

The third gene with over two-fold greater expression during senescence, TraesCS2A02G101100, is expressed far lower than the other two, peaking at just over 2 TPM in the flag leaf blade during ripening (Figure 5-7). One gene, TraesCS2B02G118200, showed nearly no expression across the developmental time course, with only very low expression (approximately 1 TPM) in the early seedling stages. The homoeologs TraesCS5A02G049100 and TraesCS5A02G054200 were principally expressed in the tillering stage, with the B homoeolog reaching 60 TPM and the A homoeolog reaching just over 30 TPM in the first leaf sheath. They are also highly expressed in roots during the flag leaf and ear emergence, though the B homoeolog has consistently higher levels of expression. Notably, the homoeologs only have low to moderate levels of expression in senescence, reaching a maximum of 6 TPM in the fifth leaf blade at milk grain stage. This contrasts with their Arabidopsis ortholog, *ANAC021*, which is highly upregulated during leaf senescence (Balazadeh et al. 2008). The final candidate, TraesCS3B02G371200, was expressed throughout development, but principally during anthesis and grain filling. It was particularly highly expressed in the lemma and peduncle during milk grain stage, reaching 28 TPM and 24 TPM respectively.

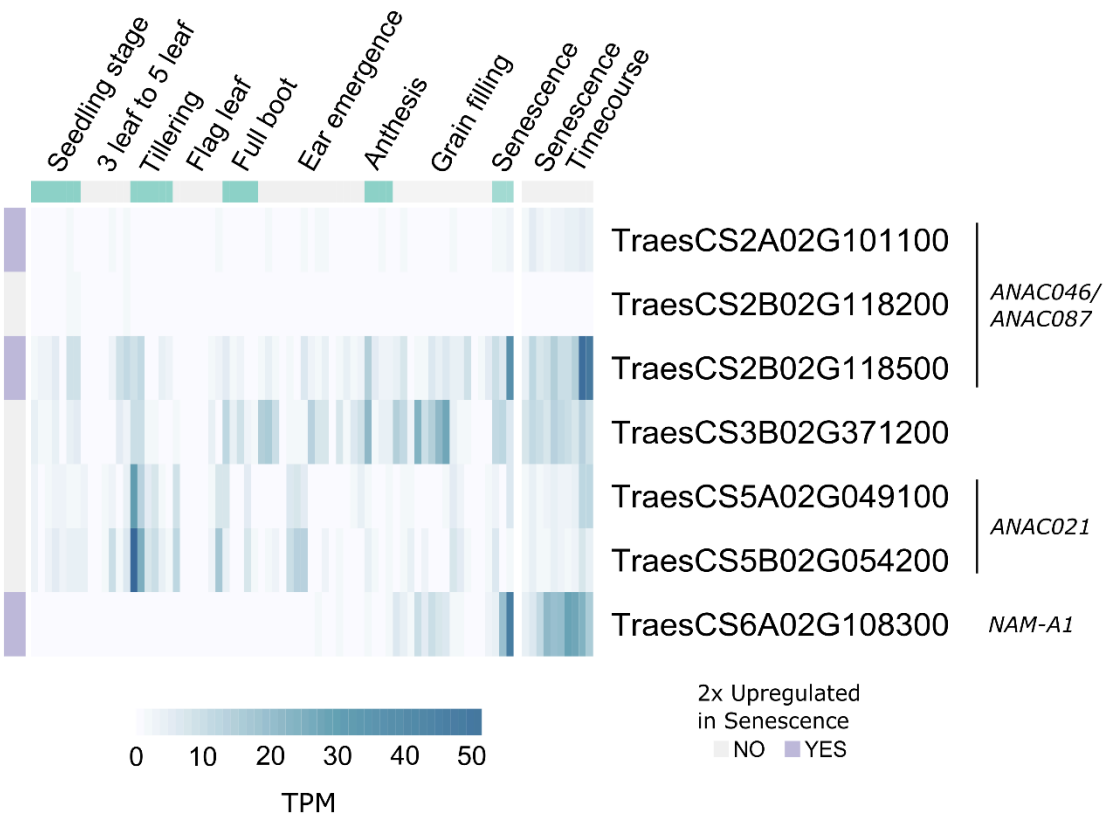

Figure 5-7: Expression profiles of the NAC transcription factors which interact with NAM-B1 across development. Gene expression is shown in transcripts per million (TPM), derived from the cv. Azhurnaya developmental timecourse (Ramírez-González et al. 2018) and the cv. Bobwhite flag leaf senescence timecourse (Borrill et al. 2019a). Genes which are two times (2x) upregulated in the "Senescence" stage of the Azhurnaya timecourse, compared to the average expression level across the rest of the timecourse, are marked with purple. The corresponding gene names of the Arabidopsis orthologues (ANAC0XX) or wheat gene are provided where possible.

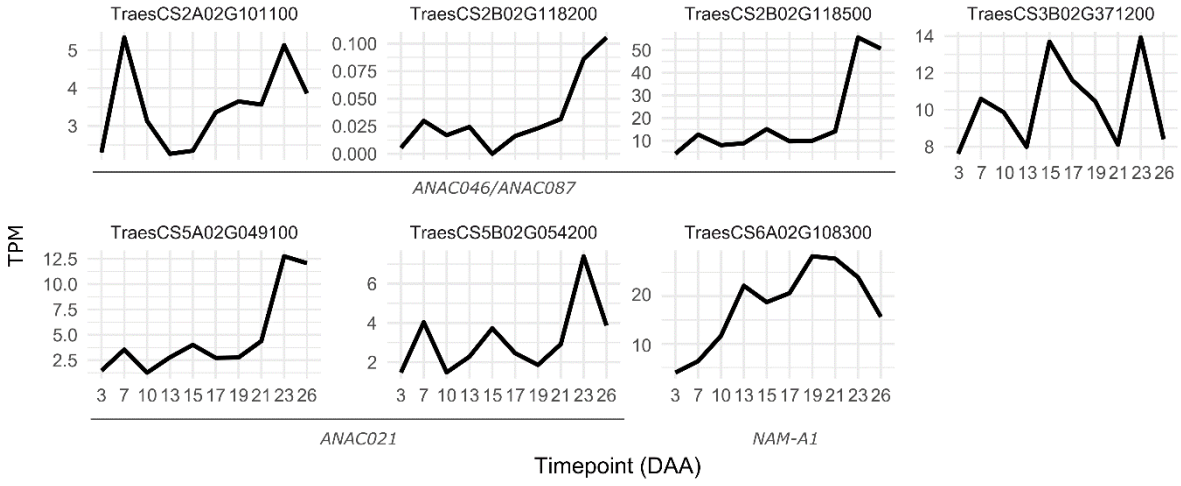

Figure 5-8: Expression values of candidate interactors across the flag leaf senescence time course. Expression values are shown in transcripts per million (TPM), derived from the (Borrill et al. 2019a) timecourse data. The progression of expression is shown as days after anthesis (DAA). Note that the y-axes are not the same for each gene. The corresponding Arabidopsis orthologs and wheat gene names are shown below, where possible.

We looked in more detail at the senescence-specific timecourse and investigated whether the NAC transcription factors show similar profiles of expression across senescence (Figure 5-8). Four of the seven genes show peaks in expression at the end of the senescence timecourse. However, one of these, TraesCS2B02G118200, has very low gene expression overall. The remaining three include TraesCS2B02G118500, which has very high expression at 23 DAA, reaching approximately 55 TPM from 14 TPM two days earlier. In contrast, the closely related gene TraesCS2A02G101100 has a very different expression pattern, peaking at 7 DAA at approximately five TPM and again at 23 DAA.

TraesCS5A02G049100 and TraesCS5B02G054200, both homoeologs, have similar patterns of expression to each other across the timecourse (Figure 5-8). Both have local peaks at seven and 15 DAA of approximately four TPM. However, the A genome homoeolog has a higher peak of expression at 23 DAA, reaching 12.5 TPM compared to just over 7 TPM for the B homoeolog. TraesCS3B02G371200 also has three clear peaks of expression during the timecourse, again at seven, 15, and 23 DAA. Following each peak, its expression drops back to under 8 TPM.

In contrast to the other genes, *NAM-A1* (TraesCS6A02G108300) displays a more consistent increase in expression level across the timecourse, peaking at 19 DAA, with a smaller earlier peak at 13 DAA (Figure 5-8). This pattern is consistent with earlier work on *NAM-A1* where its expression was observed at high levels in the flag leaf at 14 DAA (Guttieri et al. 2013). Knocking out *NAM-A1* leads to greater differential expression of genes compared to a wild-type control at 12 DAA compared to 22 DAA, also supporting a role for *NAM-A1* early in the regulation of senescence, before visual symptoms appear. In contrast, the later peaks in expression for the other candidate genes suggests that they may have a role in regulating later stages of senescence.

5.3.3 *Four candidate genes can induce cell death when expressed in N. benthamiana*

To test whether any of the candidate genes had the potential to induce cell death, a key aspect of monocarpic senescence (Lim et al. 2007a), we established a transient screening system in *Nicotiana benthamiana*. Here, we over-expressed the NAC transcription factors using the same Gateway vectors as previously used in the Co-immunoprecipitation experiment, pGWB12 and pGWB21. Each NAC transcription factor was infiltrated as a patch into the leaf, as depicted in Figure 3-2, p.103. At five days post-infiltration, cell death was scored based on chlorosis and necrosis of the leaf, using a previously published scale (Maqbool et al. 2015) and reproduced in Figure 3-3 (p.103). This method was also used to score the effect of *NAM-A1* mutations in Chapter 3 (Harrington et al. 2019c).

We found that TraesCS2A02G101100, TraesCS2B02G118200, and TraesCS2B02G118500 induced cell death in the leaf to the greatest extent of any of the genes tested (Figure 5-9). *NAM-A1* (TraesCS6A02G108300) and *NAM-B1* were also able to induce cell death, though to a lower

extent. We saw no evidence of cell death in response to expression of the other NAC transcription factors. As the data takes the form of highly skewed count data, we were unable to accurately fit any statistical approximation to these results. However, based on the trends visible in the data, TraesCS2A02G101100, TraesCS2B02G118200, TraesCS2B02G118500, and TraesCS6A02G108300 (*NAM-A1*) could all induce cell death to a similar extent as *NAM-B1*. Indeed, the NAC transcription factors on chromosome 2 appear to induce cell death to an even greater extent than *NAM-A1* and *NAM-B1*. However, the remaining NAC transcription factors failed to show any indication of a cell death response when over-expressed in this heterologous system.

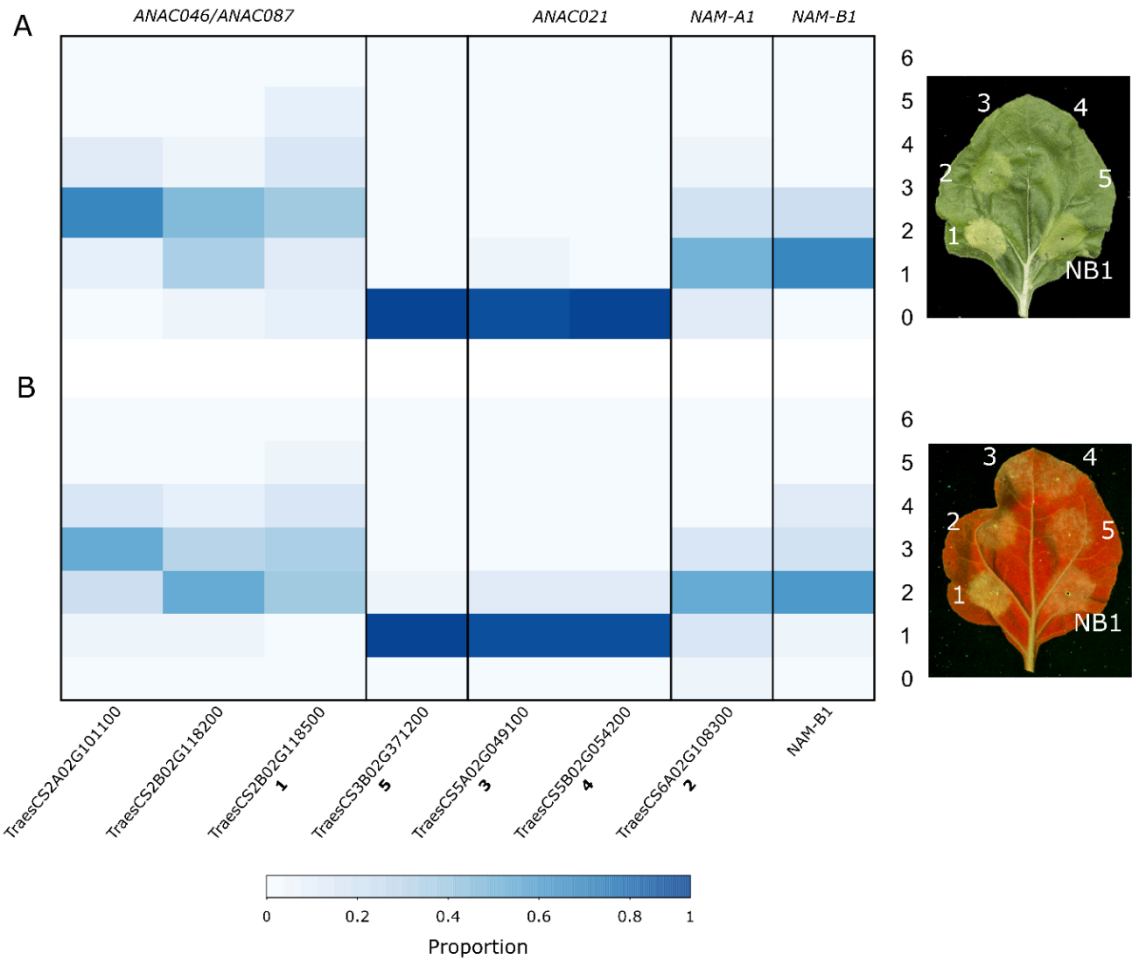

Figure 5-9: Four of the candidate genes and *NAM-B1* induce cell death when overexpressed in *N. benthamiana.* All seven candidate genes, alongside *NAM-B1*, were infiltrated into *N. benthamiana* and scored for cell death responses, based on chlorosis (A) and cellular lysis (B). Infiltrated patches were scored on a scale of 0 – 6, where 0 is no cell death, and 6 is very high cell death, derived from the scale used in Maqbool et al. (2015). The fraction of replicates which fell in the seven bins is shown by the heatmap; N = 20 for all genes. A characteristic image is provided for both chlorosis (top) and cellular lysis scoring (bottom); note that the cellular lysis image was taken on the abaxial side of the leaf and has been flipped horizontally to align with the adaxial chlorosis image. The numbers on the photos correspond to numbers beneath the gene names on the heatmap; NB1 refers to NAM-B1. The corresponding Arabidopsis orthologs or wheat gene names are provided where possible.

5.3.4 *The roles of the candidate NAC transcription factors in senescence.*

To test whether the candidate genes had a role in the regulation of senescence in wheat, we used the Kronos TILLING population (Uauy et al. 2009, Krasileva et al. 2017) to develop double mutants in the A and B genome homoeologs of each gene. The homoeologs were determined based on triad assignments carried out previously (Ramírez-González et al. 2018).We chose premature termination codon (PTC) or start-lost mutations where possible, and otherwise selected missense mutations that had a low SIFT score (0) (Ng and Henikoff 2003) and were located in a highly conserved residue of the NAC domain, as determined using the PSSM viewer (https://www.ncbi.nlm.nih.gov/Class/Structure/pssm/pssm_viewer.cgi) (Table 5-9). Where possible, two independent TILLING lines were selected for each homoeolog. In the following section, we will discuss the effect of the mutations for each homoeolog pair on the onset and progression of senescence. Unfortunately, we were unable to identify any deleterious mutations in TraesCS3B02G371200 or its A-genome homoeolog at the beginning of the project. As a result, this candidate gene was dropped from further study in favour of focussing on those with TILLING mutants.

Table 5-9: Mutations from the Kronos TILLING population selected for the candidate NAC transcription factors and their homoeologs. Homoeolog pairs are grouped by colour.

Gene	Present in Y2H screen?	Arabidopsis Ortholog	Kronos TILLING Line	Mutation	Predicted Effect	SIFT
TraesCS2A02G101100	Yes	*ANAC046/ ANAC087*	K0996	atG/atA	Start lost; predicted 7 amino acid deletion at 5' end.	-
			K1308	tGg/tAg	Premature termination codon, W55*	-
TraesCS2B02G118200	Yes	*ANAC046/ ANAC087*	K2169	tGg/tAg	Premature termination codon, W31*	-
TraesCS2A02G101400	No	*ANAC046/ ANAC087*	K2236	cCc/cTc	Missense mutation, P65L	0
			K3842	cGc/cAc	Missense mutation, R93H	0
TraesCS2B02G118500	Yes	*ANAC046/ ANAC087*	K4485	aCc/aTc	Missense mutation, T28I	0
			K4681	tgG/tgA	Premature termination codon, W62*	-
TraesCS5A02G049100	Yes	*ANAC021*	K3964	tgG/tgA	Premature termination codon, W72*	-
			K2825	Ggg/Agg	Missense mutation, G95R	0
TraesCS5B02G054200	Yes	*ANAC021*	K2132	cCg/cTg	Missense mutation, P14L	0

5.3.4.1 Mutations in TraesCS2A02G101100 and TraesCS2B02G118200 do not affect senescence timing.

For the first pair of homoeologs, TraesCS2A02G101100 and TraesCS2B02G118200, two sets of crosses were made between two independent mutations in the A genome (K0996 and K1308) and a single mutation in the B genome copy (K2169). Based on the predicted effects of the TILLING mutations, we expected the K1308 x K2169 cross to have the strongest phenotype, as both mutations are for premature termination codons (Table 5-9). These homoeologs were orthologous to the Arabidopsis genes *ANAC046* and *ANAC087*.

Following crossing and bulking of seed, we grew the double mutants and their wild-type siblings in the F_3 generation under glasshouse conditions. The plants were individually scored for visual senescence onset in the flag leaf and the peduncle. Neither cross showed any variation in flag leaf senescence or peduncle senescence (Figure 5-10A, Figure 5-10B; Table 5-10). This qualitative score was complemented with quantitative measurements of leaf greenness using a SPAD meter (Figure 5-10C). Here, no significant differences in leaf greenness were found between the wild-type and mutants of either line during the progression of senescence, suggesting that the mutations have no effect. This experiment was repeated under controlled environment conditions, and again no difference was seen in senescence timing between the wild-type and double mutant plants.

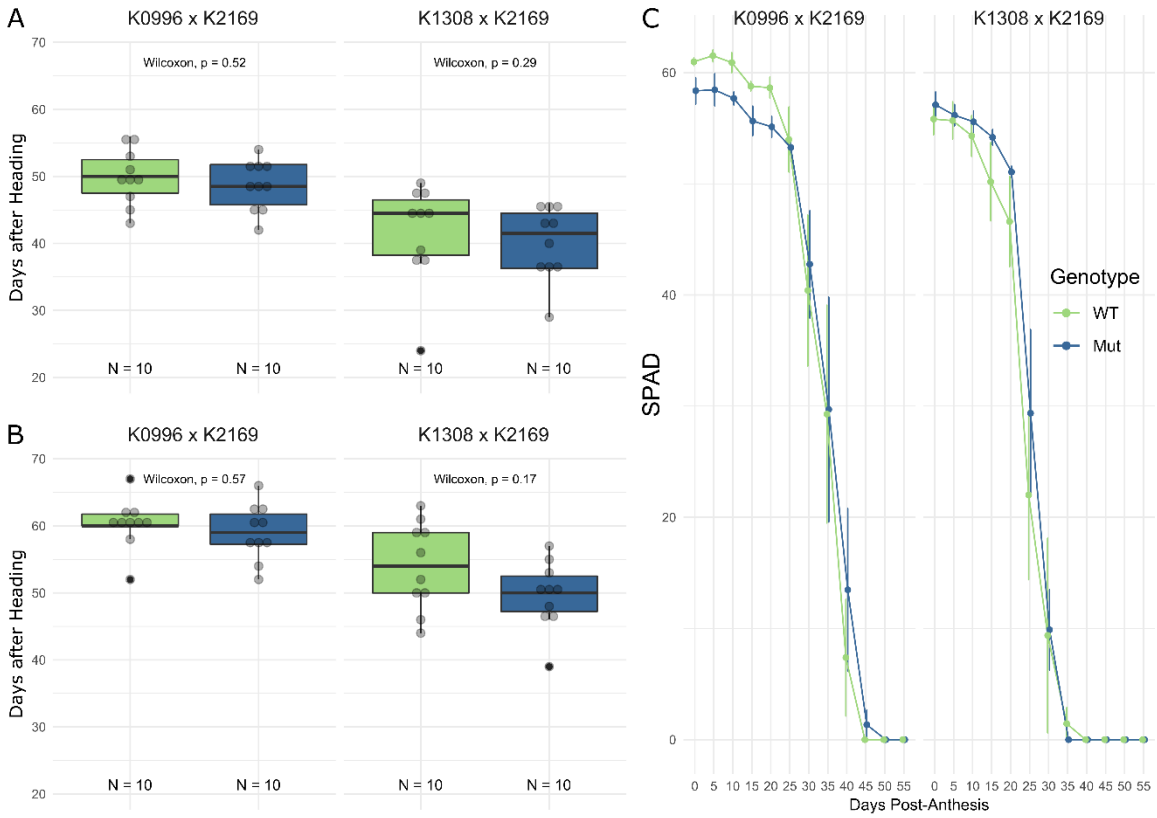

Figure 5-10: Mutations in TraesCS2A02G101100 and TraesCS2B02G118200 show no substantial variation in senescence timing. Plants were scored for visual senescence in the flag leaf (A) and peduncle (B), reported as days after heading where wild-type is green and mutant is blue. Chlorophyll levels in the flag leaf were also scored using a SPAD meter for five plants of each genotype (C) at five-day intervals post-anthesis. Statistical comparisons between the wild-type and mutant lines were carried out using the two-sample Wilcox test for A and B, and the pairwise Wilcoxon rank-sum test for C.

Table 5-10: Summary statistics for F_3 phenotyping of TraesCS2A02G101100 and TraesCS2B02G118200 TILLING mutants.

Cross	Variable	Genotype	N	Mean	Standard Deviation
K0996 x K2169	Days from Heading to Flag Leaf Senescence	aabb	10	48.6	3.78
		AABB	10	49.9	4.15
	Days from Heading to Peduncle Senescence	aabb	10	59.1	4.20
		AABB	10	60.2	3.74
K1308 x K2169	Days from Heading to Flag Leaf Senescence	aabb	10	40	5.44
		AABB	10	41.6	7.49
	Days from Heading to Peduncle Senescence	aabb	10	49.6	5.08
		AABB	10	54	6.53

We then carried out a backcross (BC) for the K1308 x K2169 line to the wild-type parent Kronos. This cross was selected for the backcross as the mutations in each homoeolog were most likely to have a deleterious effect on gene function. The BC_1F_3 lines were also scored for visual flag leaf and

peduncle senescence. In neither case was the difference in senescence timing significantly different from the wild-type plants (Figure 5-11A, Figure 5-11B; Table 5-11). However, the scores of leaf greenness using the SPAD meter found that the double mutants contained less chlorophyll from anthesis until approximately 18 days post-anthesis ($p < 0.05$, pairwise Wilcoxon rank-sum test; Figure 5-11C). It may be the case, therefore that the slight trend towards early senescence observed in the visual scoring is due not to early senescence itself, but rather to lower levels of chlorophyll in the plant overall. This experiment was repeated a second time under glasshouse conditions, again showing no variation in senescence timing between the wild-type and mutant plants.

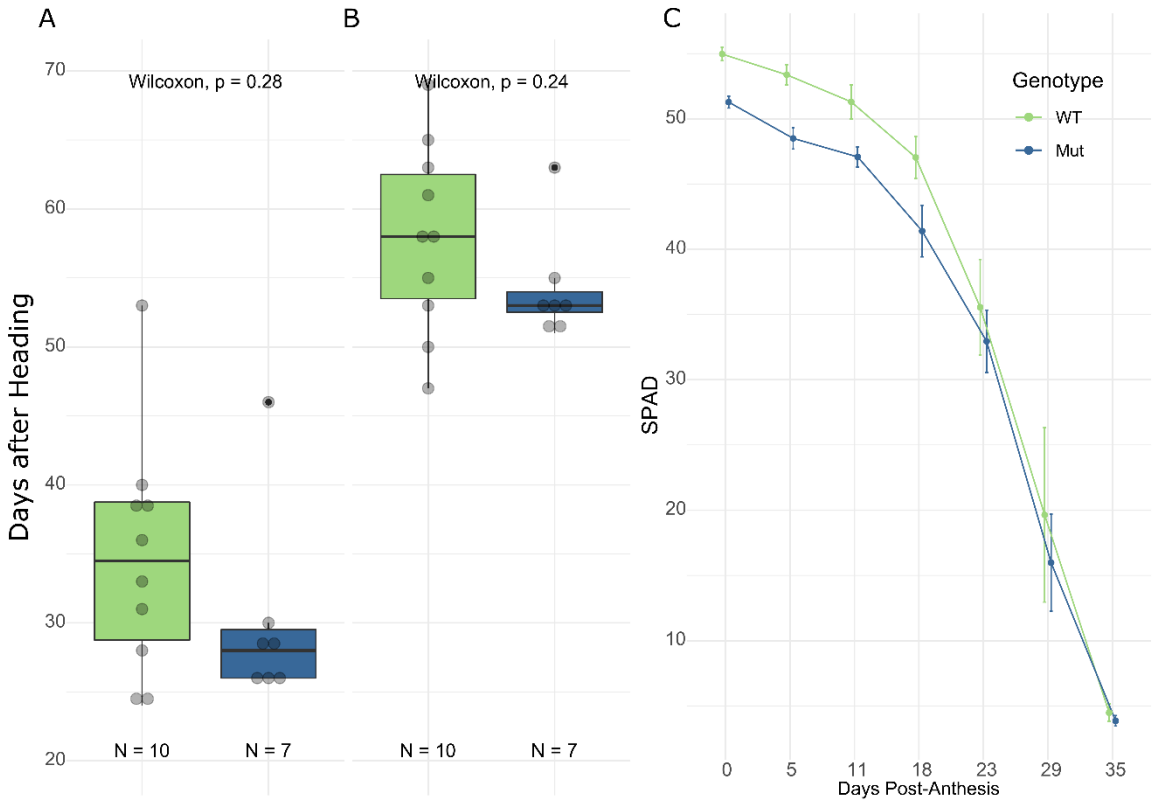

Figure 5-11: Mutations in TraesCS2A02G101100 and TraesCS2B02G118200 show no substantial variation in senescence timing following backcrossing to cv. Kronos. Plants were scored for visual senescence in the flag leaf (A) and peduncle (B), with wild-type shown in green and mutant in blue. Chlorophyll levels in the flag leaf were also scored using a SPAD meter (C) at intervals post-anthesis (N = 5 for all genotypes at each timepoint). The mutant plants contained significantly ($p < 0.05$) fewer SPAD units (as a proxy for chlorophyll content) than the wild-type plants until 23 days post-anthesis. Statistical comparisons between the wild-type and mutant lines were carried out using the two-sample Wilcox test for A and B, and the pairwise Wilcoxon rank-sum test for C.

These results suggest that the two homoeologs, despite physically interacting with NAM-B1, do not appear to be involved in regulating senescence. This correlates with the low levels of expression seen for the B homoeolog during senescence, and indeed the rest of development (Figure 5-7, Figure 5-8). However, the A homoeolog does show upregulation in the later stages of development, behaving like a candidate positive regulator of senescence. It is instead possible that

the genes may act to regulate senescence, but in a redundant or subtle manner such that deleterious mutations of the genes are not sufficient to perturb the normal senescence process.

Table 5-11: Summary statistics for BC_1F_3 phenotyping of TraesCS2A02G101100 and TraesCS2B02G118200 TILLING mutants.

Cross	Variable	Genotype	N	Mean	Standard Deviation
K1308 x K2169	Days from Heading to Flag Leaf Senescence	aabb	7	30.14	7.17
		AABB	10	34.70	8.59
	Days from Heading to Peduncle Senescence	aabb	7	54.29	4.03
		AABB	10	57.90	6.85

5.3.4.2 Mutations in TraesCS5A02G049100 and TraesCS5B02G054200 may lead to earlier senescence.

Crosses were carried out between TILLING mutants of the homoeologs TraesCS5A02G049100 and TraesCS5B02G054200, both of which were identified in the Y2H screen and are orthologs of the Arabidopsis gene *ANAC021*. At the time of crossing, only one missense mutation could be identified in the B homoeolog which was predicted to be highly deleterious. This single mutation was crossed to two independent mutations in the A genome, one a premature termination codon (K3964) and the other a missense mutation with a SIFT score of 0 (K2825) (Table 5-9).

Initial screens of the F_3 mutant population suggested that the genes may act redundantly as negative regulators of senescence, as the double mutant in the K2825 x K2132 cross displayed significantly earlier peduncle senescence ($p < 0.05$, two-sample Wilcox test; Figure 5-12B). While flag leaf senescence was not significantly earlier in the double mutants, there was a trend towards earliness overall with the double mutant senescing on average 3-4 days before wild type ($p = 0.33$, two-sample Wilcox test; Figure 5-12A, Table 5-12). As we were unable to recover double mutants in the second cross, K3964 x K2132, we could not confirm this with an independent A-genome mutation. Measurement of leaf greenness using the SPAD meter failed to identify any clear difference between the mutant and wild-type plants (Figure 5-12C), though this could be in part due to the substantial variation observed in SPAD measurements within each genotype. These data suggested that the effect of the gene may be fully redundant between the homoeologs, such that the double mutant is required to see any phenotypic effect (Borrill et al. 2019b).

The initial glasshouse experiment was also repeated under controlled environment conditions. In this case, no significant variation was observed between the wild-type plants and the double mutants (data not shown). These results suggested that perhaps the initially observed early senescence phenotype was not representative of the effect of the mutations. However, the timing and progression of senescence is known to be highly dependent on environmental conditions (see

further discussion in section 6.2.1). As a result, the loss of the phenotype may have been due to variation in light or temperature conditions.

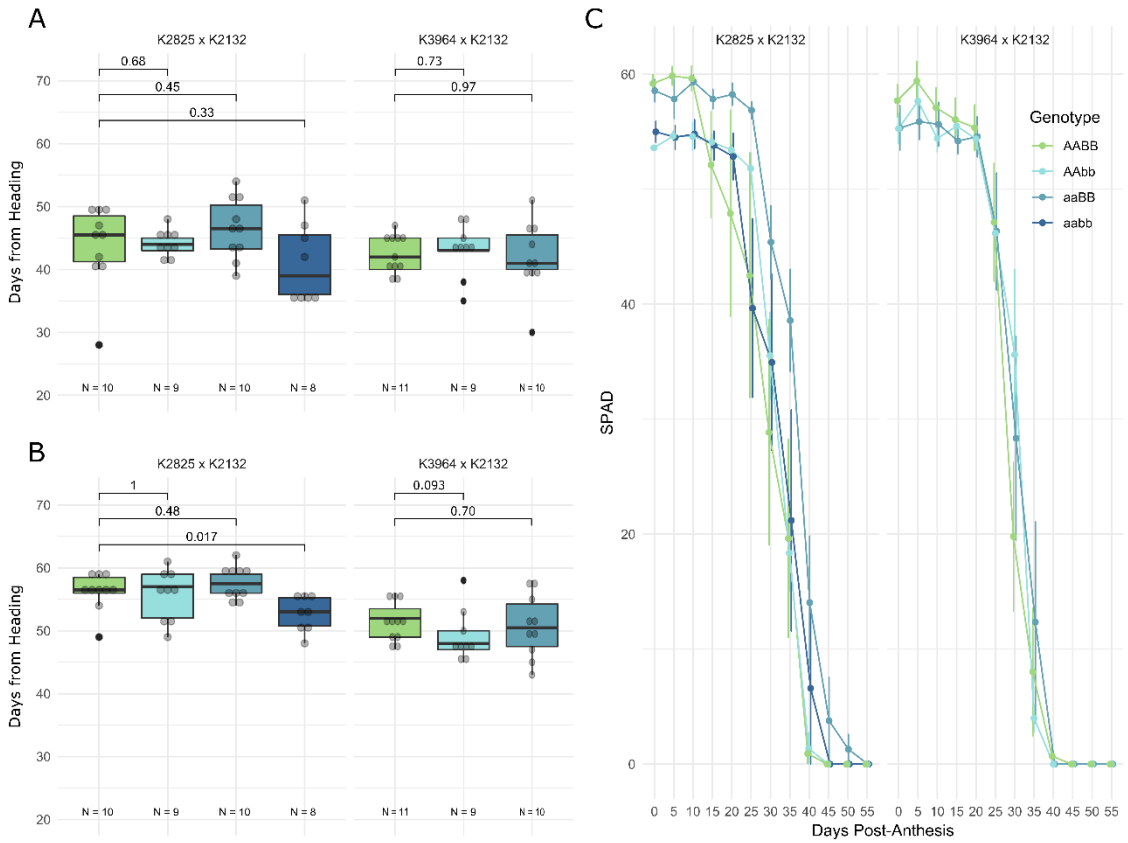

Figure 5-12: Mutations in TraesCS5A02G049100 and TraesCS5B02G054200 may lead to premature peduncle senescence. Plants were scored for visual senescence in the flag leaf (A) and peduncle (B), where the genotype colours match the legend in panel C. Chlorophyll levels in the flag leaf were also scored using a SPAD meter (C) at intervals post-anthesis for five plants for each genotype at each timepoint. Note that no double mutants could be recovered for the K3964 x K2132 cross. Statistical comparisons between the wild-type and mutant lines were carried out using the two-sample Wilcox test for A and B, and the pairwise Wilcoxon rank-sum test for C.

Table 5-12: Summary statistics for F_3 phenotyping of TraesCS5A02G049100 and TraesCS5B02G054200 TILLING mutants.

Cross	Variable	Genotype	N	Mean	Standard Deviation
K3964 x K2132	Days from Heading to Flag Leaf Senescence	AABB	11	42.45	3.05
		AAbb	9	43.00	4.24
		aaBB	10	41.90	5.67
	Days from Heading to Peduncle Senescence	AABB	11	51.55	3.14
		AAbb	9	49.11	4.08
		aaBB	10	50.70	4.97
K2825 x K2132	Days from Heading to Flag Leaf Senescence	AABB	10	43.70	6.53
		AAbb	9	44.11	2.15
		aaBB	10	46.50	4.88
		aabb	8	41.00	6.14
	Days from Heading to Peduncle Senescence	AABB	10	56.20	3.01
		AAbb	9	55.67	4.09
		aaBB	10	57.60	2.55
		aabb	8	52.75	2.92

To pursue this further, we backcrossed the double mutant of the K2825 x K2132 cross with the wild-type cultivar Kronos. Following bulking, we scored the BC_1F_3 generation for senescence under glasshouse conditions, more similar to those of the initial experiment (Figure 5-12). Unlike at the F_3 generation, we saw no evidence for early peduncle senescence in the double mutants compared to wild-type ($p = 0.27$, pairwise Wilcoxon rank-sum test; Figure 5-13B) nor was there any difference in flag leaf senescence (Figure 5-13A,

Table 5-13). The single B-genome mutant did show early peduncle senescence compared to the wild-type ($p < 0.05$), however there was no significant variation in SPAD levels for any of the mutant lines (Figure 5-13C). The experiment was repeated under glasshouse conditions, and again failed to recapitulate the early peduncle senescence phenotype (Data not shown; $p = 0.36$, pairwise Wilcoxon rank-sum test).

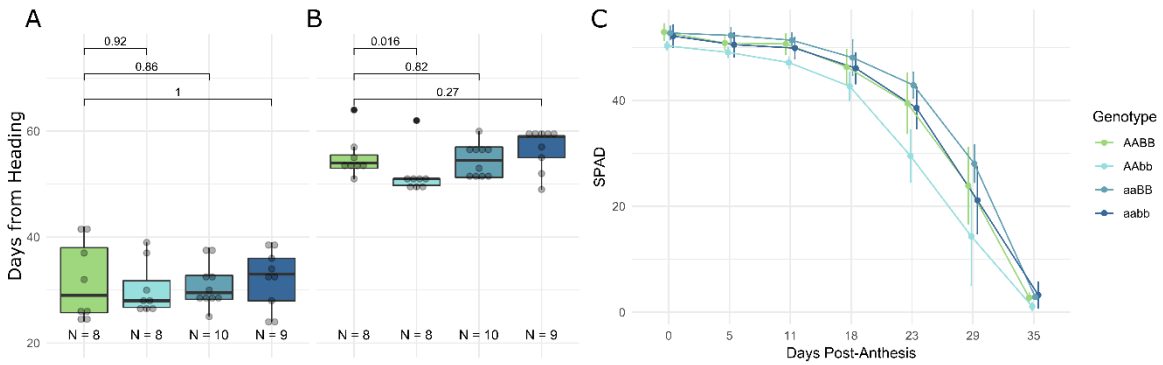

Figure 5-13: Double mutants in TraesCS5A02G049100 and TraesCS5B02G054200 did not maintain the early peduncle senescence phenotype following backcrossing to cv. Kronos. As previously, plants were scored for visual senescence in the flag leaf (A) and peduncle (B) and for chlorophyll levels in the flag leaf using a SPAD meter (C; N = 5 for all genotypes at each timepoint). All genotypes are depicted using the colours in the legend in panel C. Statistical comparisons between the wild-type and mutant lines were carried out using the pairwise Wilcoxon rank-sum test.

As the initial F_3 phenotype was not recapitulated in the BC_1F_3 generation, it is possible that the observed early senescence was due to a background mutation in the TILLING lines. This background mutation may have been segregated away as a result of the backcross, leading to a loss of the early senescence phenotype. Alternatively, the original early senescence phenotype observed may have been due to local environmental heterogeneity which was avoided in the further controlled environment and glasshouse experiments.

Table 5-13: Summary statistics for BC_1F_3 phenotyping of TraesCS5A02G049100 and TraesCS5B02G054200 TILLING mutants.

Cross	Variable	Genotype	N	Mean	Standard Deviation
K2825 x K2132	Days from Heading to Flag Leaf Senescence	AABB	8	31.63	7.46
		AAbb	8	30.13	5.06
		aaBB	10	30.90	4.12
		aabb	9	32.00	5.59
	Days from Heading to Peduncle Senescence	AABB	8	55.13	3.98
		AAbb	8	51.75	4.23
		aaBB	10	54.50	3.27
		aabb	9	56.67	3.91

5.3.4.3 Mutations in TraesCS2A02G101400 and TraesCS2B02G118500 lead to delayed leaf and peduncle senescence

The final pair of homoeologs which were tested using the Kronos TILLING population were TraesCS2A02G101400 and TraesCS2B02G118500, orthologs of the Arabidopsis genes *ANAC046* and *ANAC087*. The B-genome homoeolog had been identified in the Y2H screen, but the A genome homoeolog had not been. To test whether the A-genome homoeolog could also interact

with NAM-B1, we cloned TraesCS2A02G101400 into the Y2H vectors, pDEST22 and pDEST32, and carried out a Y2H experiment (Figure 5-14). We found that the A-genome homoeolog could indeed interact with NAM-B1, though only when it was fused to the binding domain. This contrasts with the B-genome homoeolog, which only interacted with NAM-B1 when fused to the activation domain (Figure 5-5).

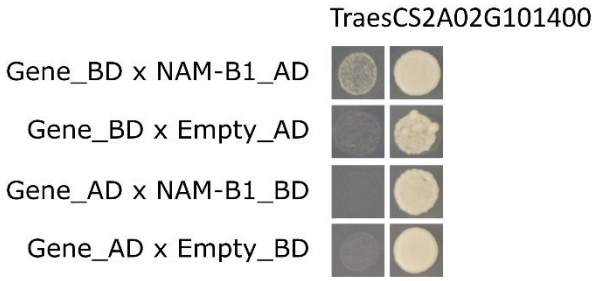

Figure 5-14: TraesCS2A02G101400 interacts with NAM-B1 in a Y2H assay. Yeast colonies were grown on selective media (SC-LTH + 10 mM 3AT, left) and control media (SC-LT, right). Growth on the selective media in rows 1 and 3 indicates the presence of an interaction between NAM-B1 and TraesCS2A02G101400. Growth on selective media for the controls (rows 2 and 4) would indicate that the presence of TraesCS2A02G101400 alone is sufficient to induce growth, and thus not due to an interaction between TraesCS2A02G101400 and NAM-B1.

Initial phenotyping of the TILLING mutants was carried out at the F_2 generation for two of the crosses, K2236 x K4681 and K3842 x K4485 (Figure 5-15, Table 5-14). Only visual senescence was scored for senescence onset in the flag leaf and peduncle. No significant difference was observed for the leaf senescence timings (Figure 5-15A), though on average the K2236 x K4681 double mutant plants were delayed in senescence by ten days compared to wild-type (Table 5-14). In contrast, the K3842 x K4485 double mutants were only delayed very slightly compared to the wild-type plants (76 to 72 days, Table 5-14).

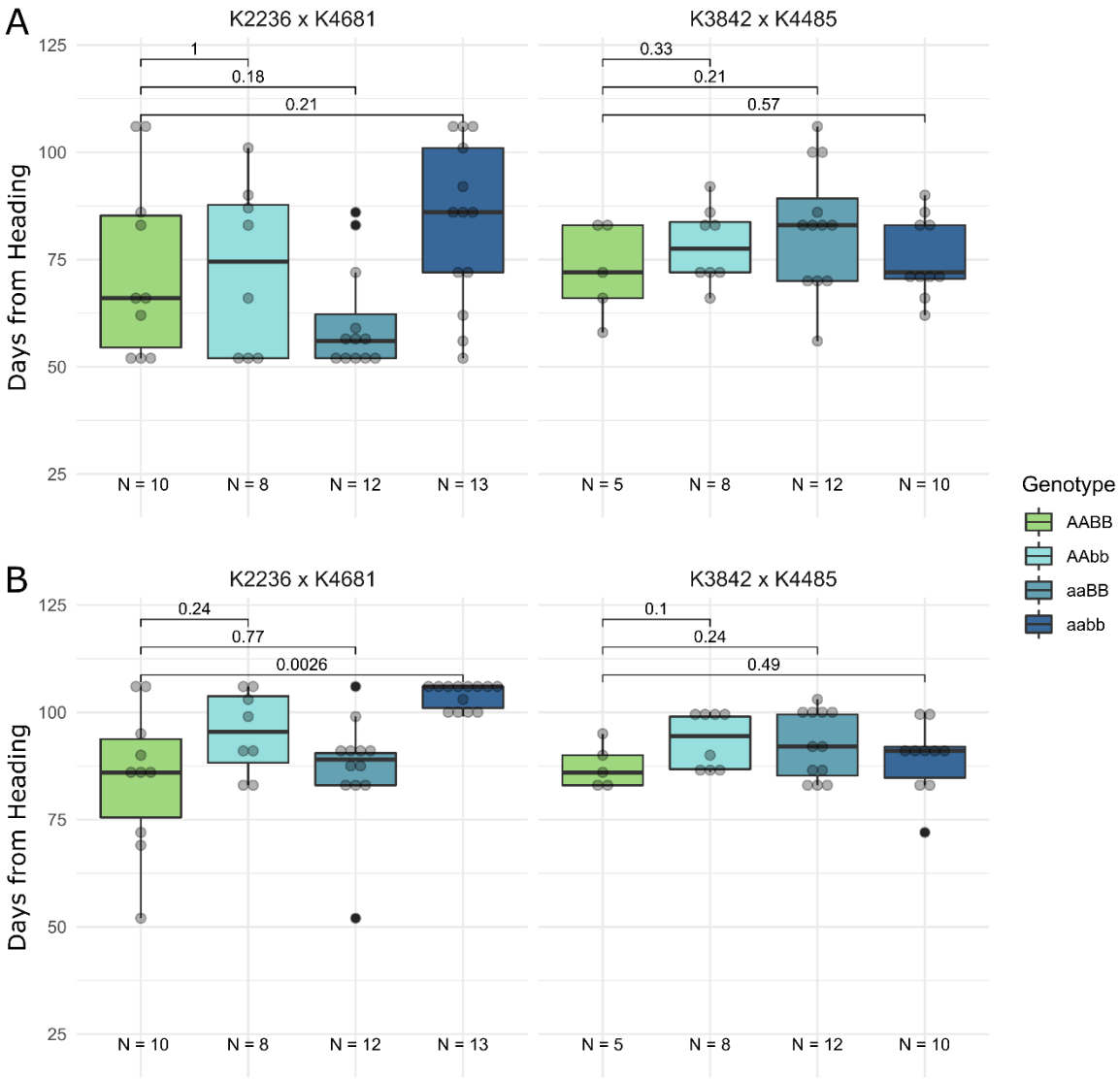

Figure 5-15: Peduncle senescence is delayed in the K2236 x K4681 double mutant. F_2 plants were phenotyped for visual flag leaf (A) and peduncle (B) senescence, presented as days to senescence after heading. Pairwise statistical comparisons were carried out between each mutant and the corresponding wild-type line using the two-sample Wilcox test for A and B, and the pairwise Wilcoxon rank-sum test for C.

We observed a significant delay of 19 days in peduncle senescence for the K2236 x K4681 double mutants ($p < 0.01$, pairwise Wilcoxon rank-sum test; Figure 5-15B, Table 5-14). The single mutant in the B genome, derived from the K4681 premature stop codon mutation, also showed a delay in peduncle senescence of approximately 20 days compared to the mean wild-type value. This delay was not statistically significant, principally due to the large spread in senescence timings for the wild-type plants (standard deviations of 21 and 17 days for the flag leaf and peduncle senescence, respectively, Table 5-14). Note that, due to time constraints, the experiment was terminated approximately 105 days after heading. Not all of the K2236 x K4681 double mutants had completed senescence at that time, leading to an artificial clump of datapoints at approximately 105 days post-heading.

Table 5-14: Summary statistics for F_2 phenotyping of TraesCS2A02G101400 and TraesCS2B02G118500 TILLING mutants.

Cross	Variable	Genotype	N	Mean	Standard Deviation
K2236 x K4681	Days from Heading to Flag Leaf Senescence	AABB	10	73.10	21.05
		AAbb	8	72.88	19.79
		aaBB	12	60.75	12.45
		aabb	13	83.31	19.13
	Days from Heading to Peduncle Senescence	AABB	10	84.80	16.75
		AAbb	8	95.25	9.59
		aaBB	12	86.92	12.90
		aabb	13	104.00	2.77
K3842 x K4485	Days from Heading to Flag Leaf Senescence	AABB	5	72.40	10.88
		AAbb	8	78.25	8.96
		aaBB	12	82.50	14.62
		aabb	10	75.60	9.26
	Days from Heading to Peduncle Senescence	AABB	5	87.40	5.13
		AAbb	8	93.25	6.54
		aaBB	12	92.42	7.88
		aabb	10	89.30	8.23

Following initial phenotyping at the F_2 stage, the double mutants of the two crosses (K2236 x K4681 and K3842 x K4485) were backcrossed to wild-type Kronos. A third cross, K3842 x K4681, was also produced at the F_2 generation, but too few seeds were recovered for phenotyping. Nevertheless, this cross was also backcrossed to Kronos and phenotyped in the BC_1. Following crossing, seeds were bulked in the BC_1F_2 generation, and then phenotyped for senescence timing in the BC_1F_3 generation (Figure 5-16, Table 5-15).

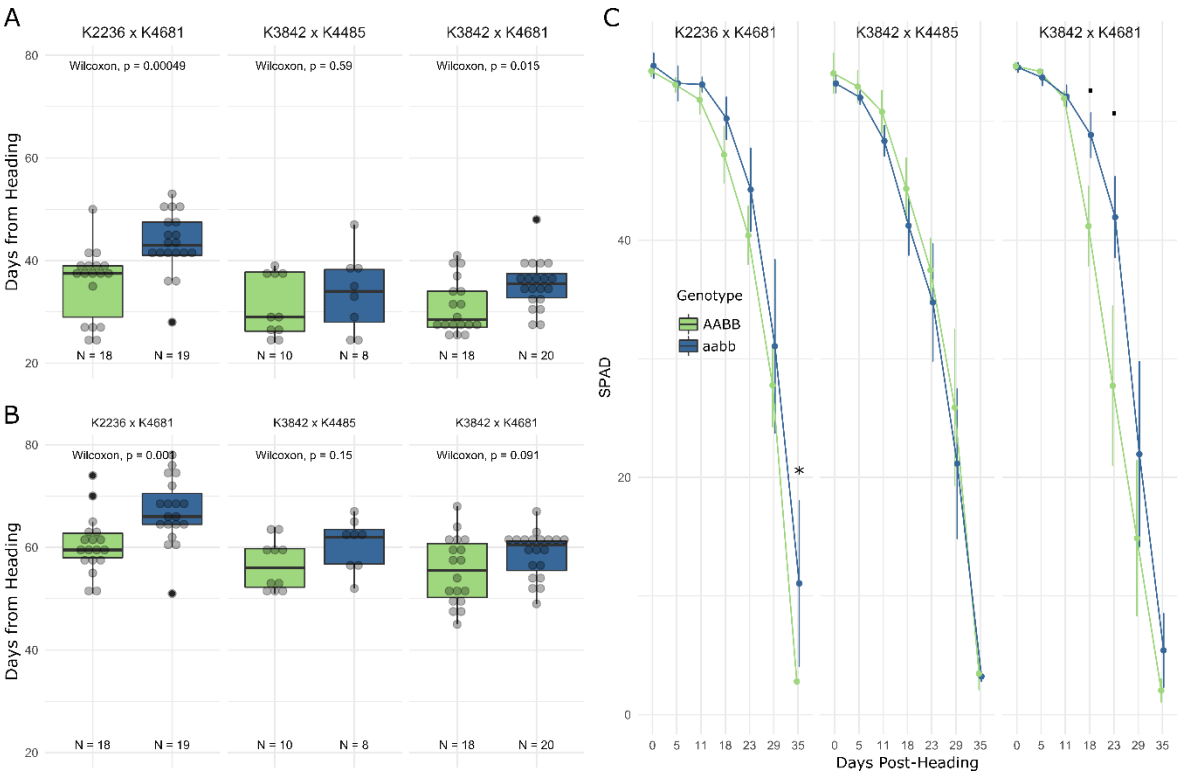

Figure 5-16: Independent TILLING mutants of TraesCS2A02G101400 and TraesCS2B02G118500 exhibit delayed flag leaf senescence. Visual senescence was scored for the flag leaf (A) and peduncle (B) of individual mutant plants, scored as days post-heading. The SPAD meter was used to measure leaf greenness of five plants of each genotype at intervals following heading (C). Statistical comparisons between the wild-type (AABB) and mutant (aabb) plants were carried out using the two-sample Wilcox test for A and B, and the pairwise Wilcoxon rank-sum test for C. For the SPAD measurements, p-values are reported as *, $p < 0.05$; •, $p < 0.1$.

Table 5-15: Summary statistics for the BC_1F_3 generation, phenotyped for visual flag leaf and peduncle senescence.

Cross	Variable	Genotype	N	Mean	Standard Deviation
K2236 x K4681	Days from Heading to Flag Leaf Senescence	AABB	18	35.61	6.90
		aabb	19	43.42	5.99
	Days from Heading to Peduncle Senescence	AABB	18	60.39	5.60
		aabb	19	67.05	6.41
K3842 x K4485	Days from Heading to Flag Leaf Senescence	AABB	10	31.20	6.07
		aabb	8	33.75	7.72
	Days from Heading to Peduncle Senescence	AABB	10	56.60	4.93
		aabb	8	60.50	5.04
K3842 x K4681	Days from Heading to Flag Leaf Senescence	AABB	18	31.00	5.27
		aabb	20	35.30	4.74
	Days from Heading to Peduncle Senescence	AABB	18	55.44	6.55
		aabb	20	58.65	4.51

Following scoring for visual flag leaf senescence, we found that for all three crosses, the double mutant lines showed at least a degree of delay in senescence onset. The strongest delay was seen for the K2236 x K4681 line, with a delay of seven days compared to wild-type ($p < 0.001$, two-

sample Wilcox test; Figure 5-16A, Table 5-15). This line exhibited a substantial delay in whole plant senescence, shown at 63 days post heading in Figure 5-17. The other cross with the K4681 premature termination codon mutation in the B-genome homoeolog, K3842 x K4681, had the next largest delay of five days ($p < 0.05$, two-sample Wilcox test; Figure 5-16A, Table 5-15). The third cross, K3842 x K4485, which only contains missense mutations in the two homoeologs, showed a delay of three to four days in flag leaf senescence, though this was not significant compared to wild-type ($p = 0.59$, two-sample Wilcox test; Figure 5-16A, Table 5-15).

Figure 5-17: The K2236 x K4681 double mutant showed delayed monocarpic senescence compared to wild-type. The mutant (right) and wild-type (left) plants were photographed at 63 days after heading. The scale (far right) is in increments of 1 cm.

We complemented the visual scoring of senescence with the quantitative SPAD measurements of leaf greenness at intervals following heading (Figure 5-16C). For the cross with the strongest leaf senescence phenotype, K2236 x K4681, the mutant plants were only significantly greener than the wild-type plants at 35 days after heading, the final measurement point, when the measured wild-type plants were all fully senesced ($p < 0.01$, pairwise Wilcoxon rank-sum test; Figure 5-16C). However, the mutant plants had consistently higher SPAD readings across the timecourse, corresponding with the delayed senescence observed in the visual scoring. This difference was not present at anthesis, indicating that the variation in chlorophyll content was not a result of the mutant plants starting with higher chlorophyll content. The other cross with significantly delayed senescence onset (K3842 x K4681) did not show any significant differences at the $p < 0.05$ level in SPAD readings between the wild-type and mutant plants. However, at 18 and 23 days post-heading the SPAD levels were significant at the $p < 0.1$ level ($p = 0.075$ and $p = 0.056$, respectively,

pairwise Wilcoxon rank-sum test; Figure 5-16C). No significant differences were seen at any stage for the K3842 x K4485 cross, in keeping with the small, non-significant delay in senescence onset observed.

Onset of peduncle senescence was also scored for all three lines. Here, similar patterns were observed as for the flag leaf senescence. The strongest delay in senescence was seen for the K2236 x K4681 line, with a delay of approximately seven days in the mutant lines compared to wild-type ($p < 0.01$, two-sample Wilcox test; Figure 5-16B, Table 5-15). K3842 x K4681 also showed a delay in senescence of approximately four days (Table 5-15), though this was not significant to the $p < 0.05$ level ($p = 0.09$, two-sample Wilcox test; Figure 5-16B). The third cross, K3842 x K4485, showed a similar delay in senescence (approximately four days, Table 5-15), though this was again not significant ($p = 0.15$, two-sample Wilcox test).

The BC_1F_3 experiments were repeated twice under similar glasshouse conditions (Appendix 8.3.3). Each experiment recapitulated the delay in flag leaf and peduncle senescence observed for the K2236 x K4681 cross, though to a lesser extent than in the initial experiment reported above. In line with this, the senescence delays observed for the other two lines, with less extreme senescence phenotypes, were typically smaller or non-existent. As a result, we decided to take the plants forward to a second backcross generation in an attempt to further clean-up the genetic background, removing more of the undesired EMS-induced mutations present in the TILLING lines.

At the BC_2F_2 generation, we selected for the single and double mutants of each of the three lines previously phenotyped in the BC_1F_3 generation. As previously, the plants were phenotyped for visual flag leaf (Figure 5-18A) and peduncle (Figure 5-18B) senescence. A significant delay in flag leaf senescence was seen for both single mutants in the K2236 x K4682 cross ($p < 0.05$, pairwise Wilcoxon rank-sum test; Figure 5-18A) of four and eight days compared to wild-type (Table 5-16). The double mutant was not significantly delayed compared to wild-type ($p = 0.07$, pairwise Wilcoxon rank-sum test; Figure 5-18A), though on average the plants showed a delay of five days compared to the wild-type plants. However, the delay in senescence previously seen for the K3842 x K4681 line was completely lost in the BC_2F_2 generation, with the double mutant actually showing a one to two day advance in senescence onset compared to wild-type (Figure 5-18A, Table 5-16). The third cross, K3842 x K4485, also failed to show any significant delay in flag leaf senescence compared to wild-type. Both single mutant lines did show slight delays in senescence of four and three days for the A and B genome mutants, respectively, though no delay was observed for the double mutant (Table 5-16).

SPAD measurements of the three different lines were carried out on plants from each genotype, and which correspond with the visual scoring observed. For the K2236 x K4681 cross, the double mutant maintained higher leaf greenness than the wild-type plants until 17 days post-heading ($p < 0.05$, pairwise Wilcoxon rank-sum test; Figure 5-18C). Following this, while there was no

significant difference the double mutant plants had higher SPAD readings on average than the wild-type plants until the completion of senescence (Figure 5-18C). This trend was also seen for both the K2236 (aaBB) and K4681 (AAbb) single mutants. Curiously, the K3842 x K4681 cross showed significantly reduced leaf greenness in the double mutant compared to wild-type until 24 days post-heading, a trend which was not observed in the previous BC_1F_3 experiments (Figure 5-16C).

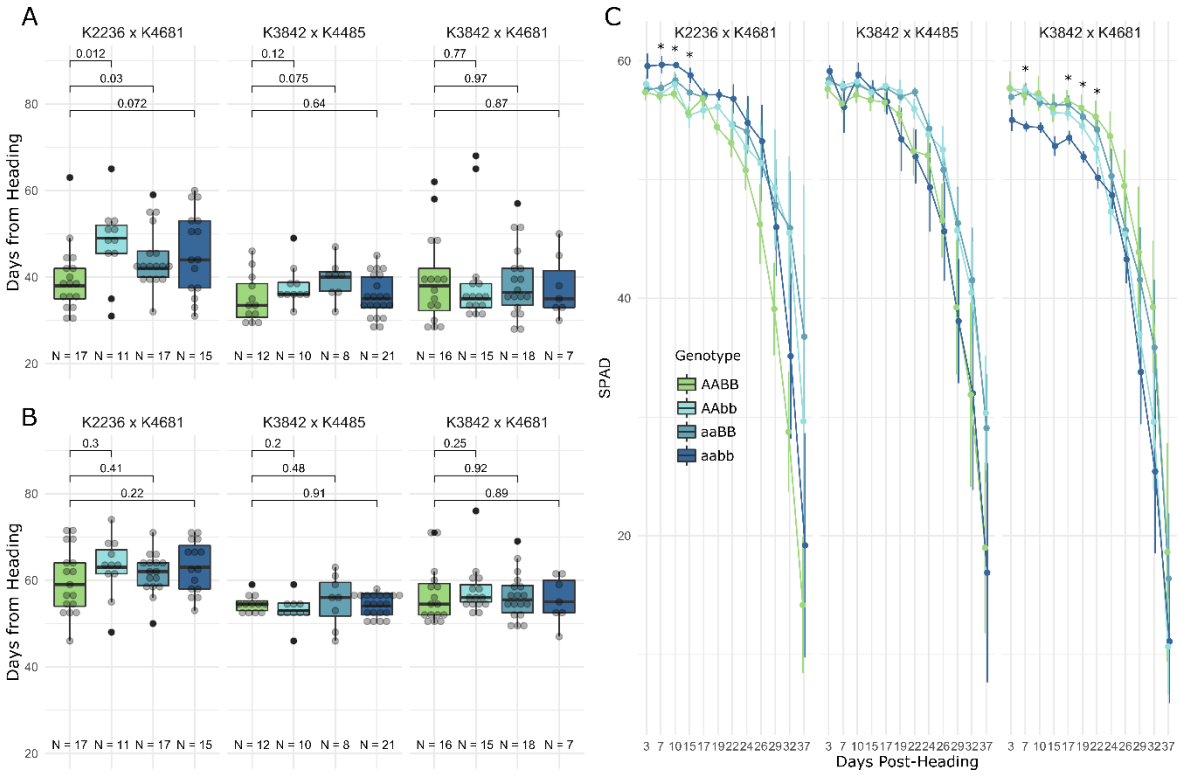

Figure 5-18: Delayed leaf senescence is only retained in the K2236 x K4681 cross in the BC2F2 generation. Plants were scored for flag leaf (A) and peduncle (B) senescence onset, reported as days post-heading. Leaf greenness was also measured using a SPAD meter for five plants from each genotype at intervals following heading (C). All statistical comparisons were carried out using the pairwise Wilcoxon rank-sum test. Only comparisons between wild-type (AABB) and double mutant (aabb) plants are shown in (C); *, $p < 0.05$.

A similar situation is seen for the peduncle senescence, though the strength of the delayed senescence phenotype in the K2236 x K4681 line is substantially reduced. None of the single or double mutants for this line are significantly delayed in peduncle senescence onset, though they do show delays on average compared to wild-type of up to three days (Table 5-16, Figure 5-18B). The remaining two lines, K3842 x K4485 and K3842 x K4681, show no delays in peduncle senescence compared to wild-type (Figure 5-18B, Table 5-16).

Table 5-16: Summary statistics of senescence onset for the TraesCS2A02G101400 and TraesCS2B02G118500 BC_2F_2 crosses.

Cross	Variable	Genotype	N	Mean	Standard Deviation
K2236 x K4681	Days from Heading to Flag Leaf Senescence	AABB	17	39.59	7.95
		AAbb	11	47.91	9.10
		aaBB	17	44.41	7.10
		aabb	15	45.87	9.75
	Days from Heading to Peduncle Senescence	AABB	17	59.88	7.69
		AAbb	11	63.00	6.97
		aaBB	16	61.56	4.84
		aabb	15	63.07	6.01
K3842 x K4485	Days from Heading to Flag Leaf Senescence	AABB	12	35.00	5.61
		AAbb	10	38.00	4.64
		aaBB	8	39.38	4.47
		aabb	21	35.57	4.88
	Days from Heading to Peduncle Senescence	AABB	12	54.58	2.07
		AAbb	10	53.10	3.28
		aaBB	8	55.25	5.99
		aabb	21	54.33	2.50
K3842 x K4681	Days from Heading to Flag Leaf Senescence	AABB	16	39.38	10.25
		AAbb	15	39.07	11.45
		aaBB	18	38.83	8.25
		aabb	7	37.71	7.25
	Days from Heading to Peduncle Senescence	AABB	16	56.56	6.71
		AAbb	15	57.80	5.78
		aaBB	18	56.22	5.40
		aabb	7	55.57	5.41

In summary, of the three homoeolog pairs we phenotyped for senescence, only one showed a significant and consistent senescence-related phenotype. Mutations in TraesCS2A02G101400 and TraesCS2B02G118500 were significantly delayed in both flag leaf and peduncle senescence, and this was repeated across multiple experiments and following backcrosses (Figure 5-15, Figure 5-17, Figure 5-16). While the senescence phenotype was weaker in the BC_2F_2 generation, these homoeologs remain strong candidates for positive regulators of senescence. A second homoeolog pair, TraesCS5A02G049100 and TraesCS5B02G054200, may also act in senescence regulation as the double mutants did originally show an early senescence phenotype (Figure 5-12). However, this phenotype could not be recovered in further experiments (Figure 5-13). This suggests that either the initial phenotype was an artefact, or that the phenotype itself is highly environmentally variable. Finally, the homoeologs TraesCS2A02G101100 and TraesCS2B02G118200 showed no variation in senescence onset or progression when mutated (Figure 5-10), corresponding with their low expression levels (Figure 5-7, Figure 5-8). Taking these results together, we decided to focus our further work on the TraesCS2A02G101400 and TraesCS2B02G118500 homoeologs, with the

strongest and most consistent senescence phenotype. Following this point, we will refer to these homoeologs as *NAC3-A1* and *NAC3-B1*, respectively.

5.3.5 *NAC3-A1 and NAC3-B1 are part of an expanded clade of NAC transcription factors, the NAC-3 Clade.*

Following the phenotypic results which suggest *NAC3-A1* and *NAC3-B1* are likely to be positive regulators of senescence, we further investigated the family of NAC transcription factors related to these genes. As previously described, two NAC transcription factors with very similar protein sequences to these genes (Figure 5-4) were also found to interact with NAM-B1 (Figure 5-5, Figure 5-6) and could also induce cell death when transiently overexpressed in *N. benthamiana* (Figure 5-9). To identify other closely related NAC transcription factors, we carried out a BLAST originally against the TGAC v1 gene models (Clavijo et al. 2017b), which was later repeated with the RefSeqv1.1 gene models (IWGSC et al. 2018). From this, we identified 12 NAC transcription factors with highly similar sequences (Figure 5-19). This set of closely related NAC transcription factors will be referred to as the NAC3 clade for the remainder of the chapter.

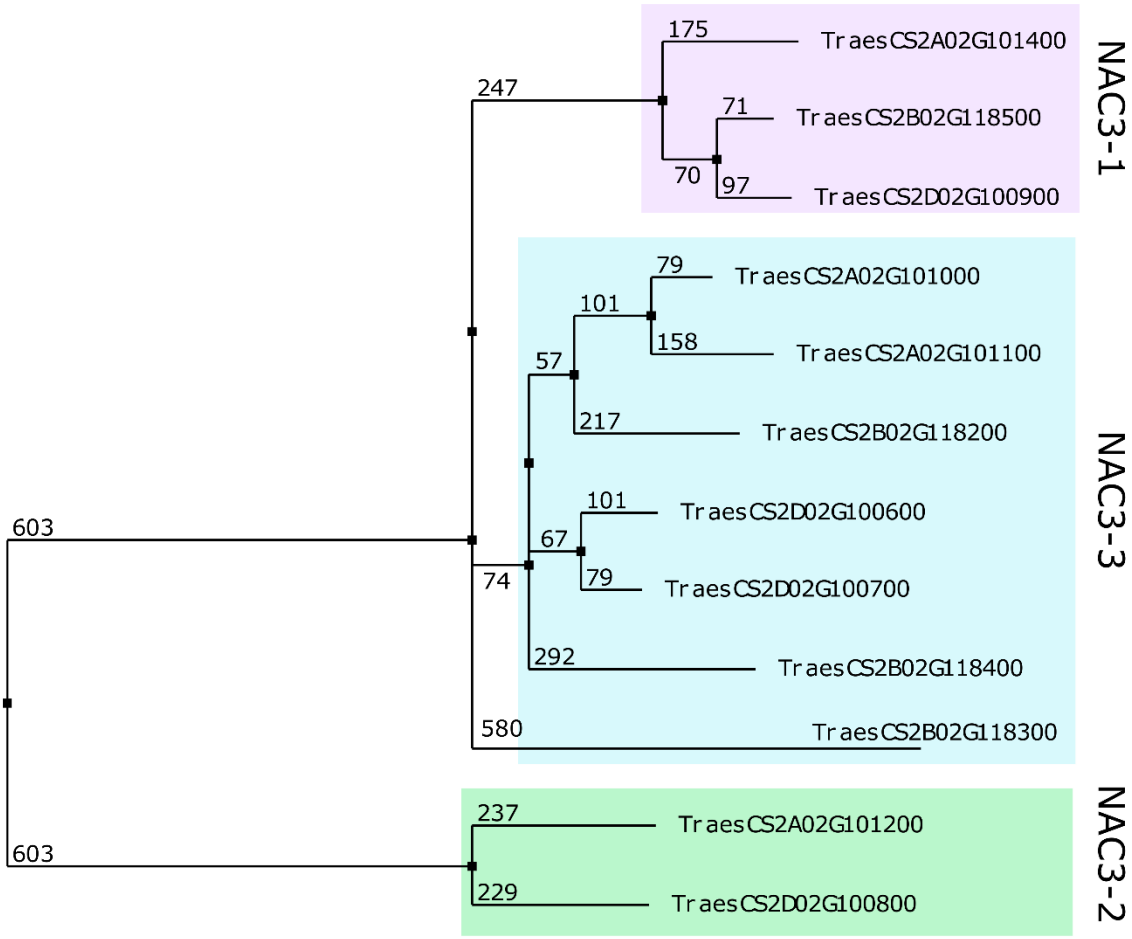

Figure 5-19: Protein sequence phylogeny of the NAC3 clade in Triticum aestivum resolves three distinct sets of genes. Branches are coloured to indicate the three distinct groups within the NAC3 clade— NAC3-1 (purple), NAC3-2 (green), and NAC3-3 (blue). The branch numbers are calculated from the neighbour joining distance matrix as first described in (Saitou and Nei 1987) based on the PAM250 amino acid similarity matrix (Schwartz and Dayhoff 1978). Lower numbers indicate more similar sequences.

The protein sequences for all 12 NAC transcription factors were aligned using Clustal Omega, from which a Neighbour Joining tree was calculated using the PAM250 amino acid substitution matrix (Figure 5-19). This tree separates the 12 genes into three distinct clusters. One set, coloured purple in Figure 5-19, forms a canonical triad, containing homoeologous copies on the A, B, and D genomes. This triad includes *NAC3-A1* (TraesCS2A02G101400) and *NAC3-B1* (TraesCS2B02G118500), the genes which, when mutated, lead to delayed flag leaf and peduncle senescence. This triad will be referred to as the NAC3-1 triad for the remainder of the chapter.

5.3.5.1 Expression levels vary substantially within the NAC3 Clade

Having identified the full set of 12 closely-related genes within the NAC3 clade, we then explored their expression profiles across development and flag leaf senescence in particular. We had previously found that three of the members of the clade had different expression profiles, with very different levels of expression (Figure 5-7 and Figure 5-8). *NAC3-B1* (TraesCS2B02G118500) was

highly expressed in the late stages of flag leaf senescence, while TraesCS2B02G118200 was hardly expressed at all. This suggested that there may be substantially varying levels of expression across the entire NAC3 clade.

Within the clade, we found that the NAC3-1 triad has the highest levels of expression at various stages of development, particularly during senescence. Each of the homoeologs is upregulated in senescence by between 4.3 and 5.2 times compared to their expression levels throughout the rest of the developmental timecourse (Figure 5-20). By the ripening stage of the Azhurnaya data set (Ramírez-González et al. 2018), expression of the A, B, and D homoeologs reaches on average 22, 38, and 36 TPM, respectively. The three homoeologs show similar patterns of expression during the progression of flag leaf senescence, becoming highly upregulated at 23 days after anthesis (DAA) (Figure 5-21). The homoeologs also show slight increases in expression earlier during senescence, at seven and 15 DAA, but the major change in expression levels appears to be at the later stage of senescence.

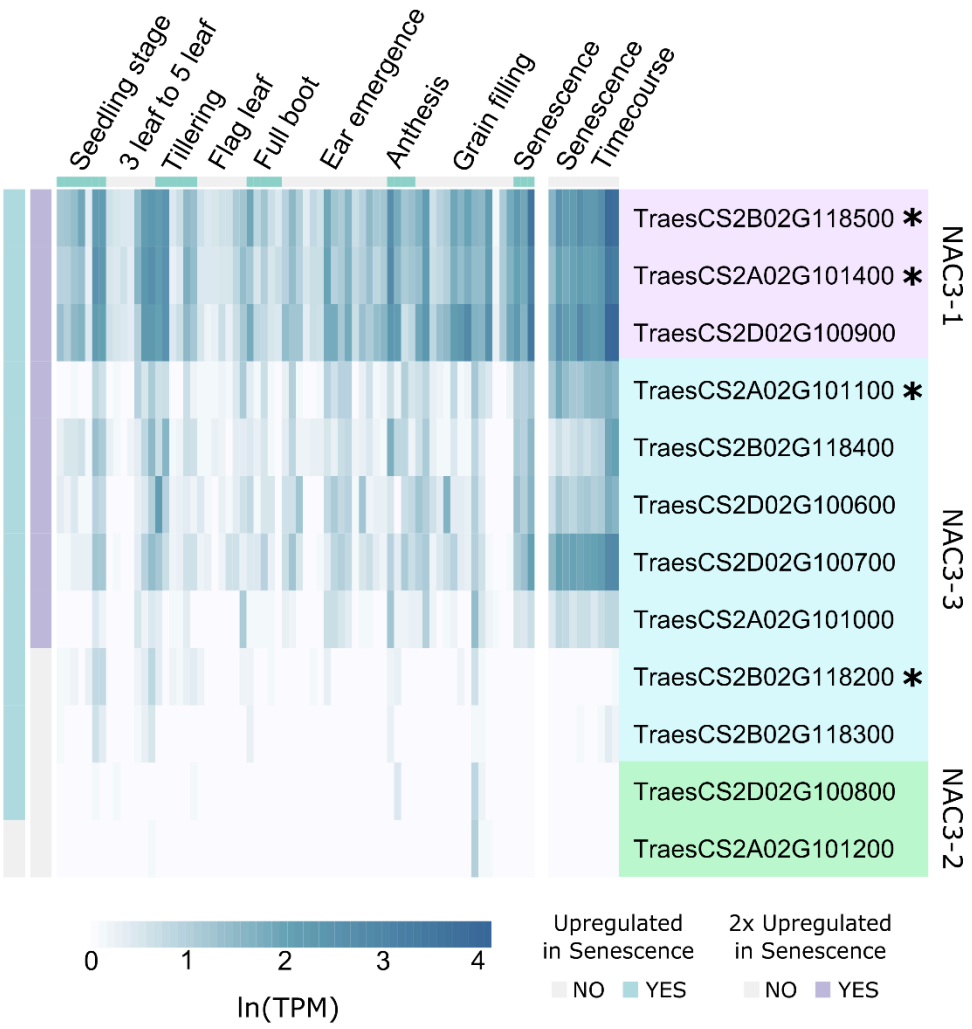

Figure 5-20: Expression profiles of the NAC transcription factors in the NAC3 clade. Gene expression is presented as the logarithm of the TPM derived from the cv. Azhurnaya developmental timecourse (Ramírez-González et al. 2018) and the cv. Bobwhite flag leaf senescence timecourse (Borrill et al. 2019a). The NAC3 genes are split into three subgroups, corresponding to those identified in Figure 5-19. The average was taken from the three replicates of each timepoint and presented here. The natural logarithm (ln) transformation of the expression data was necessary due to the substantial variation in expression levels within the clade. Genes which are two times (2x) upregulated in the "Senescence" stage of the Azhurnaya timecourse, compared to the average expression level across the rest of the timecourse, are marked with purple, while those which are upregulated to any extent (> 0) are marked with light blue. Genes which can physically interact with NAM-B1 are indicated with an asterisk.

In contrast to the complete triad, the remaining members of the NAC3 clade show comparatively low levels of expression throughout development. The pair of genes which cluster separately, TraesCS2A02G101200 and TraesCS2D02G100800, hereafter referred to as the NAC3-2 pair, have the lowest levels of expression of the clade, only showing low expression (~1-2 TPM) in the grain at the milk grain stage (Figure 5-20). They also have essentially no expression during the flag leaf senescence timecourse (Figure 5-21). The remaining members of the NAC3 clade form an unresolved cluster, the NAC3-3 group, containing multiple copies on each genome- two on the A, three on the B, and a further two on the D genome. These genes show more variable patterns of

expression across the developmental timecourse, with some such as TraesCS2D02G100700 also showing relatively high levels of expression (7 TPM) in the flag leaf blade during ripening and senescence. Like the highly-expressed NAC3-1 triad, this gene, along with TraesCS2B02G118400 and TraesCS2D02G100600, also shows an increase in expression at 23 DAA in the flag leaf senescence timecourse (Figure 5-21).

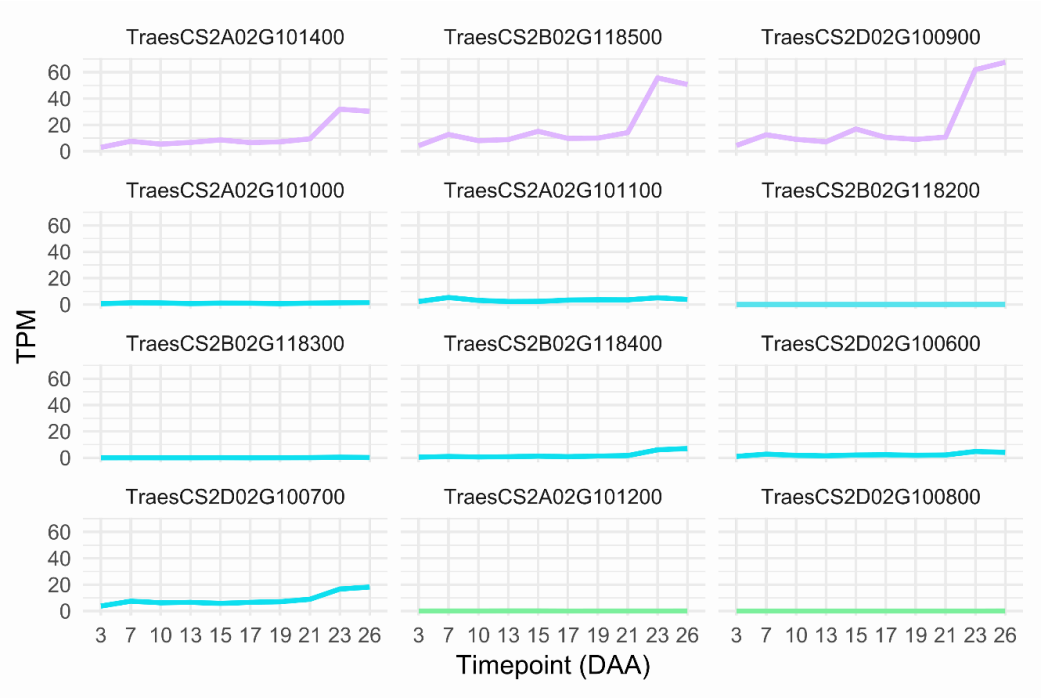

Figure 5-21: The complete NAC3-1 triad is highly expressed late in flag leaf senescence. Expression values are derived from a timecourse of senescence in the flag leaf (Borrill et al. 2019a), presented as transcripts per million (TPM). The average value of three replicates at each timepoint is presented here. The progression of expression is shown as days after anthesis (DAA). Colours indicate the NAC3 group the gene is a member of—NAC3-1 (purple), NAC3-2 (green), and NAC3-3 (blue).

5.3.5.2 The NAC3 clade is present in the Triticeae in a syntenic cluster on Chromosome 2

To ascertain whether the NAC3 clade represents an expansion of genes unique to wheat, we investigated whether orthologs to the NAC3 clade are also present in multiple copies in the rest of the Triticeae. Using the EnsemblPlants gene tree, developed using the Gene Orthology/Paralogy prediction pipeline (Vilella et al. 2009), we identified the predicted orthologues of the NAC3 clade in the Triticeae (Figure 5-22). We found that all A genome members of the NAC3 clade have orthologous copies in both *Triticum urartu* (Ling et al. 2018), the A-genome progenitor of wheat, and in *Triticum turgidum* subsp. *dicoccoides*, or emmer wheat (Avni et al. 2017), the wild tetraploid grass which was domesticated into durum wheat and was the source of the A and B genomes in hexaploid bread wheat. Most B-genome members of the NAC3 clade were also represented by orthologs in the emmer wheat genome, with the exception of TraesCS2B02G118300. Only one copy was identified by the Ensembl pipeline in *Aegilops tauschii* (Luo et al. 2017), orthologous to TraesCS2D02G100900, the D genome member of the NAC3-1

triad. We also found orthologous copies of the NAC3-1 triad and the NAC3-2 pair in the *Hordeum vulgare*, or barley, genome (Mascher et al. 2017), as well as an independent triplet of genes which did not cluster within any of the three wheat clusters. Only a single ortholog was identified in *Brachypodium distachyon*, *Leersia perrieri* and across the *Oryza* genus, suggesting that the expansion of the NAC3 clade occurred exclusively in the Triticeae.

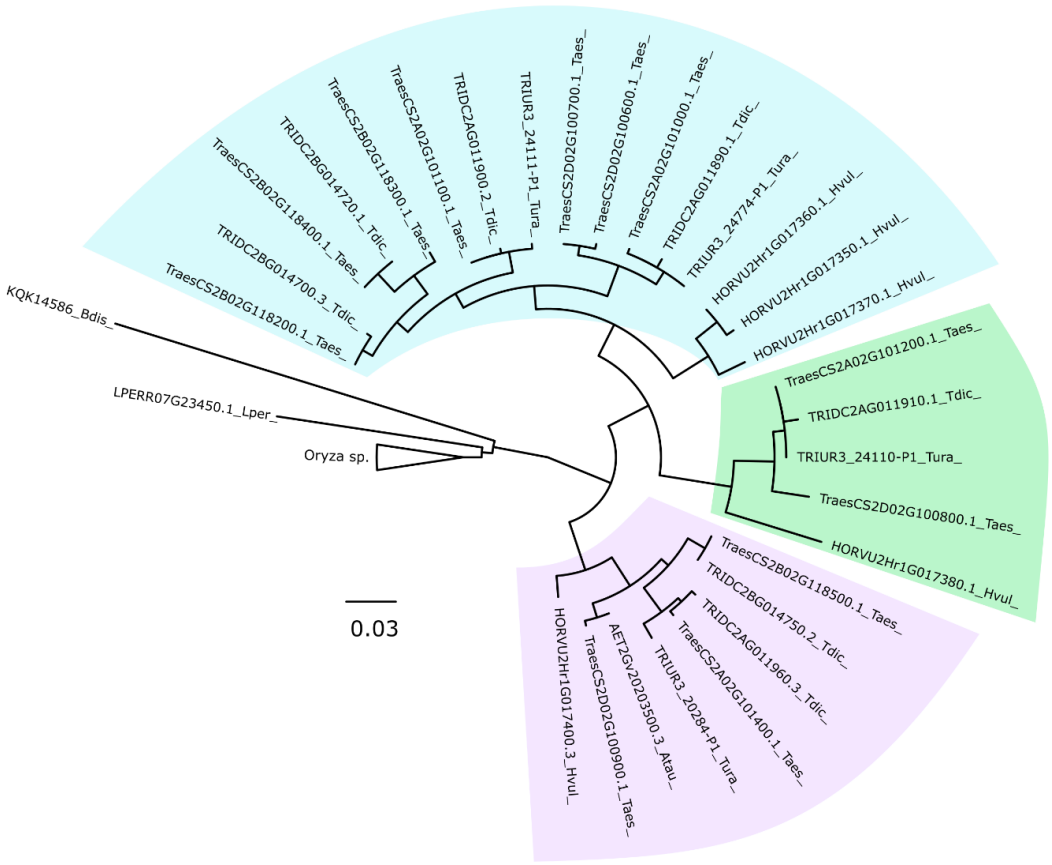

Figure 5-22: Phylogeny of NAC3 genes from the Triticeae, highlighting the existence of three distinct clades conserved across species. Background colours of each clade match those distinguishing the clades in the previous protein phylogeny figure. The phylogenetic relationships were obtained from the EnsemblPlants gene tree for TraesCS2B02G118500, based on a maximum likelihood phylogenetics gene trees approach (Vilella et al. 2009). Tips are labelled with the gene name, as used in EnsemblPlants, followed by an abbreviation of the species name: Taes, T. aestivum; Tdic, T. dicoccoides; Tura, *T. urartu*; Atau, *A. tauschii*, Hvul, *H. vulgare*, Bdis, *B. distachyon*, Lper, *L. perrieri*. The *Oryza* genus is compressed into a single branch ("Oryza sp.").

We also observed that the wheat members of the NAC3 clade are located in close physical proximity to each other on each of the genomes. On all three genomes, the four annotated members of the NAC3 clade are located in syntenic clusters. The most recent RefSeqv1.1 gene annotation incorporates the physical position of genes into their names, whereby gene TraesCS1A02G000002 is positioned in between TraesCS1A02G000001 and TraesCS1A02G000003. From this, we can see that the four NAC3 genes on the B and D genomes are located in a single cluster on each genome, to the exclusion of any other genes (Figure 5-23). The genes on the A genome appear to sit in a similar pattern, though a gap appears between TraesCS2A02G101200 and TraesCS2A02G101400.

Further investigation revealed that, while a small, single exon gene was originally annotated between the two genes in the v1.0 gene annotation, as TraesCS2A01G101300, this gene model was removed in the v1.1 annotation. As a result, it appears that across all three genomes the NAC3 genes are present in tandem arrays to the exclusion of other genes (Figure 5-23).

We also investigated whether a B-genome homoeolog of the NAC3-2 triad existed in the *T. aestivum* RefSeqv1.0 genome but had not been identified in the original annotation. We identified a likely homoeolog in the TGAC v1 gene models, TRIAE_CS42_U_TGACv1_641123_AA2086190, based on best BLAST hits to the A and D homoeologs. However, there was no corresponding annotated gene in the RefSeqv1.0 or 1.1 gene models. We then manually annotated the corresponding region of the RefSeqv1.0 genome, from position 83210464 to 83212468 on chromosome 2B. Based on conservation with the NAC3 clade, we found that the shared start codon of the putative NAC3-2B gene was mutated and lost leading to the likely use of a later start codon, and a seven-residue deletion. The splice donor site at the beginning of exon 2 was also mutated, and the next two potential splice sites both led to frameshifts and premature stop codons early in exon 2. As a result, the putative *NAC3-2B* gene appears to be a pseudogene, and as such is depicted in a light green colour in Figure 5-23.

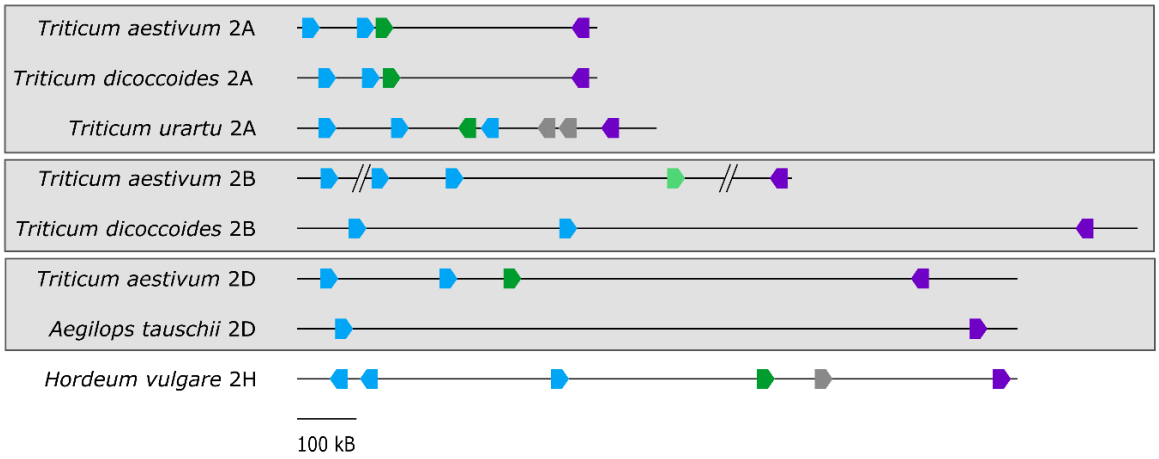

Figure 5-23: The NAC3 clade is expanded across a small region of chromosome 2 in the Triticeae. This schematic depicts the syntenic clustering of the NAC3 genes in wheat and their orthologs within the Triticeae. Genes are coloured according to their sub-groupings- the NAC3-1 triad (purple), the NAC3-2 pair (green) and the remaining NAC3 genes (blue). The barley (*Hordeum vulgare*) genes which grouped separately in Figure 5-22 are coloured in light grey. The pseudogene *NAC3-2B* is shown in light green. Genes which are not members of the NAC3 clade are coloured in dark grey. The *Triticum urartu* genes and gene positions are derived from the 2018 release of the genome (Ling et al. 2018) which is not yet in use on Ensembl and thus differs from that shown in Figure 5-22. Approximate distances between genes are shown to scale, though the genes themselves are not. Chromosome 2B for *Triticum aestivum* was truncated due to the length of the region; the gaps indicated by the double lines are approximately 350 Kb and 750 Kb, respectively.

We then investigated whether the expansion of the NAC3 genes in other members of the Triticeae was also due to tandem duplications. Using the most up-to-date genomes and gene annotations, we were able to determine the relative physical positions of the NAC3 orthologs present in *T.*

dicoccoides (emmer wheat) (Avni et al. 2017), *T. urartu* (Ling et al. 2018), *A. tauschii* (Luo et al. 2017), and *H. vulgare* (Mascher et al. 2017). In all cases, the NAC3 orthologs sit in close proximity to each other on the equivalent chromosome 2 (Figure 5-23). The size of the region encompassing the NAC3 orthologs varies, from 500 Kb in the A genome of *T. dicoccoides* to 1,400 Kb on its B genome (Table 5-17). In nearly all cases, with the exception of *H. vulgare* and *T. urartu*, the NAC3 genes are present in a tandem array to the exclusion of any other genes. A full list of the NAC3 genes in the Triticeae is provided in Table 5-18.

Table 5-17: Summary details of regions in the Triticeae containing orthologs of the NAC-3 clade.

Species	Chromosome	Region Coordinates (nearest 100 Kb)	Size of Region	Genes in Region	*NAC3* genes in region
Triticum aestivum	2A	54,000,000-54,500,000	500 Kb	4	4
	2B	82,300,000-84,000,000	1,700 Kb	5[1]	5[1]
	2D	52,600,000-53,800,000	1,200 Kb	4	4
Triticum dicoccoides	2A	56,500,000-57,000,000	500 Kb	5[2]	4[2]
	2B	89,800,000-91,200,000	1,400 Kb	3	3
Triticum urartu	2	47,600,000-48,200,000	600 Kb	7	5[3]
Aegilops tauschii	2D	53,500,000-54,700,000	1,200 Kb	2	2
Hordeum vulgare	2H	40,800,000-42,000,000	1,200 Kb	6	5

[1]This includes pseudogene TraesCS2B02G118450 which is missing from the RefSeqv1.1 annotation
[2]Two genes which are annotated separately in the *T. dicoccoides* genome (TRIDC2AG011910 and TRIDC2AG011920) each encode half of a full NAC transcription factor in the NAC-3 clade; these two genes are represented as a single "NAC3" gene in the table, and in Figure 5-23.
[3]Two genes which are annotated separately in the *T. urartu* genome (TuG1812G0200001021 and TuG1812G0200001022) encode the same gene.

The conserved synteny of the NAC3 genes within the Triticeae supports the hypothesis that a single gene, present in the common ancestor of the BEP clade (containing the Bambusoideae, Ehrhartoideae, and Pooideae) and the Hordeeae, underwent a series of duplications leading to three tandem genes on chromosome 2. Each of the identified NAC3 groups, NAC3-1, NAC3-2, and NAC3-3, is derived from one of those original three genes. After the split between barley and the Triticineae, further expansions occurred within the NAC3-3 clade (blue in Figure 5-23). These independent expansions contributed to the variable numbers of NAC3 genes present in the Triticeae genomes. At the same time, the process of pseudogenisation appears to have been taking place. We see signatures of this within the wheat NAC3 genes, where the B genome homoeolog of the NAC3-2 group is only present as a gene fragment (Table 5-18). Given this evolutionary history, and the phenotypic data obtained in the previous sections, we decided to move forward on the assumption that the comparatively stable NAC3-1 triad is most likely to regulate the highly controlled process of senescence. In the remainder of this chapter, we will address the predicted role of the NAC3-1 triad in positively regulating senescence.

Table 5-18: All NAC3 genes present in the Triticeae on the group 2 chromosomes.

Species	Chromosome	Gene	Start	End	Orientation	NAC3 Group	Notes
Triticum aestivum	2A	TraesCS2A02G101000	54,012,215	54,014,046	+	NAC3-3	
Triticum aestivum	2A	TraesCS2A02G101100	54,111,812	54,113,406	+	NAC3-3	
Triticum aestivum	2A	TraesCS2A02G101200	54,139,933	54,141,073	+	NAC3-2	
Triticum aestivum	2A	TraesCS2A02G101400	54,471,565	54,474,879	-	NAC3-1	
Triticum aestivum	2B	TraesCS2B02G118200	82,361,173	82,362,952	+	NAC3-3	
Triticum aestivum	2B	TraesCS2B02G118300	82,713,779	82,715,637	+	NAC3-3	
Triticum aestivum	2B	TraesCS2B02G118400	82,851,554	82,853,290	+	NAC3-3	
Triticum aestivum	2B	TraesCS2B02G118450	83,210,464	83,212,468	+	NAC3-2	Not annotated; pseudogene
Triticum aestivum	2B	TraesCS2B02G118500	83,953,621	83,955,548	-	NAC3-1	
Triticum aestivum	2D	TraesCS2D02G100600	52,646,885	52,648,518	+	NAC3-3	
Triticum aestivum	2D	TraesCS2D02G100700	52,851,898	52,853,720	+	NAC3-3	
Triticum aestivum	2D	TraesCS2D02G100800	52,959,503	52,960,664	+	NAC3-2	
Triticum aestivum	2D	TraesCS2D02G100900	53,648,855	53,650,859	-	NAC3-1	
Hordeum vulgare	2H	HORVU2Hr1G017350	40,851,300	40,853,415	-	NAC3-3	
Hordeum vulgare	2H	HORVU2Hr1G017360	40,902,535	40,902,992	-	NAC3-3	Possibly truncated NAC
Hordeum vulgare	2H	HORVU2Hr1G017370	41,249,925	41,252,008	+	NAC3-3	
Hordeum vulgare	2H	HORVU2Hr1G017380	41,571,285	41,572,514	+	NAC3-2	
Hordeum vulgare	2H	HORVU2Hr1G017390	41,642,477	41,643,166	+	NA	
Hordeum vulgare	2H	HORVU2Hr1G017400	41,955,172	41,958,298	+	NAC3-1	
Triticum urartu	2A	TuG1812G0200001021	47,643,980	47,645,437	+	NAC3-3	Encode same gene
Triticum urartu	2A	TuG1812G0200001022	47,643,929	47,645,343	+	NAC3-3	
Triticum urartu	2A	TuG1812G0200001023	47,779,445	47,780,953	+	NAC3-3	
Triticum urartu	2A	TuG1812G0200001024	47,897,044	47,897,968	-	NAC3-2	

Species	Chromosome	Gene	Start	End	Orientation	NAC3 Group	Notes
Triticum urartu	2A	TuG1812G0200001025	47,913,156	47,914,853	-	NAC3-3	
Triticum urartu	2A	TuG1812G0200001026	48,137,715	48,139,055	-	NA	
Triticum urartu	2A	TuG1812G0200001027	48,139,473	48,141,769	-	NA	
Triticum urartu	2A	TuG1812G0200001028	48,170,411	48,172,348	-	NAC3-1	
Aegilops tauschii	2D	AET2Gv20202900	53,586,613	53,588,544	+	NAC3-3	
Aegilops tauschii	2D	AET2Gv20203500	54,628,732	54,630,828	+	NAC3-1	
Triticum dicoccoides	2A	TRIDC2AG011890	56,549,187	56,550,867	+	NAC3-3	
Triticum dicoccoides	2A	TRIDC2AG011900	56,613,093	56,615,141	+	NAC3-3	
Triticum dicoccoides	2A	TRIDC2AG011910	56,630,853	56,631,387	+	NAC3-2	Same gene, first half
Triticum dicoccoides	2A	TRIDC2AG011920	56,631,451	56,632,353	+	NAC3-2	Same gene, second half
Triticum dicoccoides	2A	TRIDC2AG011960	56,971,823	56,973,711	-	NAC3-1	
Triticum dicoccoides	2B	TRIDC2BG014700	89,849,450	89,898,242	+	NAC3-3	
Triticum dicoccoides	2B	TRIDC2BG014720	90,260,285	90,262,182	+	NAC3-3	
Triticum dicoccoides	2B	TRIDC2BG014750	91,102,344	91,104,247	-	NAC3-1	

5.3.6 *Predicted targets of NAC3-1 triad are enriched for senescence-related GO terms.*

Having shown that the NAC3-1 triad is highly expressed during senescence, and that mutations in the A and B genome homoeologs lead to a delay in flag leaf and peduncle senescence, we next began investigating the pathways through which NAC3-1 regulates senescence. To do this, we utilised a GENIE3 network developed from 850 independent RNA-Seq samples covering a wide range of developmental, tissue, and stress conditions (Ramírez-González et al. 2018). This network was created using an algorithm that applies a Random Forests approach to predict the strength of possible regulatory links between transcription factors and target genes (Huynh-Thu et al. 2010). The algorithm considers the similarity of the gene expression profiles between the transcription factors and putative target genes, producing a ranked list of each pairwise comparison where the higher ranked pairs are most likely to represent true regulatory links.

5.3.6.1 The GENIE3 network provides biologically relevant predictions of transcription factor targets.

Before carrying out our analyses using the GENIE3 network, we needed to demonstrate that the predicted targets in the network contained biologically relevant information. To do this, we compared the differentially expressed genes between wild-type and *nam-a1* mutant plants during senescence with the predicted targets of *NAM-A1* in the GENIE3 network (Harrington et al. 2019a). We calculated the level of overlap, or "shared ratio," between different sets of genes as follows:

$$\frac{|A \cap B|}{|A|}$$

Where A and B are sets of genes, and |A| < |B|. For example, given two sets of genes A and B, where A contains 5 genes and B contains 10 genes, if they share two genes in common the shared ratio is 2/5, or 0.4. We found that the differentially expressed genes overlapped significantly more than expected by chance with the GENIE3 targets both in early (12 DAA) and late (22 DAA) flag leaf senescence ($p < 0.001$, Sign Test; Figure 5-24A and B). This demonstrated that the GENIE3 network provides biologically relevant predictions for targets of *NAM-A1*, a positive regulator of senescence.

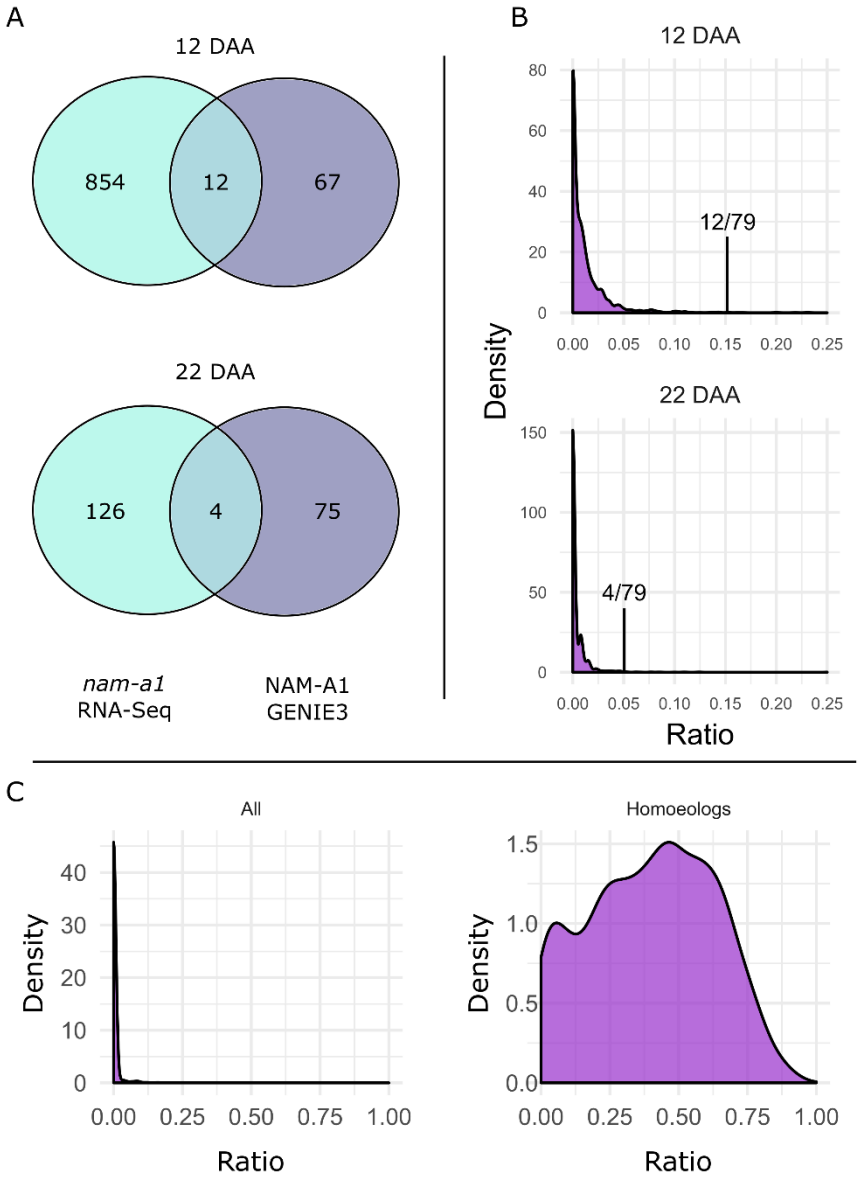

Figure 5-24: GENIE3 targets and differentially expressed genes of *NAM-A1* overlap, and the GENIE3 network identifies shared targets between homoeologs. (A) Overlapping genes between the *nam-a1* RNA-Seq dataset and the predicted targets of *NAM-A1* are identified at 12 DAA (15.2% of the GENIE3 targets) and 22 DAA (5.1%). (B) A randomly selected transcription factor shares very few predicted targets with the *nam-a1* differentially expressed genes, with a distribution of 1000 pairwise comparisons skewed strongly towards 0. At both 12 and 22 DAA, the predicted targets of *NAM-A1* overlap significantly more with the *nam-a1* DEG than would be expected by chance (Sign Test, $p < 0.001$). Note that the x-axis is capped at 0.25, to emphasize the skew of the distributions towards zero. (C) Homoeologs share substantially more targets in common than any two randomly selected transcription factors, with an average shared ratio of 0.4. "Ratio" refers to the shared ratio of targets between the DEGs and the GENIE3 transcription factors. The figure and figure caption were adapted from Harrington et al. (2019a), under CC-BY-NC 4.0.

5.3.6.2 The NAC3 triad share predicted targets

We next used the GENIE3 network to investigate the putative function of the NAC3-1 triad. We found that the NAC3-1 homeologs have a high level of overlap between their respective targets, as expected for homoeologs (Figure 5-24C) (Harrington et al. 2019a). 50 target genes are shared

between all three of the homoeologs, 38 between the A and B homoeologs, 27 between the A and D homoeologs, and 22 between the B and D homoeologs (Figure 5-25). Overall, the *NAC3-A1* homoeolog has the largest number of predicted targets, with 1123, while *NAC3-B1* has 256 and *NAC3-D1* has 157. The list of the top 100 predicted targets for each of the three homoeologs is available in Appendix 8.3.4. Notably, each homoeolog is predicted to regulate the other two members of the triad. It may be the case that the NAC3-1 genes do regulate each other. However, it is also possible that these connections are solely an artefact of their similar expression patterns. This highlights the fact that the GENIE3 network predictions are based upon expression patterns; as such, predicted connections in the network may be genes which are only co-expressed with the given transcription factor, such as homoeologs, rather than true regulatory targets.

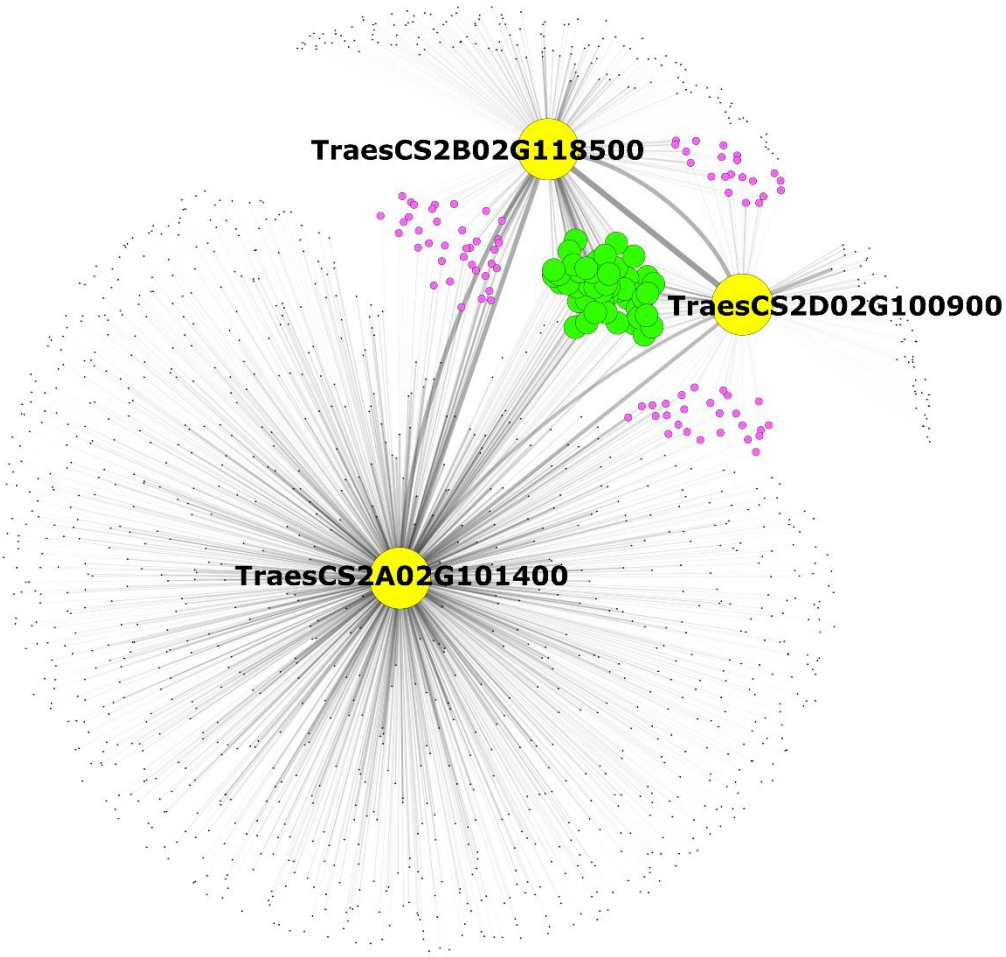

Figure 5-25: The NAC3-1 homoeologs share many predicted targets from the GENIE3 network. The weight of the lines connecting the transcription factors to each gene corresponds to the strength of the predicted connection, where darker lines indicate stronger predictions. Green targets are shared between all three homoeologs; purple targets are shared between two of the three homoeologs.

5.3.6.3 The NAC3 triad targets are enriched for senescence-related GO terms

While it is tempting to investigate the individual genes that are predicted targets of the NAC3-1 triad in the GENIE3 network, it is likely that some, if not many, of these genes are not true targets

of the transcription factors. This is because the GENIE3 network relies solely on expression data to produce predictions of transcription factor targets. As a result, the individual genes predicted by the GENIE3 network may not be directly regulated by the transcription factor in question. However, as we demonstrated previously (Harrington et al. 2019a), a substantial portion of the predicted target of *NAM-A1* were differentially expressed in a *nam-a1* mutant RNA-Seq dataset.

To make the most of the GENIE3 dataset, therefore, we carried out GO-term enrichment analyses on the predicted targets of the NAC3-1 homoeologs. Enrichment analyses such as this allow us to reduce the confounding effect of incorrect predictions by highlighting overall trends supported by larger sets of targets. We can then use these analyses to predict what pathways the NAC3-1 transcription factors may be acting in.

We therefore extracted GO terms which were over-represented amongst the targets of all three homoeologs ($p < 0.05$, adjusted for false discovery rate using the Benjamini-Hochberg correction). The top twenty over-represented GO terms for each homoeolog are reported in the tables below (Table 5-19, Table 5-20, and Table 5-21). Overall, we found that the targets of all NAC3-1 homoeologs are enriched for senescence-related GO terms. *NAC3-A1* is highly enriched for catabolism-related terms, including protein, chlorophyll, and nucleotide breakdown. *NAC3-B1* is enriched explicitly for senescence terms, including floral organ and leaf senescence and programmed cell death. It is also enriched for protein and chlorophyll catabolism, as is the D-homoeolog *NAC3-D1*.

Table 5-19: The top 20 GO terms enriched in the GENIE3 targets of *NAC3-A1*. "BP" refers to Biological Process GO terms, "MF" to Molecular Function, and "CC" to Cellular Compartment. "N" refers to the number of genes present in the GENIE3 targets which are associated with the given GO term.

GO term	**Description**	**Ontology**	**P- value (adj)**	**N**
GO:0006511	ubiquitin-dependent protein catabolic process	BP	2.20E-07	43
GO:2000083	negative regulation of L-ascorbic acid biosynthetic process	BP	1.48E-06	14
GO:0016844	strictosidine synthase activity	MF	9.21E-06	7
GO:0004842	ubiquitin-protein transferase activity	MF	7.45E-05	39
GO:0043765	T/G mismatch-specific endonuclease activity	MF	0.000217	5
GO:0015996	chlorophyll catabolic process	BP	0.000217	10
GO:0045910	negative regulation of DNA recombination	BP	0.000296	6
GO:0009821	alkaloid biosynthetic process	BP	0.000306	5
GO:0004222	metalloendopeptidase activity	MF	0.000374	10
GO:0016567	protein ubiquitination	BP	0.000382	47
GO:0007275	multicellular organism development	BP	0.00077	31
GO:0005787	signal peptidase complex	CC	0.001038	4

GO term	Description	Ontology	P- value (adj)	N
GO:0000014	single-stranded DNA endodeoxyribonuclease activity	MF	0.001118	5
GO:0006308	DNA catabolic process	BP	0.001764	5
GO:0033897	ribonuclease T2 activity	MF	0.002788	4
GO:0016891	endoribonuclease activity, producing 5'-phosphomonoesters	MF	0.003126	5
GO:0016788	hydrolase activity, acting on ester bonds	MF	0.003165	15
GO:0090502	RNA phosphodiester bond hydrolysis, endonucleolytic	BP	0.003579	9
GO:0010304	PSII associated light-harvesting complex II catabolic process	BP	0.003898	6
GO:0007140	male meiosis	BP	0.005132	6

Table 5-20: The top 20 GO terms enriched in the GENIE3 targets of *NAC3-B1*. "BP" refers to Biological Process GO terms, "MF" to Molecular Function, and "CC" to Cellular Compartment. "N" refers to the number of genes present in the GENIE3 targets which are associated with the given GO term.

GO term	Description	Ontology	P- value (adj)	N
GO:0015996	chlorophyll catabolic process	BP	7.92E-09	12
GO:0080187	floral organ senescence	BP	3.83E-08	9
GO:1903647	negative regulation of chlorophyll catabolic process	BP	2.55E-05	3
GO:0006572	tyrosine catabolic process	BP	2.55E-05	5
GO:0004587	ornithine-oxo-acid transaminase activity	MF	2.55E-05	3
GO:0006593	ornithine catabolic process	BP	2.55E-05	3
GO:0010121	arginine catabolic process to proline via ornithine	BP	2.55E-05	3
GO:0032441	pheophorbide a oxygenase activity	MF	2.55E-05	3
GO:0003868	4-hydroxyphenylpyruvate dioxygenase activity	MF	8.46E-05	3
GO:0007275	multicellular organism development	BP	0.000208	22
GO:0009507	chloroplast	CC	0.000208	69
GO:0010277	chlorophyllide a oxygenase [overall] activity	MF	0.000208	4
GO:0006570	tyrosine metabolic process	BP	0.000226	3
GO:0009415	response to water	BP	0.000305	5
GO:0043765	T/G mismatch-specific endonuclease activity	MF	0.000305	4
GO:0035265	organ growth	BP	0.000344	10
GO:0010150	leaf senescence	BP	0.000359	15
GO:0010282	senescence-associated vacuole	CC	0.000438	5

GO term	Description	Ontology	P- value (adj)	N
GO:0010260	animal organ senescence	BP	0.000639	7
GO:0010623	programmed cell death involved in cell development	BP	0.000639	5

Table 5-21: The top 20 GO terms enriched in the GENIE3 targets of *NAC3-D1*. "BP" refers to Biological Process GO terms, "MF" to Molecular Function, and "CC" to Cellular Compartment. "N" refers to the number of genes present in the GENIE3 targets which are associated with the given GO term.

GO term	Description	Ontology	P- value (adj)	N
GO:0015996	chlorophyll catabolic process	BP	3.74E-09	10
GO:0004587	ornithine-oxo-acid transaminase activity	MF	6.03E-06	3
GO:0006593	ornithine catabolic process	BP	6.03E-06	3
GO:0010121	arginine catabolic process to proline via ornithine	BP	6.03E-06	3
GO:0006570	tyrosine metabolic process	BP	7.33E-05	3
GO:0019544	arginine catabolic process to glutamate	BP	0.000335	3
GO:0016491	oxidoreductase activity	MF	0.000376	24
GO:0006559	L-phenylalanine catabolic process	BP	0.000475	5
GO:0034214	protein hexamerization	BP	0.001395	3
GO:0055114	oxidation-reduction process	BP	0.002591	26
GO:0004053	arginase activity	MF	0.002591	2
GO:0008783	agmatinase activity	MF	0.002591	2
GO:0009445	putrescine metabolic process	BP	0.002591	2
GO:1903647	negative regulation of chlorophyll catabolic process	BP	0.003072	2
GO:0004043	L-aminoadipate-semialdehyde dehydrogenase activity	MF	0.003072	2
GO:0080124	pheophytinase activity	MF	0.003072	2
GO:0009695	jasmonic acid biosynthetic process	BP	0.003473	8
GO:0016629	12-oxophytodienoate reductase activity	MF	0.004029	3
GO:0006572	tyrosine catabolic process	BP	0.004429	3
GO:0003868	4-hydroxyphenylpyruvate dioxygenase activity	MF	0.004725	2

These enrichment results corroborate our hypothesis that the NAC3 triad is involved in regulating senescence, particularly as positive regulators. However, having gained insight into the overall function of the NAC3 triad, we wanted to identify candidate downstream targets of the transcription factors. While, as mentioned previously, the GENIE3 network is likely to contain many spurious predictions of targets, we hypothesized that targets shared across all three homoeologs of the NAC3 triad were more likely to have a biologically relevant role in senescence.

To investigate this, we extracted the list of the 50 genes which were shared between each of the homoeologs (Appendix 8.3.2, Table 8-2). Of the 50 genes, only seven were transcription factors, all of which were NAC transcription factors. We calculated the fold change in expression that the genes exhibit during senescence and found that 32 of the 50 genes were upregulated more than two-fold. For the 36 genes which were included in the senescence networks previously developed (Borrill et al. 2019a), we also identified the timepoint at which they first became up or down regulated. All genes included in the networks were initially upregulated during the timecourse. One was upregulated early, at 3 DAA, but the largest set were upregulated at 17 DAA (15/36). One of the other members of the NAC3 clade was identified in the set of shared genes, TraesCS2B02G118400, and was also upregulated first at 17 DAA. We were able to identify the ortholog(s) of 33 of the target genes in *A. thaliana*. 27 of those identified orthologs were expressed at least 75% of their maximum TPM in the senescent leaf in the Klepikova Arabidopsis Expression Atlas (Klepikova et al. 2016), and we listed those as being highly expressed during senescence. A subset of the orthologs are already known to have senescence-related functions in *A. thaliana*, such as *Stay Green 1* (*SGR1*), which balances chlorophyll catabolism with remobilisation of other macromolecules in the senescent leaf (Park et al. 2007), and *ANAC046* itself, orthologous to the NAC3 triad and with a known role in Arabidopsis leaf senescence (Oda-Yamamizo et al. 2016).

5.3.6.4 The NAC3 triad shares predicted targets with its ortholog AtANAC046

Having shown that the predicted targets of the NAC3 triad included many orthologous to Arabidopsis genes with known functions in senescence, we then considered whether we identified any targets which were orthologous with the known targets of the Arabidopsis orthologs *ANAC046* and *ANAC087*. *ANAC046* was found to bind a set of promoters for genes involved in chlorophyll degradation in Arabidopsis (Oda-Yamamizo et al. 2016). These genes were *NYC1* (AT4G13250), a chlorophyll catabolic enzyme (Horie et al. 2009), *SGR1* (AT4G22920) and *SGR2* (At4g11910), magnesium dechetalases (Park et al. 2007, Sakuraba et al. 2014b), and *PaO* (AT3G44880), a pheophorbide a oxygenase (Pruzinska et al. 2003). All four of these genes act at separate stages of the chlorophyll catabolic pathway (Figure 5-26). The ability of ANAC046 to bind the promoters of the other genes listed in the pathway was also tested, but only the four listed above were confirmed.

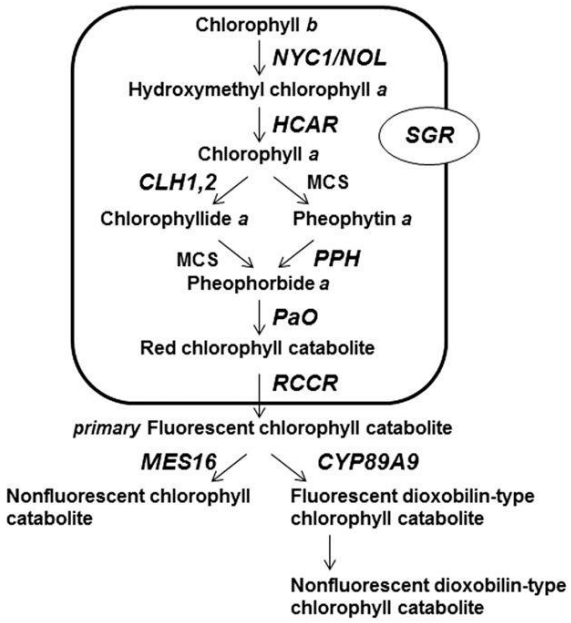

Figure 5-26: A simplified representation of the chlorophyll catabolic pathway in Arabidopsis. Reproduced from Oda-Yamamizo et al. (2016) under a CC-BY-4.0 license**.**

Using the GeneTree function in EnsemblPlants (Vilella et al. 2009), we identified the most closely related wheat genes to each of the four chlorophyll catabolism genes which ANAC046 binds (Table 5-22). We could identify a single wheat triad for three of the four genes – *PaO*, *SGR1*, and *SGR2* – while no orthologs were identified for *NYC1*. We then extracted the protein sequence of *NYC1* and identified the best BLASTp hits against the wheat proteome. This identified a single triad on chromosome 3 with 100% identity over approximately 75% of the sequence. We took this to be our closest ortholog. After identifying the best wheat orthologs, we then searched for these genes in the predicted targets of the NAC3 triad. The triads orthologous to *PaO*, *SGR1* and *SGR2* were found in the targets of the NAC3 triad; indeed, the A and D homoeologs of *SGR1* were predicted targets of all three NAC3 homoeologs, as was the A homoeolog of *PaO* (Table 8-2). However, we did not identify the *NYC1* orthologs in the predicted targets.

Table 5-22: Orthologs of ANAC046 targets in T. aestivum. Orthologs highlighted in green were present in the predicted targets of at least one member of the NAC3-1 triad.

	Genome		
	A	**B**	**D**
NYC1	TraesCS3A01G151900	TraesCS3B01G179100	TraesCS3D01G159800
PaO	TraesCS4A01G411000	TraesCS4B01G311100	TraesCS4D01G309000
SGR1	TraesCS5A01G319900	TraesCS5B01G320200	TraesCS5D01G325900
SGR2	TraesCS5A01G319900	TraesCS5B01G320200	TraesCS5D01G325900

5.3.7 *NAC3-1 shares downstream targets with NAM-A1.*

We also used the GENIE3 network to predict novel transcription factors which may regulate senescence. To do this, we looked for transcription factors which had a higher overlap between their predicted targets and the differentially expressed genes found in the *nam-a1* mutant. We identified 20 unique transcription factors which, based on this criterion, we considered candidate regulators of senescence (Figure 5-27). Of these 20 transcription factors, four were upregulated at least two-fold during senescence (Figure 5-27A) and nine shared at least one predicted target from the GENIE3 network with the predicted targets of *NAM-A1* (Figure 5-27B). Six of the transcription factors had previously been identified (Borrill et al. 2019a) as candidate senescence regulators based on their expression profiles across the flag leaf senescence timecourse, which was not incorporated into the GENIE3 network.

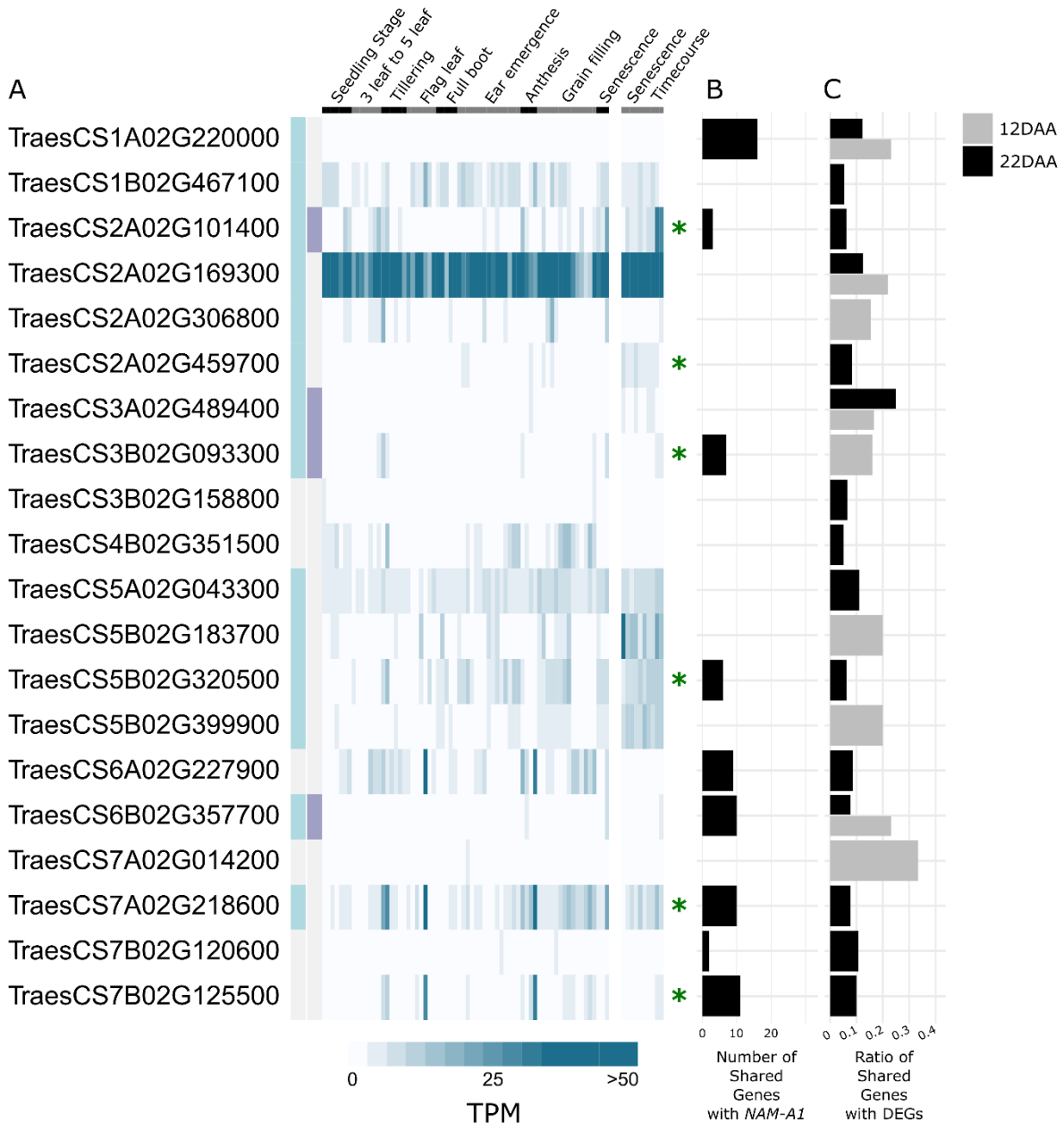

Figure 5-27: Candidate senescence regulators. Twenty transcription factors were identified which had a higher shared ratio between their GENIE3-predicted targets and the *nam-a1* RNA-Seq DEGs than *NAM-A1* itself. (A) Their expression pattern is shown across a developmental time course (Ramírez-González et al. 2018) and a senescence-specific time course (Borrill et al. 2019a). The TPM reported is averaged across three samples of the same tissue and timepoint from each dataset. Genes upregulated in senescence are highlighted with a light blue box, and those with a greater than 2-fold increase are highlighted with a purple box (left side). Genes present in the list of 341 candidate transcription factors based on the senescence time course in Borrill et al. 2019 are indicated with a green asterisk (Borrill et al. 2019a). (B) The number of targets shared between the transcription factors and *NAM-A1*. (C) The shared ratio for the GENIE3 targets of each candidate gene against the 12 DAA (grey) and 22 DAA (black) DEGs. Note that genes which had a higher shared ratio at both 12 and 22 DAA are shown with split bars. This figure was reproduced from Harrington et al. (2019a) under a CC-BY-NC 4.0 license.

Of particular interest was the fact that the A-homoeolog of the NAC3-1 triad, TraesCS2A02G101400, was also present in this list of candidate senescence regulators. At 22 DAA, the list of differentially expressed genes between the wild-type and *nam-a1* flag leaf contained eight genes which were also present in the predicted targets of *NAC3-A1* (6%; $p < 0.001$, Sign Test). Three predicted targets of *NAM-A1* were also shared with the *NAC3-A1* gene. We then investigated whether *NAM-A1* shares predicted targets with the rest of the NAC3-1 triad and found that this was indeed the case. Two of the three targets were shared between *NAC3-A1*, *NAC3-B1*, and *NAM-A1*, while a third target was uniquely shared between *NAC3-A1* and *NAM-A1* (Figure 5-28). A further five targets were shared solely between *NAC3-B1* and *NAM-A1*, which is again significantly more than would be expected purely by chance (6%, $p < 0.001$ Sign Test).

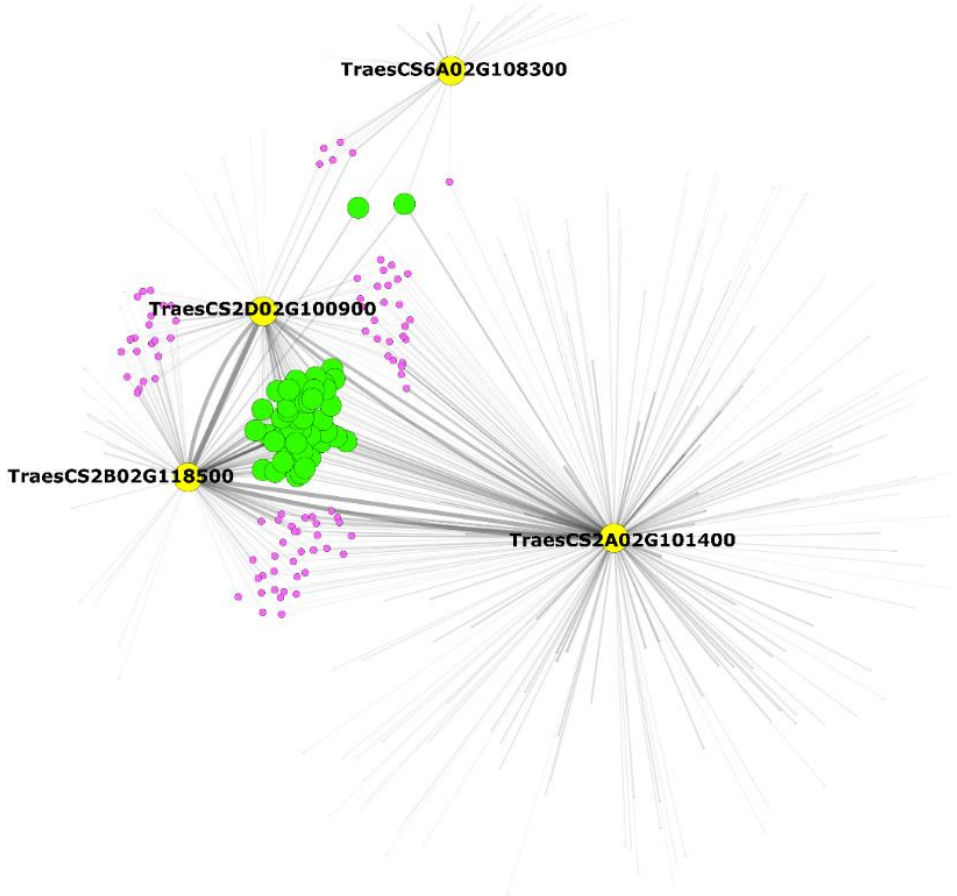

Figure 5-28: The NAC3 triad shares predicted targets with *NAM-A1*. Predicted targets of the four transcription factors, the NAC3-1 homoeologs and *NAM-A1*, are depicted in this network schematic. Targets shared between three transcription factors are shown in green, those shared between two are shown in purple. The width of the lines connecting nodes indicates the strength of the predicted connection.

The fact that members of the NAC3 triad shared targets with *NAM-A1* was particularly interesting, as we had previously found that NAC3-A1 and NAC3-B1 could physically interact with NAM-B1,

the B-homoeolog of *NAM-A1* (Figure 5-5, Figure 5-6, and Figure 5-14). Could these homoeologs also interact with NAM-A1? Using the Y2H system we established earlier, we found that both NAC3-A1 and NAC3-B1 are capable of physically interacting with NAM-A1 (Figure 5-29).

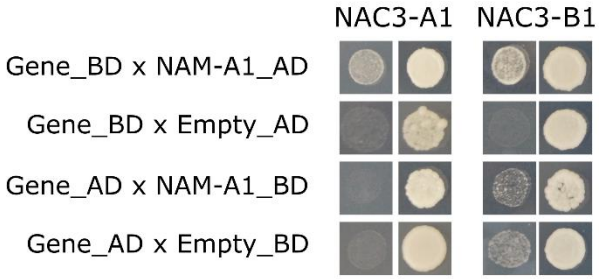

Figure 5-29: NAC3-A1 and NAC3-B1 physically interact with NAM-A1 in a yeast two-hybrid system. Colonies were grown on selective media (SC-LTH + 3AT, left) and control media (SC-LT). Growth in rows 1 and 3 indicates the ability of the protein of interest to interact with NAM-A1. Rows 2 and 4 are controls for the row above; growth in these rows indicates that the presence of the gene alone is sufficient to induce growth irrespective of the presence of NAM-A1. Note that the images shown in rows 2 and 4 were also shown in Figure 5-5 and Figure 5-14 as the experiments were carried out in tandem.

These results suggest that the shared targets identified between the NAC3 triad and *NAM-A1* may be biologically relevant. We extracted the IDs of these eight genes and found that six of the eight are upregulated at least 2-fold in senescence (Table 5-23). Two of the genes are homoeologous, TraesCS4A02G336400 and TraesCS5D02G533700, likely stemming from the 5AL/4AL translocation (Liu et al. 1992). Of the six genes included in the senescence networks (Borrill et al. 2019a), half were first upregulated at 13 DAA, and the other half at 17 DAA, similar in timing to the genes shared between all three NAC3 homoeologs (Table 8-2). Finally, we investigated the Arabidopsis orthologs of the shared genes, and found orthologs for four of the eight genes. Based on expression levels in the senescing leaf, we found that three of the four sets of Arabidopsis orthologues may have roles in Arabidopsis leaf senescence (Table 5-23).

We then looked in detail at the expression profiles of the shared genes in the flag leaf senescence timecourse (Borrill et al. 2019a). We first compared the profiles of the NAC3 triad with that of *NAM-A1* (Figure 5-30). *NAM-A1* is upregulated earlier during senescence, reaching an early peak at around 13 DAA, before rising higher again at 19 and 21 DAA, approaching 30 TPM at 19 DAA. In contrast, the NAC3 triad is stably expressed during the early stages of senescence before peaking sharply at 23 DAA, reaching over 60 TPM in the case of *NAC3-D1*.

Of the shared target genes, most follow a profile more similar to that of the NAC3 triad (Figure 5-30). One gene, TraesCS5A02G230300, is very lowly expressed and will not be considered further. The remaining seven genes all show either flat or slightly increasing expression levels during the early stages of senescence, before rising substantially between either 19 and 21 DAA, or, more commonly, between 21 and 23 DAA.

Table 5-23: Predicted targets shared between the NAC3 triad and *NAM-A1*. Change in expression levels between senescing and non-senescing samples was calculated from the Azhurnaya developmental RNA-Seq dataset (Ramírez-González et al. 2018), as detailed in the methods. The timing of gene upregulation was obtained from the senescence regulatory networks developed in (Borrill et al. 2019a). Arabidopsis orthologs were extracted from Ensembl Plants (Vilella et al. 2009), and gene function was obtained from TAIR (Berardini et al. 2015). The senescence expression in Arabidopsis was determined using the Klepikova Expression Atlas (Klepikova et al. 2016), as implemented in TAIR; genes were considered highly expressed during senescence when at least 75% of the maximum TPM was reached in the senescing leaf.

Gene ID (v1.1)	Fold change in senescence	Upregulated 2x in senescence	Timing of upregulation	Arabidopsis Orthologue ID	Arabidopsis Orthologue Name	Arabidopsis Function	Arabidopsis Senescence Expression?
TraesCS2B02G150600	1.6	NO	UP 13 DAA	AT3G59140	ABCC10	Transporter; member of MRP subfamily	YES
TraesCS2D02G507800	5.1	YES	UP 17 DAA	NA	NA	NA	NA
TraesCS4A02G336400	13.8	YES	UP 13 DAA	AT4G14070, AT3G23790, AT4G14070, AT3G23790, AT4G14070, AT3G23790, AT4G14070, AT3G23790	AAE15, AAE16, AAE15, AAE16, AAE15, AAE16, AAE15, AAE16	Plastidic acyl activating enzyme involved in the elongation of exogenous medium-chain fatty acids to 16- and 18-carbon fatty acids.	YES
TraesCS5A02G230300	0	NO	NA	NA	NA	NA	NA
TraesCS5D02G237600	25.1	YES	UP 17 DAA	NA	NA	NA	NA
TraesCS5D02G533700	19.0	YES	UP 13 DAA	AT4G14070, AT3G23790, AT4G14070, AT3G23790, AT4G14070, AT3G23790, AT4G14070, AT3G23790	AAE15, AAE16, AAE15, AAE16, AAE15, AAE16, AAE15, AAE16	Plastidic acyl activating enzyme involved in the elongation of exogenous medium-chain fatty acids to 16- and 18-carbon fatty acids.	YES
TraesCS6B02G188300	4.6	YES	UP 13 DAA	AT1G23240, AT1G23250, AT1G70680, AT1G70670	PXG5, PXG4	Caleosin-related family protein, role in oxidation-reduction process	NO
TraesCSU02G100900	29.3	YES	NA	NA	NA	NA	NA

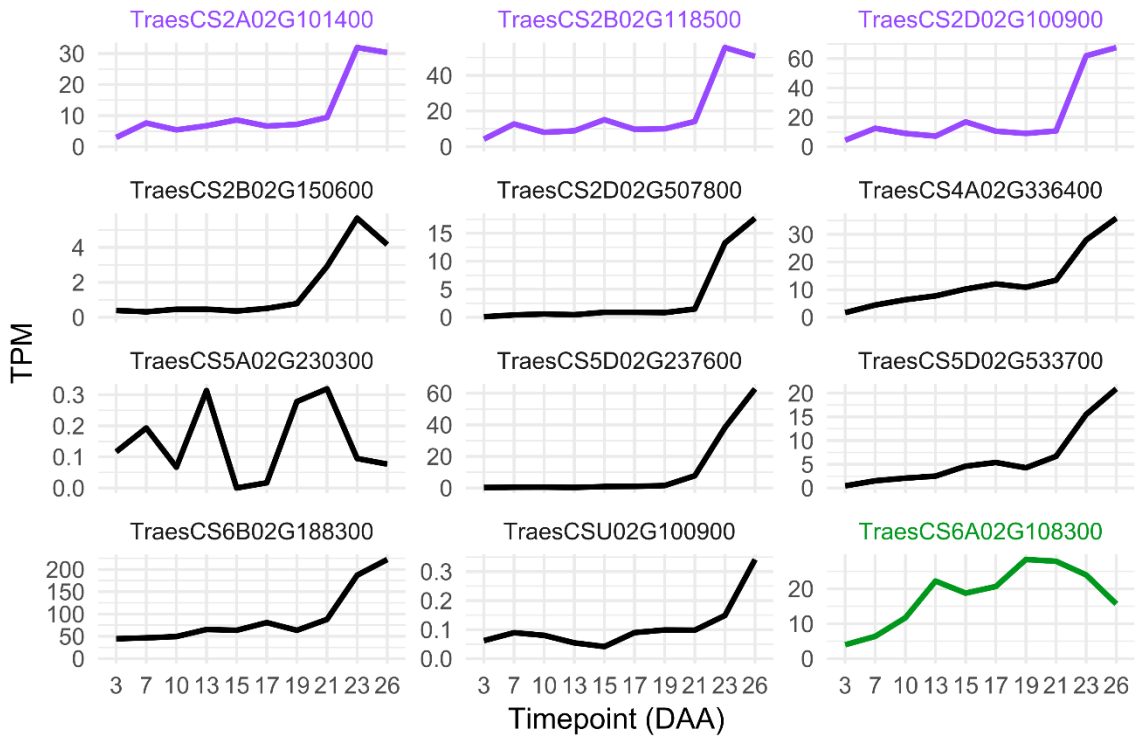

Figure 5-30: Expression levels of the shared targets of the NAC3 triad and *NAM-A1* most closely resemble the NAC3 triad. Expression data was taken from the flag leaf senescence timecourse in (Borrill et al. 2019a). The homoeologs from the NAC3 triad are shown in purple, and *NAM-A1* is shown in green. Expression levels are presented in transcripts per million (TPM). Note that the y-axes are different for each gene.

5.4 Discussion

5.4.1 *NAM-B1 may form heterodimers with other NAC transcription factors to regulate developmental senescence*

5.4.1.1 The possible role of heterodimers in transcriptional regulation

We presented preliminary data in Chapter 4 which demonstrated that *NAM-B1* is unable to induce premature senescence when ectopically over-expressed in wheat. This suggested that there may be other components required for *NAM-B1* and its homoeologs to initiate monocarpic senescence. Based on a yeast two-hybrid library screen, we identified a set of candidate genes that may interact with NAM-B1, including other NAC transcription factors (Table 5-8). Various studies have demonstrated that NAC transcription factors are able to form heterodimers with other NACs, as detailed in the introduction. For example, the NAC transcription factors VNI2 and VND7 were shown to interact in a Y2H screen (Yamaguchi et al. 2010), while the rice NAC OsNAC5 was found to interact with three other NAC transcription factors (Jeong et al. 2009). Based on this previous knowledge, we chose to prioritise the NAC transcription factors identified from the library screen for further validation and functional characterisation.

The biological importance of these heterodimers, however, is not always known. In the case of VNI2 and VND7, VNI2 was found to repress genes that were regulated by VND7. This suggests

that VNI2 and VND7 may interact in an antagonistic manner, whereby VNI2 binds VND7 and prevents it from activating its target genes, thereby controlling xylem differentiation (Yamaguchi et al. 2010). Heterodimerisation may also be a mechanism to increase the range of possible DNA binding sites for a given NAC transcription factor. Olsen et al. previously demonstrated that ANAC019 and ANAC092 could not only form heterodimers *in vitro*, but that they could also bind DNA as a heterodimer (Olsen et al. 2005a). Heterodimeric NAC transcription factors have since been identified in many different species (Hegedus et al. 2003, Olsen et al. 2005a, Mathew et al. 2016). The observation that NAC transcription factors can bind palindromic sequences, though this is not always the case (Lindemose et al. 2014), not only supports the fact that NACs bind DNA as dimers, as detailed in (Olsen et al. 2005a), but also leads to the hypothesis that non-palindromic binding sites may be indicative of heterodimer binding. Studies in bHLH transcription factors have found that heterodimerisation can alter binding specificities, suggesting this as a mechanism for expanding the range of binding sites accessible to transcription factors (Grove et al. 2009, De Masi et al. 2011, Lindemose et al. 2014). Future work on the binding site specificities of NAC transcription factors, particularly comparing homo- and heterodimer preferences, would begin to shed light on whether a similar system exists for NAC transcription factors.

5.4.1.2 Are the NAC3-1 proteins likely to bind the NAM-1 proteins in planta?

The NAC transcription factors we have identified, in particular the NAC3-1 triad, are strong candidate interactors of NAM-B1. However, while we were able to show that they could interact in artificial conditions, we were unable to prove that these proteins interact with NAM-B1 during senescence in wheat itself. To do so would require us to develop transgenic lines containing transgenes of both NAM-B1 and the given candidate under their native promoters, each appropriately tagged for a co-immunoprecipitation experiment. Alternatively, antibodies could be raised against the native proteins in wheat. However, it is both difficult and expensive to raise antibodies against native grass proteins as the majority of animals used to raise antibodies are natural grass eaters and thus already contain antibodies specific for grass proteins (Uauy, personal communication). Within the timeline of this thesis, it was not possible to develop transgenic lines for the most promising candidate genes. Nevertheless, as transformation efficiencies are steadily increasing while the cost of transformation is decreasing, this is an avenue of investigation that should be revisited in future work.

As it was impractical to verify these interactions in wheat itself, we could instead turn to expression data as an initial indication of whether these protein-protein interactions were likely to take place. In particular, we were interested in whether these proteins were likely to be in physical proximity in the flag leaf during senescence. While gene expression is an imperfect measure of protein abundance (Liu et al. 2016c), it can be used as a first approximation for protein levels. Using this measure, we find that the NAC3-1 triad and *NAM-A1* have distinct gene expression patterns during the process of flag leaf senescence (Figure 5-30; Figure 5-31). While *NAM-A1* begins to be

upregulated shortly after anthesis, with an early peak at 13 DAA, the NAC3-1 triad only peaks in expression 10 days later at 23 DAA. By this stage, the expression levels of *NAM-A1* are already decreasing from their second peak at 19 DAA.

Given the discrepancy in expression patterns between the NAC3-1 triad and *NAM-A1*, are their proteins likely to physically interact in the flag leaf during senescence? It may be the case that the proteins form heterodimers throughout the progression of senescence, as the NAC3-1 triad is also expressed during mid-senescence (~ 15 DAA; Figure 5-31A), though at lower levels than in the later stages of senescence. Alternatively, given that we have shown NAM-A1 and NAM-B1 can physically interact, the NAM proteins may preferentially form homodimers or homoeodimers – dimers between homoeologous proteins – earlier in senescence, when their protein abundance is theoretically higher than that of the NAC3-1 triad. By 23 DAA, the abundance of the NAM proteins may be dropping, in line with their expression levels, while the abundance of the NAC3-1 proteins may be increasing. As the concentration of the NAC3-1 proteins increases relative to that of the NAM proteins, heterodimers may be more likely to form. Importantly, we have not demonstrated here that the NAC3-1 proteins are able to form dimers amongst themselves, either as homo- or homoeodimers. Thus, while we can hypothesise that the NAC3-1 proteins may homo- or homoeodimerise following the significant increase in expression at 23 DAA (Figure 5-31A), further work will be needed to test this possibility.

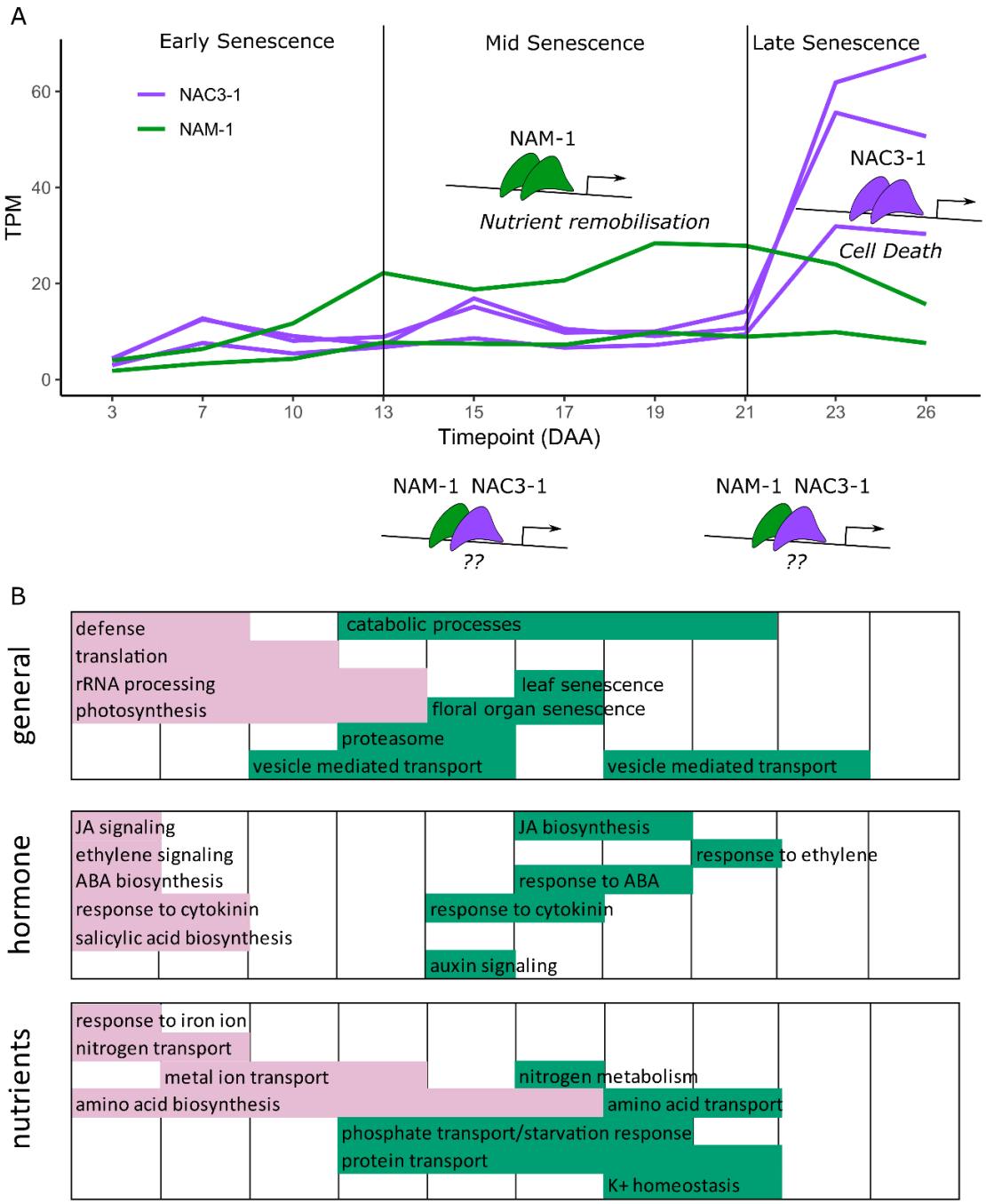

Figure 5-31: NAM-1 and NAC3-1 regulate different stages of senescence. (A) Expression profiles of the NAM-1 (green; A and D homoeologs) and NAC3-1 (purple; A, B, and D) homoeologs demonstrate that, while the NAM-1 genes are upregulated across the senescence process, the NAC-3 genes are primarily upregulated in the later stages of senescence. Based on gene expression profiles, NAM-1 and NAC-3 may form heterodimers when both genes are upregulated to similar levels, such as at 15 DAA or 21-23 DAA. In contrast, the transcription factors may be more likely to form homo- or homoeodimers when they are more highly expressed, such as from 23 DAA for NAC3-1 or between 13 and 21 DAA for NAM-1. (B) Genes associated with specific biological processes are up- (green) and down- (pink) regulated during senescence. Filled rectangles refer to GO terms associated with groups of genes which begin to be upregulated or downregulated at that timepoint. The upregulation of transport-related GO terms during the middle stage of senescence corresponds to the ongoing process of remobilisation, and the corresponding upregulation of NAM-1. This figure has been adapted from Figure 1 in Borrill et al. (2019a), under license ID 4671870949174.

Another consideration is that the RNA-Seq data as captured does not adequately define the spatial patterns of gene expression. Previous work in the lab, using *in situ* hybridisation, has showed that the NAM genes are expressed particularly around leaf vascular bundles, particularly in the mestome sheath cells (Borrill 2014). A similar pattern is observed in the peduncle, with high levels of NAM expression observed in the sclerenchyma and outer vascular bundles. In both cases, expression patterns varied according to the position of the flag leaf or peduncle which was tested. These cell-type-specific differences are not visible within the higher-level RNA-Seq data, which instead clusters these distinct cell types into single tissues such as "flag leaf" and "peduncle" (Ramírez-González et al. 2018). It is possible, therefore, that while the high-level tissue expression levels of the NAC3-1 triad and the NAM-1 genes do not completely correlate, their expression patterns on a cellular level may be more similar. Further work investigating the cell type-specificity of the NAC3-1 triad would shed further light on this hypothesis.

It's also possible that the NAM proteins form heterodimers with different proteins throughout the progression of senescence, perhaps only interacting with NAC3-1 in the later stages of senescence. For instance, the homoeologs TraesCS5A02G049100 and TraesCS5B02G054200 and the unrelated NAC TraesCS3B02G371200 peak earlier in senescence (at 7 and 15 DAA) as well as at the later 23 DAA timepoint. We have shown that they can physically interact with NAM-B1; could they form heterodimers with the NAM homoeologs in wheat as well? We were not able to test the effect of TILLING mutants in TraesCS3B02G371200, while our experiments with TraesCS5A02G049100 and TraesCS5B02G054200 suggested that their function is not necessary for the normal onset and progression of senescence. However, these results do not preclude a role for the genes in regulating senescence, but merely show that, for the conditions tested, any effect caused by mutating these genes does not lead to a visible senescence phenotype.

One clear caveat in the use of a Y2H library screen to identify candidates is that the potential pool of interacting partners is influenced entirely by the choice of cDNA library used. As illustrated in this chapter, the expression levels of genes can vary dramatically across development and tissues. The cDNA library used for the Y2H library screen in this case was taken from spike, peduncle and flag leaf tissues harvested at 8, 14, and 20 days post-anthesis. These timepoints were chosen based on the known expression profile of *NAM-A1*, which peaks at approximately 13 and 21 DAA (Figure 5-30, Figure 5-31). While these three timepoints cover a large portion of senescence they necessarily exclude any genes which are expressed during senescence but not during one of the three selected timepoints. Another caveat to using Y2H screens to identify interacting partners of NAC transcription factors is that the C-terminal domain of NACs is often auto-activating (Jeong et al. 2009). In such cases, only the DNA-binding domain can be used to validate interactions. As a result, any interacting partners that require the C-terminal domain to interact will be discarded during the validation process. Intrinsically disordered regions such the NAC C-terminal domain have been proposed to be essential in the ability of hub proteins in network to bind to many

different targets, suggesting that the study of these regions should not be overlooked (Dunker et al. 2005). These caveats should be kept in mind when designing strategies to identify protein interacting partners of NAC transcription factors, and indeed other activating transcription factors.

5.4.2 *The NAC3-1 triad positively regulates senescence in wheat*

5.4.2.1 The action of the NAC3-1 triad is conserved between monocots and dicots

Of the NAC transcription factors which were found to interact with NAM-B1, the NAC-3 triad is the most promising candidate senescence regulator. These NAC transcription factors appear to function as part of a highly conserved programmed cell death (PCD) pathway. During the course of this thesis their Arabidopsis orthologs, *ANAC046* and *ANAC087*, were both found to regulate PCD in the root cap columella by inducing the expression of PCD-promoting genes (Huysmans et al. 2018). *ANAC046* also acts as a positive regulator of leaf senescence by activating genes involved in chlorophyll breakdown (Oda-Yamamizo et al. 2016). This mirrors the likely action of the NAC3-1 triad which induces senescence by the induction of cell death and catabolic genes (see Table 5-19, Table 5-20, Table 5-21, and Table 8-2).

5.4.2.2 The NAC3-1 genes may act to regulate programmed cell death in non-senescence contexts

The role of the Arabidopsis orthologs of the NAC3-1 triad, *ANAC046* and *ANAC087*, also highlights the fact that the NAC3-1 triad may not only have roles in the regulation of senescence. *ANAC046* was first shown to have a role in regulating leaf senescence (Oda-Yamamizo et al. 2016), and later was found to control programmed cell death in the columella and lateral root cap, along with *ANAC087* (Huysmans et al. 2018). As the ability to induce ectopic cell death when overexpressed is conserved between the NAC3-1 genes and *ANAC046*, it may be that other functions in Arabidopsis are also conserved in wheat. For instance, the NAC3-1 genes in wheat are also expressed in the wheat root system, particularly early in development from the seedling to third leaf stage (Ramírez-González et al. 2018). It is possible that the wheat orthologs share a similar role in regulating programmed cell death during early root development. Indeed, the NAC3-1 triad may more broadly regulate programmed cell death in a cell-specific manner, which is not necessarily identified through tissue-scale sampling for RNA-Seq analysis.

5.4.2.3 The NAM and NAC3-1 genes may distinguish between the remobilisation and cell death phases of senescence

Intriguingly, the cell death caused by ectopic expression of the NAC3-1 triad in *N. benthamiana* suggests that not only is this PCD pathway conserved between monocots and dicots, but also that the wheat transcription factors are themselves able to recognise the promoters of PCD-related genes in a different species. This observation is not trivial, as it has been shown that the orthologs of senescence-associated genes in wheat are often not senescence-related themselves or show very different expression profiles across senescence (Borrill et al. 2019a). In addition, the key positive

senescence regulator in wheat, *NAM-B1* and its homoeologs, may be a unique development in the *Pooideae*. The closest orthologs of the NAM genes in both rice (Os07g37920) and Arabidopsis (*AtNAM, AtNAC2,* and *NAC025*) have no role in monocarpic senescence. Instead, they influence other developmental processes such as anther dehiscence (Os07g37920) (Distelfeld et al. 2012), shoot apical meristem development (*AtNAM*) (Duval et al. 2002), salt stress and root development (*AtNAC2*) (He et al. 2005), and seed development (*NAC025*) (Kunieda et al. 2008).

The NAM-driven senescence pathway in wheat, therefore, appears to be an evolutionarily-novel approach to inducing monocarpic senescence. This neo-functionalisation may stem from the duplication of the NAM triad from the original position on chromosome 2 (now the NAM-2 triad) to create the paralogous copies on chromosome 6 (the NAM-1 triad). One hypothesis stemming from this is that the NAM genes induce the onset of monocarpic senescence, the tightly regulated process of nutrient remobilisation from the flag leaf and other photosynthetic tissue into the developing grain. Following the onset of senescence and remobilisation of nutrients, other transcription factors such as the NAC3-1 triad are activated and begin to induce programmed cell death. In such a way the separation between the remobilisation and PCD stages of senescence would be maintained.

This hypothesis is partly supported by the upregulation of transport-related processes in the middle and later stages of senescence, as detailed in (Borrill et al. 2019a) and depicted in Figure 5-31B. The timing of these processes corresponds with the upregulation of the NAM-1 genes, known to be involved in nutrient remobilisation. However, the upregulation of the NAC3-1 triad, from 21 DAA, is not associated with a clear shift in GO-term enrichment in flag leaf tissue (Figure 5-31A and B). While the downstream targets of the NAC3-1 triad in the GENIE3 network are highly associated with senescence-related GO terms, this association is not recapitulated in the senescence timecourse (Borrill et al. 2019a). The 21 DAA timepoint does, however, coincide with upregulation of genes associated with catabolic processes, yet this process has been consistently associated with upregulated genes from 13 DAA onwards. Further work looking at variation in gene expression across later stages of senescence would be valuable in untangling this relationship, in particular investigating whether genes involved in programmed cell death beyond the NAC3-1 triad are preferentially upregulated at this stage.

5.4.3 *The NAC3 clade has expanded within the Triticeae*

The clade of closely related NAC transcription factors we identified, named NAC3, has gone through multiple rounds of tandem gene duplication within the Triticeae. We demonstrated in Figure 5-22 that each member of the outgroup, consisting of *Brachypodium distachyon*, *Leersia perrieri*, and the sequenced species within the genus *Oryza*, contains only one ortholog of the NAC3 clade. However, all sequenced members of the Triticeae with sufficiently assembled genomes to allow synteny analysis contain multiple copies of the NAC3 genes in close physical

proximity to each other (Figure 5-23). We identified one complete triad in *T. aestivum*, NAC3-1, with corresponding orthologs in *H. vulgare*, *T. urartu*, *T. dicoccoides*, and *A. tauschii*. As this group, coloured purple in the synteny and phylogeny figures, is present in all members of the Triticeae and has the highest levels of expression across development (Figure 5-20) we could hypothesize that these genes represent the original copy of the *NAC3* gene. The outgroup gene, from *B. distachyon*, also clusters more closely with the NAC3-1 group consistent with this hypothesis.

The remaining two groups of genes are likely derived from tandem duplications of the original *NAC3* gene. One group, NAC3-2, appears to be an incomplete triad, containing only A and D homoeologs in *T. aestivum* and being completely lost in *T. dicoccoides* and *A. tauschii*. As we were able to identify a pseudogene on the B genome of *T. aestivum* that falls in this group and that an ortholog is present in *H. vulgare*, this suggests that this group is derived from a tandem duplication that occurred after the Triticeae tribe split from the rest of the *Poaceae* family, but before barley split from the *Triticinae* subtribe (wheat and immediate progenitors). The NAC3-2 group appears to be undergoing a process of pseudogenisation, corroborated by the nearly non-existent levels of gene expression for the A and D wheat homoeologs (Figure 5-20) across over 200 RNA-Seq samples.

The final group of NAC3 genes, NAC3-3, is the largest of the three groups and appears to be the product of multiple rounds of tandem duplication. An independent triplication occurred in barley (Figure 5-22), while multiple rounds of independent duplications and triplications have occurred on each of the three wheat genomes and their progenitors (Figure 5-23). Notably, the wheat genes all show varying levels of expression, from nearly non-existent to approaching the levels of the NAC3-1 triad (Figure 5-21). It is possible that some of the duplicated genes are undergoing neofunctionalization; variable patterns of expression can be seen in the developmental expression timecourse (Figure 5-20), while we have yet to explore their expression and possible function under stress conditions. Further work in this area would help increase our understanding of the selective pressures, if any, acting on these duplicated genes. Given the known function of the NAC3-1 triad and the Arabidopsis orthologs in regulating programmed cell death (Oda-Yamamizo et al. 2016, Huysmans et al. 2018), it may be that some of the duplicated NAC3 genes regulate PCD under specific developmental and/or stress conditions.

The expansion of the NAC3 clade through what appears to be tandem duplication also raises questions as to the genomic context that promoted this event multiple times. The recurrence of these duplications suggests that the region surrounding the NAC3 genes may contain highly repetitive sequences which facilitate unequal crossing over, a common cause of tandem duplications (Reams and Roth 2015). Characterisation of the transposable element content of the intergenic regions between the NAC3 members may shed further light on this hypothesis.

5.4.4 *The GENIE3 network predicts that the NAC3-1 triad functions in macromolecule breakdown and programmed cell death.*

We previously showed that a GENIE3 network derived from 850 wheat RNA-Seq samples could accurately predict genes whose expression is affected by the function of *NAM-A1* (Figure 5-24) (Harrington et al. 2019a). We were then able to use the network to predict a set of transcription factors that were likely to act in similar pathways to *NAM-A1*, identifying one of the NAC3-1 homoeologs in the process (Figure 5-27). Having shown the utility of using this network to study senescence, we could then apply it to studying the pathways regulated by uncharacterised transcription factors.

Here we focused on the NAC-3 triad and recovered target genes that were highly enriched for GO terms related to senescence. Not only did this result match that expected based on the senescence phenotype observed in the TILLING mutants (e.g. in Figure 5-16), but we also identified downstream genes that were orthologous to targets of *ANAC046* (Table 5-22) (Oda-Yamamizo et al. 2016). We can now use these sets of candidate targets to inform the selection of candidate genes for further study.

It is important to note, as mentioned briefly in the results section, that the GENIE3 network will include predicted transcription factor-target relationships that do not represent a biologically valid connection. This can in part be accounted for by the weighting metric which is assigned to all connections. However, as the GENIE3 network is developed solely from expression datasets, high weights may only indicate genes with very similar expression patterns, rather than genes which are regulated by the transcription factor in question. For example, the three homoeologs in the NAC3 triad are all connected to each other with high weights in the GENIE3 network (Figure 5-25) as they share very similar expression profiles.

As a result, the GENIE3 network will be primarily useful for candidate gene prediction when used in concert with other independent datasets. For example, we integrated the GENIE3 network with two independent sets of RNA-Seq data, the *nam-a1* mutants (Pearce et al. 2014) and the flag leaf senescence timecourse (Borrill et al. 2019a), to produce a list of candidate genes which can be pursued for further investigation (Figure 5-27) (Harrington et al. 2019a). In the case of the NAC3-1 triad, we use the overlap of targets between the homoeologs to initially subset the list of candidates. This can then be supplemented by information from independent datasets, such as the flag leaf senescence timecourse (Borrill et al. 2019a), and information derived from orthology to aid in selecting candidates for phenotypic characterisation.

5.5 Summary

In this chapter we have shown that the NAC transcription factor NAM-B1, a known positive regulator of senescence, can physically interact with a set of NAC transcription factors *in vivo* and *in planta*, using yeast two-hybrid and co-immunoprecipitation techniques. We then characterised

the function of these NAC transcription factors, using transient over-expression systems in *Nicotiana benthamiana* and the chemically-mutagenised cv. Kronos TILLING lines, and found that the NAC3-1 triad is a positive regulator of senescence. This triad is part of a highly expanded syntenic group of transcription factors unique to the Triticeae and which show variable levels of expression across development. Finally, we leveraged the new genomic resources available in wheat to explore the function of the candidate genes, initially, and the NAC3-1 triad in more detail. We found that the NAC3-1 triad likely regulates genes involved in senescence-related processes, and that its predicted target genes overlap with those of *NAM-A1*. Finally, we suggest that the NAC3-1 triad positively regulates senescence in tandem with the NAM genes, possibly through physical protein-protein interactions.

6 General Discussion

6.1 Summary of the thesis

In this thesis, we have approached the complex question of how senescence is regulated in wheat through a variety of methods. We first sought to identify novel senescence-associated genes using a forward screen of the Kronos TILLING population (Chapter 2). In this screen we identified a novel chlorosis mutant and used sequencing technologies to map the causal mutation to a 4.3 Mb region, identifying a putative Mg^{2+} transporter as a strong candidate gene. However, we failed to identify any novel senescence mutants, instead only finding two lines with mutations, one missense and one a splice accepter variant, in the NAC transcription factor *NAM-A1* which is a known positive-regulator of senescence.

We then used these identified lines as a springboard for studying the effect of missense mutations on the function of the NAC domain (Chapter 3). Our work demonstrated that the wheat TILLING population, in particular the tetraploid Kronos lines, provides a wealth of mutations that can be used to characterise the roles of specific residues in protein function. Alongside this work, we also highlighted the previously underappreciated role of *NAM-A1* in regulating peduncle senescence in particular.

In doing so, this emphasized the known role of the *NAM* genes in regulating monocarpic senescence in wheat. However, the impact of ectopic expression any of the *NAM* genes had not yet been studied. We therefore developed two methods for premature expression of *NAM-B1*, using both a constitutive expression system and a heat-shock inducible system (Chapter 4). Despite the role of *NAM-B1* and its homoeologs in positively regulating senescence, premature expression of *NAM-B1* was not sufficient to induce premature senescence. This suggested that there may be one or more partners that *NAM-B1* must interact with in order to induce senescence.

To investigate this hypothesis, we carried out a yeast two-hybrid library screen and identified a set of NAC transcription factors that could bind to the NAC domain of *NAM-B1* (Chapter 5). Using the Kronos TILLING population, we demonstrated that one set of homoeologs also act as positive regulators of senescence. Using RNA-Seq datasets and a GENIE3 network, we investigated the regulatory neighbourhood of these transcription factors, suggesting that they act in the latter stages of senescence to induce cell and tissue death, following remobilisation of nutrients earlier in senescence.

The work presented in this thesis also highlights the difficulties inherent in studying a complex, multi-genic trait such as senescence, which will be discussed in detail in the remainder of this chapter.

6.2 Senescence is a complex trait that has the potential to inform crop breeding

6.2.1 Senescence is environmentally unstable

Perhaps the most notable example in this thesis of an environmentally-unstable trait is that of the *YES-1* locus (Chapter 2), which drives a chlorosis phenotype that is extremely clear in the field in the United Kingdom, but which is completely lost under both glasshouse conditions in the UK, and in the field in California. While the chlorosis phenotype is most likely not a result of premature senescence, this environmental variability does highlight the difficulties faced when working with senescence. While monocarpic senescence in annuals such as wheat is a critical, programmed part of plant development, the symptoms which are characteristic of senescence are substantially influenced by environmental conditions. Senescence is known to be accelerated under water-limited and heat-stressed conditions (Schippers et al. 2015), meaning that small differences in senescence timing that might be visible in the field in more temperate climates may be completely lost in hotter, drier climes.

This effect of environment on senescence profiles can also be seen in more controlled conditions, as was found in our characterisation of the NAC transcription factor candidates in Chapter 5. One set of homoeologs, TraesCS5A02G049100 and TraesCS5B02G054200**,** showed an inconsistent early senescence phenotype that only appeared in one out of the three experiments carried out. It may be the case that this phenotype was not truly linked to the mutations in these genes, and that it was caused by background mutations co-segregating with the mutant alleles. However, it is also possible that the mutations in these genes may truly lead to slightly premature senescence but that this earliness is so subtle as to be lost under certain environmental conditions.

We saw a similar effect with TraesCS2B02G118500 which, in most of the experiments, showed a clear delay in senescence in the double mutant lines. However, the severity of this delay varied substantially between experiments and conditions. We found that the delay was lost entirely under controlled environment conditions. At the same time, repeated experiments in the glasshouse at different times of the year, using seeds taken from identical parental lines, showed very different levels of delay for the different mutant lines (Figure 6-1). This environmental dependence was previously found for lines containing an introgressed copy of *NAM-B1* (Uauy et al. 2006a), where the early senescence phenotype was only observed under field conditions, and never in glasshouse conditions (Uauy, Personal Communication). These incidents highlight the importance, when working on senescence, of repeating the characterisation of candidate gene mutations multiple times. While no phenotype may be observed in the initial trial, this is not necessarily because the mutation has no effect on senescence but could instead be because the effect was not visible under the conditions tested.

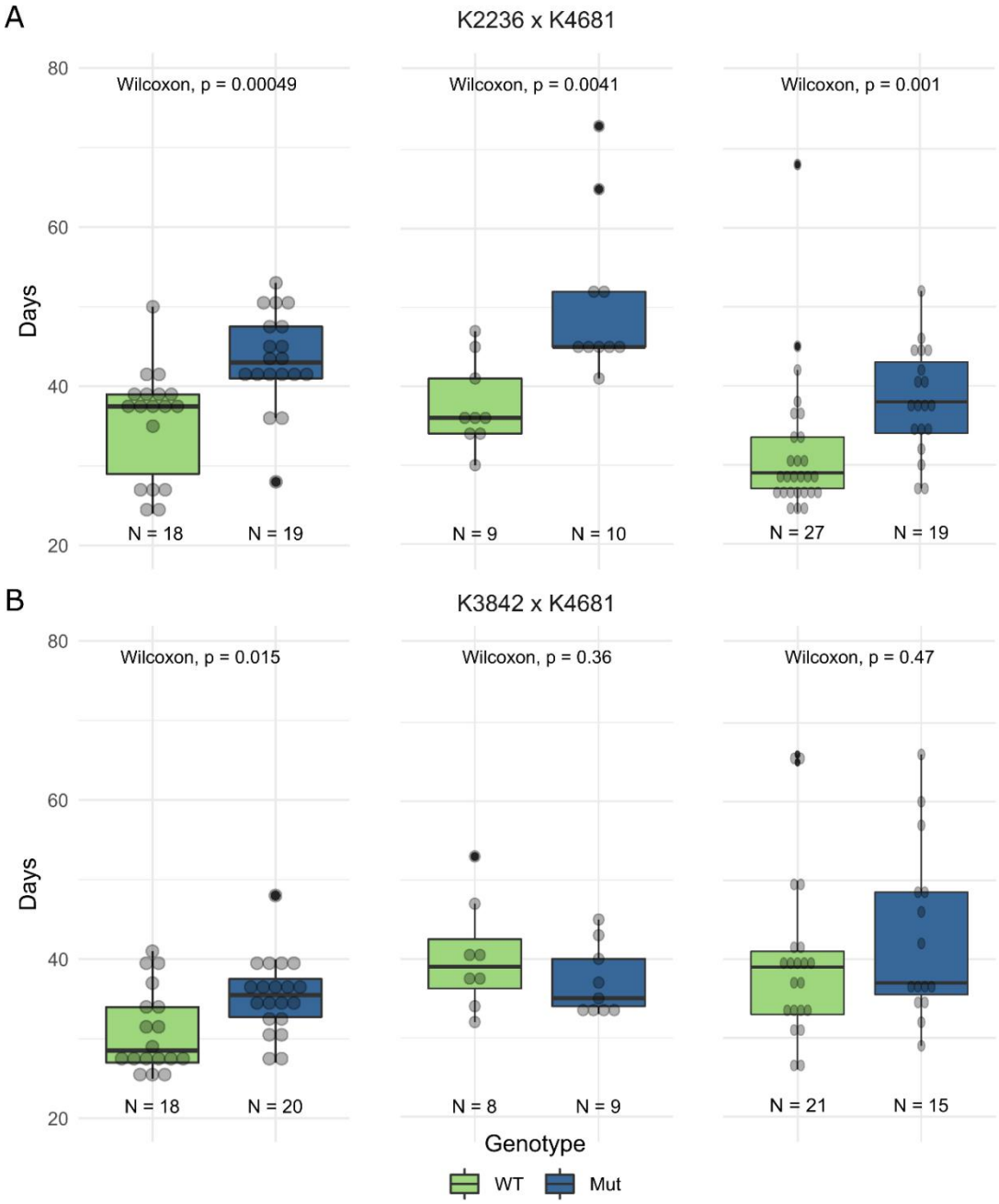

Figure 6-1: Variation in senescence timing observed for the NAC3-A TILLING mutants under glasshouse conditions. BC_1F_3 seedlings from the same BC_1F_2 parents were grown three times, sown in September 2018, December 2018, and February 2019 (left to right) and scored for flag leaf senescence (depicted as days from heading). The K2236 x K4681 (A) and K3842 x K4681 (B) crosses are depicted as they were grown in all three experiments. Statistical comparisons were carried out using the two-sample Wilcox-test. The panels are adapted from, left to right, Figure 5-16, Figure 8-1, and Figure 8-2.

6.2.2 Can environmental variability in senescence be leveraged for bespoke crop breeding?

The variability in senescence profiles as a result of environmental conditions also highlights a potential opportunity for future crop breeding, particularly when looking forward to a future that may involve precision gene editing, at least in some jurisdictions. As we demonstrated with our

allelic series in *NAM-A1*, precise single-base-pair mutations can lead to variable levels of the mutant phenotype. Some, such as the $K2711_{R110W}$ mutation, led to a severe delay of nearly 30 days for peduncle senescence. Others, such as the $K2060_{P154L}$ mutation, led to a more subtle delay of only three days. Such allelic variants could be introduced into elite varieties for environments in which they would provide the most benefit- perhaps the stronger, $K2711_{R110W}$, allele in highly arid and hot environments in which senescence progresses quickly, and thus would require a stronger delaying senescence allele to have a significant effect. In contrast, the weaker allele may be useful in more temperate locations, where a small decrease in *NAM-A1* activity is sufficient to induce the desired delay of only a few days. Base editing techniques have shown promise within the CRISPR-Cas9 system and hold out the potential for bespoke crop breeding in the future (Shimatani et al. 2017, Zong et al. 2017).

6.2.3 Improved transformation techniques are essential for a future of bespoke crop breeding

Because of our hypothesis that ectopic expression of *NAM-B1* would lead to premature senescence, and likely embryo-lethality, we needed to develop a system that would allow us to induce expression of *NAM-B1* only when desired. By using a heat-shock inducible promoter coupled with the Cre-*lox*P system, we could permanently over-express *NAM-B1* expression following heat-shock treatment. Such a system now has the potential to be used for characterisation of many other genes, particularly those which have temporally-sensitive phenotypic responses.

Yet our ability to create this system relied upon having access to the BRACT transformation system that has established a high level of expertise in Agrobacterium-mediated stable transformation of wheat. As it currently stands, stable transformation of wheat typically remains the purview of specialised labs. This is largely due to the technical challenges involved in successfully recovering transformed seedlings from tissue culture and in maintaining a supply of immature embryos for transformation (Lowe et al. 2016). This difficulty is compounded by the generally low efficiencies of transformation, often no more than 5% (Adamski et al. 2018). While attempts have been made to establish an equivalent system to the “floral dip” method in Arabidopsis, and some have claimed success (Agarwal et al. 2009), we have not been able to replicate this in the lab (data not shown). Compounding this difficulty is the fact that only a select few wheat cultivars have shown the propensity to be transformed. As a result, testing of transgenic constructs can only be carried out in varieties such as Fielder and Bobwhite, rather than in more immediately relevant elite cultivars (Adamski et al. 2018).

This inability to effectively and inexpensively generate transformed wheat lines is currently hindering the establishment of wheat as a model system for crops and polyploid species. While the genetic resources are available, characterisation of genes using transgenic approaches is hindered by the inability to easily complement mutant phenotypes, for example. Beyond the lab, new

advances in gene editing technologies have opened up the possibilities of transgene-free edited crops which will fall in line with genetic modification regulations in many parts of the world, though not currently within the European Union (Zhang et al. 2016). Yet until reliable, cost-effective methods of stable transformation in wheat are developed, these are likely to remain unfulfilled promises. Promising steps are being taken in this direction, such as the *Baby Boom*/*Wuschel2* system which has been established to great success in maize and other monocots (Lowe et al. 2016). Not only did this system dramatically increase the transformation efficiency, from only 0-2% to over 50% for some constructs, but it was also used to directly transform embryo sections from mature seeds and seedling leaf sections. The transformed cells were able to form transgenic calli and, later, mature transgenic plants with a regeneration frequency of 41% and 46% respectively. This system, therefore, not only has the possibility to substantially improve transformation efficiency, but also to reduce the need to generate immature wheat embryos for transformation via *Agrobacterium tumefaciens*. New transgenic methods such as this may also allow transformation of multiple different cultivars, rather than only those such as Fielder and Bobwhite that are possible currently. Direct transformation of elite cultivars from different regions would also facilitate the introduction of specific senescence profiles into germplasm adapted for different climactic regions, as discussed in the previous section.

6.2.4 Transient expression systems offer a flexible way to carry out preliminary characterisation of gene function.

While stable transgenic lines are still difficult and cost-prohibitive to obtain, in this thesis we were able to take advantage of the transient expression system in *Nicotiana benthamiana* to characterise the effect of specific NAC transcription factors on cell death induction. We were able to demonstrate that the NAC3 transcription factors could induce cell death long before the double mutants and back-crossed lines were produced containing TILLING mutations. In this case, the transient system provided us with a quick and consistent method of determining whether candidate NAC transcription factors had the capacity to induce cell death when expressed ectopically. The system was limited by the lack of information generated by a null result, that is when no cell death was observed. Such an outcome could be due to the gene having no role in senescence, acting as a negative regulator, or indeed acting as a positive regulator but without the ability to induce cell death in an ectopic system. As a result, the only informative outcome when testing an uncharacterised gene is the presence of cell death following infiltration.

However, given the appearance of cell death in response to expression of the wild-type allele, it is then possible to draw conclusions when cell death is not recapitulated in response to the infiltration of a mutant allele. This principle was used to demonstrate the effects of missense mutations in the NAC domain of *NAM-A1* (Chapter 3). Given the length of time required to generate the double mutants necessary to carry out phenotypic characterisation of candidate genes in wheat, transient expression systems such as this provide a useful, rapid assay for cell death activity. The assay can

provide preliminary results to help prioritise candidate genes for further investigation. Nevertheless, interpretation of any results should be carried out with great care to ensure that the presence or absence of a phenotype is not taken as diagnostic for gene function in wheat.

6.3 Untangling the wheat proteome and interactome will shed light on the regulatory systems governing senescence

6.3.1 Advances in antibody generation are required to study the roles of specific proteins in wheat

In this thesis we have begun to use wheat as a system to aid in the interrogation of protein functional domains, and their impacts on plant phenotype. As it stands, very little is known about the wheat proteome itself. Part of this is due to the difficulty in raising antibodies against grass-derived proteins; most animal systems used for raising antibodies feed on grasses and consequently produce antibodies to grass proteins themselves. Due to this, most of the studies that have looked at wheat proteins use peptide tags to allow their perception (e.g. Borrill (2014), Connorton et al. (2017), Harrington et al. (2019c)). While this is perfectly adequate in most cases when working with proteins *in vitro* or in ectopic systems, such as when expressed in *N. benthamiana*, the requirement for proteins to be tagged means that, in order to study a given wheat protein *in planta*, the protein must be tagged and then introduced through transgenic means (Borrill 2014).

This difficulty means that standard proteomics techniques, such as immunoprecipitation-mass spectrometry (IP-MS), are rarely, if ever, carried out in wheat. We therefore lack knowledge on the protein-binding landscape in wheat itself. While we can carry out *in vivo* studies of protein binding in yeast, as with the yeast two-hybrid library screen described in Chapter 5, these studies can only show the physical ability of the proteins to interact. It does not say anything about whether the two proteins are in the same temporal or physical location in the wheat plant, permitting them to actually interact *in planta*.

Recently, advances have been made in generating antibodies without the use of animal systems (Gray et al. 2016). Using systems such as naïve phage display libraries may facilitate the development of antibodies specific to native wheat proteins. Once this becomes more widely available, the protein landscape during complex pathways in wheat will be possible to study in depth. At the same time, as obtaining stable transformants in wheat becomes less expensive and more effective (Adamski et al. 2018), it will be increasingly feasible to develop transgenic lines expressing the gene of interest under the native promoter with a fused peptide tag to allow easy protein purification and isolation.

6.3.2 Large-scale proteomics can supplement expression databases

Alongside these techniques for specific protein isolation, large scale proteomics experiments in wheat will help us correlate the expression data obtained through www.wheat-expression.com with

protein levels, which often do not track exactly with gene expression. Techniques such as protein mass spectrometry can be used on complex peptide mixtures obtained from samples taken from different tissues or developmental stages (Domon and Aebersold 2006). Using the high-quality gene annotations associated with the gold-standard reference genomes, specific peptides can be identified from complex protein mixes derived from wheat tissues (Zhang et al. 2013). While analyses of these complex mixtures are unlikely to identify all or even most of the proteins present in the samples, they will begin to provide insight into protein abundances and post-translational modifications that are currently unknown in wheat.

One possibility to reduce the complexity of the sample would be to isolate specific organelles via cellular fractionation (Domon and Aebersold 2006). This could aid in studying the abundances of transcription factors, using purified nuclei as the proteome source as has previously been carried out in multiple plant species, including not only Arabidopsis (Goto et al. 2019) but also crop species such as rice (Agrawal and Rakwal 2011), maize (Casati 2012), and soybean (Yin and Komatsu 2016). In wheat itself, such an analysis has been carried out on nuclei extracted from grains during their development, identifying 114 distinct proteins across the six developmental stages considered (Bonnot et al. 2015). The majority of the proteins identified were consistent with that expected for nuclei, including histones, transcription factors, and translation factors, containing relatively few likely contaminants (6%). Mass spectrometry on whole tissues has also been applied to study variation in the proteome of developing spikes from multi-ovary wheat, identifying 90 differentially expressed proteins (Guo et al. 2019). These initial studies highlight the possibility for further interrogation of the wheat proteome across tissues and developmental stages.

6.3.3 NAC Transcription Factors are key players in the regulation of senescence

Throughout this thesis, we have highlighted NAC transcription factors that are involved in regulating senescence in wheat. As detailed at length in the introduction, NAC transcription factors have been shown to regulate many aspects of senescence in model species such as Arabidopsis. They are also the first two transcription factors of any kind found to regulate senescence in wheat—*NAM-B1* (Uauy et al. 2006b) and *NAC-S* (Zhao et al. 2015). In Chapter 5, we demonstrated that the *NAC3-1* transcription factors are positive regulators of senescence in wheat. These candidate senescence regulators were identified through a protein interaction screen against *NAM-B1*. While their role in regulating senescence was also supported through their expression profiles and orthology with Arabidopsis, they were prioritised as candidate genes due to their interaction with *NAM-B1*. Studying protein interactions in wheat, therefore, directly facilitated our identification of a novel wheat senescence regulator.

Beyond the small set of NAC transcription factors which have been explicitly characterised in wheat senescence, gene expression analysis carried out in (Borrill et al. 2019a) across a timecourse

of flag leaf senescence found that the NAC transcription factor family was, as a whole, significantly upregulated in the latter half of the timecourse. While other transcription factors, notably the CCAAT_HAP2 and RWP-RK families are significantly upregulated until 15 days after anthesis, the NACs are the only family upregulated beyond this stage. This is particularly intriguing as other transcription factor families that have roles in regulating senescence in Arabidopsis are either only slightly but non-significantly upregulated (i.e. Myb-related) or neither up or down regulated (i.e. WRKY). Interestingly, recent work in Arabidopsis identified a trio of NAC transcription factors which govern a regulatory pivot between negative and positive regulation of senescence by the senescence-associated NAC transcription factors (Kim et al. 2018a). This suggests that while the NAC family may, as a whole, promote the shift towards senescence after anthesis, specific NAC transcription factors may also have a role in preventing the premature initiation of senescence. These trends suggest that the progression of senescence is regulated in part by a network of NAC transcription factors, and that this may be characteristic of monocarpic senescence in multiple plant species.

6.3.4 Heterodimerisation of NAC transcription factors may govern changes in senescence regulation

Transcription factors can form regulatory networks through binding DNA and directly regulating the expression of other genes in the network, or alternatively through forming protein-protein interactions between network components. In our work we have shown that *NAM-B1* can, in principle, physically interact with other NAC transcription factors. While we have not verified this, it is most likely that these interactions take the form of heterodimers (Olsen et al. 2005a, Mathew et al. 2016). The question then arises—how does the formation of heterodimers influence the DNA binding properties of NAC transcription factors?

As detailed previously in the discussion in Chapter 5, it is possible that the formation of heterodimers with different NAC transcription factors can alter the target sequence bound by the NAC dimer (Grove et al. 2009, De Masi et al. 2011). This has been shown to be the case for bHLH transcription factors (De Masi et al. 2011), but it has not yet been studied for NAC transcription factors. If this were the case, the formation of heterodimers may act as a kind of regulatory control mechanism. If a heterodimer between two particular NAC transcription factors were required in order to bind and activate (or repress) the expression of a particular set of genes, then the relative expression patterns of these two NAC transcription factors would be essential in controlling the regulation of the downstream targets. In this respect, by having their regulation depend on the expression and translation of two distinct transcription factors, these targets could be considered to have a higher level of regulation ensuring their timely expression.

While directed studies such as yeast two-hybrid library screens can be used to search for possible heterodimerisation partners of specific NAC transcription factors, these techniques are

limited to the choice of bait protein. When a specific transcription factor of interest is being studied, as was the case in this thesis with *NAM-B1*, this manner of screen can provide a high-quality list of candidate interactors. However, as we have discussed in detail in Chapter 5, the fact that proteins can interact with each other does not necessarily mean that they do interact *in planta*. One method to overcome this limitation is to carry out immunoprecipitation of the protein of interest from wheat tissue, and then characterise the binding proteins using mass spectrometry (Free et al. 2009). However, this method is again limited to a single gene of interest. If we are interested in understanding the prevalence of heterodimerisation between NAC transcription factors, such techniques are too low-throughput to provide a global understanding. Alternative approaches may be able to take a snapshot of protein-protein interactions within the nucleus. In human nuclei, mass spectrometry has been coupled with crosslinking to provide a global picture of protein-protein interactions (Fasci et al. 2018). Their analysis focused primarily on interactions with histone proteins, but they did identify various transcription factors within their dataset, raising the possibility of mining such a dataset for interactions involving particularly transcription factor families.

6.4 Network analysis can shed light on gene function in wheat

6.4.1 Gene expression networks provide insight for hypothesis generation

For complex traits such as senescence, it's likely that its onset and progression is governed by a set of transcription factors that interact to induce and halt different aspects of the senescence pathway. We have already highlighted the possible role of the NAC transcription factor family in regulatory networks governing senescence. In order to study these interactions, we can turn to gene network analyses in which RNA-Seq datasets are used to predict correlations and interactions between different genes. The power of these networks lies in their ability to provide information that can help guide hypothesis generation regarding the function of uncharacterised genes. In this thesis, data obtained from the GENIE3 network (Ramírez-González et al. 2018) helped generate our hypothesis that *NAC3-1* transcription factors act in the latter stages of senescence, particularly through regulation of genes involved in various catabolic processes.

Advances in gene network analysis highlight new avenues that should be explored, particularly when working on processes such as senescence which take place over a long period of time. The dynamic GENIE3 algorithm (Dyn-GENIE3) can integrate time-series data into its prediction of transcription factor targets (Huynh-Thu and Geurts 2018). Not only would running this analysis likely provide more specific predictions of transcription factor targets, compared to predictions based solely on gene co-expression, but such an algorithm provides the opportunity to extract temporally specific transcription factor targets. That is, specific subsets of the time-series data could be analysed using the Dyn-GENIE3 algorithm. If transcription factor targets change over the time series, perhaps due to changed binding affinities due to heterodimerisation, the target predictions may also change in the different subsets of the time series.

However, the predictive power of these networks should not be overstated. While, as demonstrated in this thesis, such regulatory networks can provide useful information on putative gene function, predictions from gene networks should not be considered infallible. In all, these networks will include false positives, whether predicted transcription factor targets that are incorrect, as in the case of the GENIE3 network, or co-expressed genes that do not act in similar pathways. Networks must be complemented by biological validation of their predictions, such as through validation of individual candidate genes through low-throughput approaches such as EMSA DNA-binding assays. At a minimum, independent RNA-Seq datasets should be used to validate predictions where possible, as demonstrated in this thesis and in (Harrington et al. 2019a). In species with more mature gene network studies, gene expression data is regularly integrated with large scale biological datasets on known transcription factor binding sites and protein-protein interactions to provide a more complete set of predictions of gene function (Rackham et al. 2016).

6.4.2 New techniques will allow the characterisation of TF binding sites

Moving forwards, it will be essential to couple *in silico* predictions of transcription factor targets from networks such as the GENIE3 with large-scale, high-throughput validation. Current attempts to carry out these analyses are also hindered by the difficulties inherent in working with wheat proteins, namely the need to use tagged proteins for pull-downs (discussed above). However, new *in vitro* techniques provide promising avenues to begin testing transcription factor binding sites. The DAPseq technique, for example, incubates purified protein with extracted and purified genomic DNA libraries (Bartlett et al. 2017). Following immunoprecipitation, the fragments of DNA bound by the transcription factors can be sequenced and identified in a manner similar to that developed for the more common ChIP-Seq technique (Park 2009).

One of the benefits of using an *in vitro* technique such as DAPseq is that it allows high-throughput processing of multiple transcription factors and transcription factor combinations. Returning to our hypothesis that heterodimers of NAC transcription factors may have distinct binding preferences to homodimers of the constituent parts, DAPseq could be carried out using NAC transcription factors with different peptide tags, allowing sequential purification of only the heterodimeric form followed by identification of the corresponding binding sites. While DAPseq is carried out *in vitro* on purified genomic DNA, and thus lacks information on actual genome accessibility, it would provide essential information on the possible binding sites of transcription factors, aiding not only in predicting targets genes but also in characterising DNA binding motifs.

In tandem with DAPseq, other techniques could be applied on a genome-wide scale to characterise regions of open chromatin and transcription factor binding sites, complementing the *in vitro* approach of DAPseq with biologically-accurate information on genome accessibility. The ATAC-Seq method uses Tn5 transposase to cleave DNA in regions of open chromatin, creating a library of accessible DNA fragments (Buenrostro et al. 2015, Bajic et al. 2018). Such an

experiment could be carried out on tissue sampled from different organs, or across a timecourse of a specific developmental process, to build up a comprehensive picture of how the chromatin landscape changes across time and between organs. In the general introduction (Chapter 1), we highlighted the chromatin conformation changes that are known to happen on a large-scale during senescence in Arabidopsis, but we have no knowledge of either how chromatin changes in wheat during this process, nor how the specific patterns of accessible and inaccessible chromatin change during the progression of senescence. Correlating this kind of information with gene expression profiles and predicted transcription factor targets has the potential to provide a highly detailed understanding of how processes such as senescence are controlled *in vivo*.

6.5 Thoughts on the future of wheat and senescence research

Historically, wheat has been considered a difficult species to work with, hindered by its large, polyploid genome, long growing period, comparatively large physical size, and difficulty to transform. Yet, with the exception of size, these concerns are increasingly being overcome as the resources available to work with wheat continue to grow. There are now multiple high-quality genomes and gene annotations available not only for the model hexaploid wheat variety, cv. Chinese Spring, but also for other varieties, through the 10+ genomes project, and for many wild relatives and progenitor species (10+ Wheat Genomes Project 2016, Adamski et al. 2018). These physically-contiguous genomes will substantially accelerate the cloning of genes from QTL, which previously relied upon genetic mapping followed by tiling of BAC clones to identify the actual gene underpinning a trait. The high-quality gene annotations have also facilitated the integration of knowledge from other species, generating candidate genes based on orthology with species both closely related, such as rice, and more distantly related, such as Arabidopsis. Even the growing season of wheat is no longer insurmountable, as methods such as “speed breeding” have accelerated the growth of wheat plants under specific lighting and temperature conditions (Ghosh et al. 2018, Watson et al. 2018). This accelerated growth is particularly useful when developing crosses for validation in the field, allowing up to 6 generations to be carried out within a single calendar year when working with spring wheat.

All these resources have now positioned wheat as a viable model species, no longer reliant on knowledge generated in traditional model species that often fails to directly translate to wheat. Candidate genes identified through QTL mapping or from orthology with other species can be rapidly tested via mutants generated through the TILLING population (Uauy et al. 2009, Krasileva et al. 2017). As new transgenic technologies are introduced to wheat, such as the *Baby Boom*/*Wuschel2* system (Lowe et al. 2016), the broader wheat community will soon be able to access affordable and efficient transformation systems. Already, gene editing techniques such as CRISPR are being optimised for function in wheat and other monocots (Shan et al. 2013, Shan et al. 2014, Zong et al. 2017), laying the groundwork for their future use as fundamental tools in wheat research.

An obvious benefit to studying biological processes in wheat itself is the immediate application to breeding and food security. Yet working in wheat can also shed light on fundamental biological processes shared across the plant kingdom. The evolutionary history of wheat, derived from relatively recent hybridisation events in the Fertile Crescent (Dubcovsky and Dvorak 2007), has produced a polyploid species that has had comparatively little time to undergo the neo- and sub-functionalisation processes that are characteristic of more ancient whole genome duplications throughout the plant lineage. As such, wheat is a unique system in which to study the changes that hybridisation events and polyploidy can have on a range of parameters, including selection pressures and gene expression divergence. We explored this question in a recent paper, identifying patterns of gene expression dynamics between different sets of homoeologs (Ramírez-González et al. 2018). We hypothesized that these patterns may predict which homoeologs are in the process of neo- or sub-functionalisation.

As well as characteristics that are relatively unique to wheat, such as the recent polyploid genome, studying fundamental biological processes common to many different plant species in wheat will also greatly add to our understanding of the variation found in these processes across the Plant Kingdom. As we have shown in this thesis and elsewhere, senescence is a highly complex and tightly regulated process (Borrill et al. 2019a). Yet when even the definition of senescence itself is still up for debate, it's not surprising that the work carried out so far in different species has found that senescence is regulated in a multitude of different ways. While in some cases information from Arabidopsis and other model species may be translatable into wheat, such as seems to be the case with the *NAC3-1* genes which are orthologous to the positive senescence regulator *ANAC046* (Oda-Yamamizo et al. 2016), in other cases, such as the *NAM* genes, no corresponding ortholog either exists or plays a role in monocarpic senescence (Uauy et al. 2006b, Distelfeld et al. 2012).

We would contend that, in an era of increasingly sophisticated genetic and genomic studies, it is no longer sufficient to claim wheat is too complex and difficult to work in. The possible benefits arising from studying traits like senescence in a variety of different species are immeasurable. Not only does this work shed light on the unique regulatory pathways governing monocarpic senescence in different annual species, but we can also begin to ask questions about the different pressures influencing the evolution of senescence. How was senescence influenced during domestication, for example? The *Gpc-B1* allele, corresponding to *NAM-B1*, has been named as a domestication allele in some contexts (Dubcovsky and Dvorak 2007, Lundström et al. 2017), yet while it was initially cloned through screening for high protein content, such screens were hardly in use during early domestication. Was senescence timing selected for as part of the domestication process? The new resources, both genomic and germplasm, from wild grasses and the wheat progenitors will help us begin to explore these questions.

We can also consider senescence as occurring on a spectrum, in which annual, monocarpic plants form an extreme camp where reproduction is intrinsically coupled to plant death. Many monocarps

are annuals, but not all annuals are monocarps, instead reproducing multiple times often until the beginning of a season with adverse environmental conditions. Given the right conditions, some annual plants have been shown to continue growing beyond a single year (Dahlgren and Roach 2017), suggesting that in their case it is the environment, rather than an inherent biological pathway, that is inducing the onset of senescence in these plants. To what extent, then, is senescence regulated by age and by environment?

These questions become more complex when we move beyond annual plants, into the varied realm of perennials. Some perennials, like bamboo (Makita 1992), still act as monocarps, reproducing only once before death (though the extent to which bamboo is a true monocarp is under debate (Miyazaki et al. 2009)). Other perennials can reproduce more than once, but only two or three times before undergoing a decline in reproductive success culminating in death within only a few seasons. Yet others persist for decades, if not centuries, producing seeds across a vast range of temporal and environmental conditions. Do these plants even show senescence (Munné-Bosch 2008)? If they do, they may only show signs of senescence in ways that we cannot yet measure, not least because of our own short lifespans, and (perhaps more importantly) short funding cycles.

This may seem unrelated to our understanding of senescence in wheat, but the fundamental point remains which is that we cannot expect to comprehend even a small part of the regulation of senescence by focusing on only one or few plant species. Our aim should be to cultivate as wide a range of species to study and utilise them to the best of their possibilities. With wheat now supported by vast genomic and genetic resources and high-quality germplasm collections, it has now reached a point where fundamental studies into the genetic regulation of senescence can be carried out. In other plants, perhaps perennial herbs or long-lived trees, we should endeavour to study them in their own optimal way, whether through ecological, physiological, or molecular approaches.

Such an approach will take a concerted effort for coordination between researchers in different fields, but only by embracing the diversity within the Plant Kingdom can we hope to untangle the mechanisms and evolutionary history of senescence. In this thesis, we have demonstrated that research in wheat has the potential to substantially increase our understanding of how senescence is regulated in grasses. In the near future, this could be extended to include how senescence was (if at all) influenced during domestication and how senescence differs in annuals bred for agricultural demands compared to annuals grown under natural conditions.

6.6 Concluding Statement

"The boundaries which divide Life from Death are at best shadowy and vague.
Who shall say where the one ends, and where the other begins?"
Edgar Allan Poe, *The Premature Burial*

In this thesis we have attempted to increase our understanding of the regulation of senescence in wheat. We have identified a new candidate positive regulator of senescence, NAC3-1, which has raised a new set of questions regarding the co-regulation of senescence by multiple NAC transcription factors. Do NAC3-1 and NAM-1 interact? And if so, how does this interaction govern senescence? As we continue to study senescence in wheat, using more high-throughput techniques, our understanding of the biological relevance of these kinds of protein-protein interactions will increase. We stand at the cusp of a new era of wheat research, where we move beyond single-gene studies towards an understanding of the unified regulation of developmental processes by multiple genes, proteins, and indeed physical constraints. In the context of senescence, our work in wheat should be integrated with the knowledge of senescence regulation in other species to begin delineating an understanding of shared, universal properties of senescence and those which are unique to wheat alone.

7 References

10+ Wheat Genomes Project (2016). "The Wheat 'Pan Genome'." 2019, from http://www.10wheatgenomes.com/.

Acevedo-Garcia, J., et al. (2017). "*mlo*-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic TILLING approach." Plant Biotechnology Journal **15**(3): 367-378.

Adamski, N. M., et al. (2018). "A roadmap for gene functional characterisation in wheat." PeerJ Preprints **6**: e26877v26871.

Aerts, R. (1996). "Nutrient Resorption from Senescing Leaves of Perennials: Are there General Patterns?" Journal of Ecology **84**(4): 597-608.

Agarwal, S., et al. (2009). Floral Transformation of Wheat. Transgenic Wheat, Barley and Oats: Production and Characterization Protocols. Jones, H. D. and Shewry, P. R. Totowa, NJ, Humana Press**:** 105-113.

Agrawal, G. K. and Rakwal, R. (2011). "Rice proteomics: A move toward expanded proteome coverage to comparative and functional proteomics uncovers the mysteries of rice and plant biology." Proteomics **11**(9): 1630-1649.

Aida, M., et al. (1997). "Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant." The Plant Cell **9**(6): 841.

Allen, A. M., et al. (2017). "Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (Triticum aestivum)." Plant Biotechnology Journal **15**(3): 390-401.

Altschul, S. F., et al. (1990). "Basic local alignment search tool." Journal of Molecular Biology **215**(3): 403-410.

Aoyama, S., et al. (2014). "Regulation of senescence under elevated atmospheric CO_2 via ubiquitin modification." Plant Signaling & Behavior **9**(5): e28839.

Aoyama, T. and Chua, N.-H. (1997). "A glucocorticoid-mediated transcriptional induction system in transgenic plants." The Plant Journal **11**(3): 605-612.

Ashburner, M., et al. (2000). "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium." Nature Genetics **25**(1): 25-29.

Avila-Ospina, L., et al. (2017). "Metabolite Profiling for Leaf Senescence in Barley Reveals Decreases in Amino Acids and Glycolysis Intermediates." Agronomy **7**(1): 15.

Avni, R., et al. (2017). "Wild emmer genome architecture and diversity elucidate wheat evolution and domestication." Science **357**(6346): 93-97.

Avni, R., et al. (2014). "Functional characterization of *GPC-1* genes in hexaploid wheat." Planta **239**(2): 313-324.

Ay, N., et al. (2009). "Epigenetic programming via histone methylation at *WRKY53* controls leaf senescence in *Arabidopsis thaliana*." The Plant Journal **58**(2): 333-346.

Bajic, M., et al. (2018). "Identification of Open Chromatin Regions in Plant Genomes Using ATAC-Seq." Methods in Molecular Biology **1675**: 183-201.

Balazadeh, S., et al. (2011). "ORS1, an H_2O_2-responsive NAC transcription factor, controls senescence in *Arabidopsis thaliana*." Molecular Plant **4**(2): 346-360.

Balazadeh, S., et al. (2008). "Transcription factors regulating leaf senescence in *Arabidopsis thaliana*." Plant Biol (Stuttg) **10 Suppl 1**: 63-75.

Balk, J., et al. (2019). "Improving wheat as a source of iron and zinc for global nutrition." Nutrition Bulletin **44**(1): 53-59.

Bang, W. Y., et al. (2008). "The C-terminal region (640–967) of *Arabidopsis* CPL1 interacts with the abiotic stress- and ABA-responsive transcription factors." Biochemical and Biophysical Research Communications **372**(4): 907-912.

Bartlett, A., et al. (2017). "Mapping genome-wide transcription-factor binding sites using DAP-seq." Nature Protocols **12**: 1659.

Beales, J., et al. (2007). "A pseudo-response regulator is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* L.)." Theoretical and Applied Genetics **115**(5): 721-733.

Becraft, P. W. and Yi, G. (2010). "Regulation of aleurone development in cereal grains." Journal of Experimental Botany **62**(5): 1669-1675.

Belete, F., et al. (2018). "Effect of nitrogen fertilizer rates on grain yield and nitrogen uptake and use efficiency of bread wheat (*Triticum aestivum* L.) varieties on the Vertisols of central highlands of Ethiopia." Agriculture & Food Security **7**(1): 78.
Berardini, T. Z., et al. (2015). "The arabidopsis information resource: Making and mining the "gold standard" annotated reference plant genome." genesis **53**(8): 474-485.
Biesiekierski, J. R. (2017). "What is gluten?" Journal of Gastroenterology and Hepatology **32**(S1): 78-81.
Biffen, R. H. (1905). "Mendel's Laws of Inheritance and Wheat Breeding." The Journal of Agricultural Science **1**(1): 4-48.
Black, R. (2003). "Micronutrient deficiency--an underlying cause of morbidity and mortality." Bulletin of the World Health Organization **81**(2): 79-79.
Boden, S. A., et al. (2015). "*Ppd-1* is a key regulator of inflorescence architecture and paired spikelet development in wheat." Nature Plants **1**(2): 14016.
Bonnot, T., et al. (2015). "Changes in the nuclear proteome of developing wheat (*Triticum aestivum* L.) grain." Frontiers in Plant Science **6**: 905-905.
Borrell, A. K., et al. (2000). "Does Maintaining Green Leaf Area in Sorghum Improve Yield under Drought? II. Dry Matter Production and Yield." Crop Science **40**(4): 1037-1048.
Borrill, P., et al. (2015a). "Genomics as the key to unlocking the polyploid potential of wheat." New Phytologist **208**(4): 1008-1022.
Borrill, P., et al. (2015b). "Wheat Grain Filling Is Limited by Grain Filling Capacity rather than the Duration of Flag Leaf Photosynthesis: A Case Study Using *NAM* RNAi Plants." PLOS ONE **10**(8): e0134947.
Borrill, P., et al. (2019a). "Identification of transcription factors regulating senescence in wheat through gene regulatory network modelling." Plant Physiology: pp.00380.02019.
Borrill, P., et al. (2017). "Genome-Wide Sequence and Expression Analysis of the NAC Transcription Factor Family in Polyploid Wheat." G3: Genes|Genomes|Genetics **7**(9): 3019-3029.
Borrill, P., et al. (2019b). "Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat." The Plant Journal **97**(1): 56-72.
Borrill, P., et al. (2016). "expVIP: a Customizable RNA-seq Data Analysis and Visualization Platform." Plant Physiology **170**(4): 2172.
Borrill, P. G. (2014). The *NAM-B1* transcription factor and the control of grain composition in wheat, University of East Anglia. **Doctoral Thesis**.
Brassley, P. (2000). "Output and Technical Change in Twentieth-Century British Agriculture." The Agricultural History Review **48**(1): 60-84.
Bray, N. L., et al. (2016). "Near-optimal probabilistic RNA-seq quantification." Nature Biotechnology **34**: 525.
Breeze, E., et al. (2011). "High-Resolution Temporal Profiling of Transcripts during *Arabidopsis* Leaf Senescence Reveals a Distinct Chronology of Processes and Regulation." The Plant Cell **23**(3): 873.
Broadberry, S., et al. (2015). British Economic Growth, 1270–1870. Cambridge, Cambridge University Press.
Brown, T. A., et al. (2009). "The complex origins of domesticated crops in the Fertile Crescent." Trends in Ecology & Evolution **24**(2): 103-109.
Bruce, W., et al. (2000). "Expression Profiling of the Maize Flavonoid Pathway Genes Controlled by Estradiol-Inducible Transcription Factors CRC and P." The Plant Cell **12**(1): 65.
Brückner, A., et al. (2009). "Yeast two-hybrid, a powerful tool for systems biology." International journal of molecular sciences **10**(6): 2763-2788.
Brusslan, J. A., et al. (2012). "Genome-Wide Evaluation of Histone Methylation Changes Associated with Leaf Senescence in Arabidopsis." PLOS ONE **7**(3): e33151.
Buchanan-Wollaston, V., et al. (2005). "Comparative transcriptome analysis reveals significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in Arabidopsis." The Plant Journal **42**(4): 567-585.
Buenrostro, J. D., et al. (2015). "ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide." Current Protocols in Molecular Biology **109**(1): 21.29.21-21.29.29.
Caddick, M. X., et al. (1998). "An ethanol inducible gene switch for plants used to manipulate carbon metabolism." Nature Biotechnology **16**(2): 177-180.

Capra, J. A. and Singh, M. (2007). "Predicting functionally important residues from sequence conservation." Bioinformatics **23**(15): 1875-1882.
Casati, P. (2012). "Recent advances in maize nuclear proteomic studies reveal histone modifications." Frontiers in Plant Science **3**: 278.
Charnley, B. (2009). Rowland Biffen, Little Joss and Yeoman: looking back at two successes behind the birth of NIAB in 1919. Landmark Bulletin, NIABTAG**:** 3-4.
Chen, D., et al. (2018). "Overexpression of a predominantly root-expressed NAC transcription factor in wheat roots enhances root length, biomass and drought tolerance." Plant Cell Reports **37**(2): 225-237.
Chen, X., et al. (2016). "POWERDRESS interacts with HISTONE DEACETYLASE 9 to promote aging in Arabidopsis." eLife **5**: e17214.
Cho, E. J., et al. (2016). "A Mutation in Plant-Specific SWI2/SNF2-Like Chromatin-Remodeling Proteins, DRD1 and DDM1, Delays Leaf Senescence in *Arabidopsis thaliana*." PLOS ONE **11**(1): e0146826.
Clark, D. G., et al. (2004). "Drought-induced Leaf Senescence and Horticultural Performance of Transgenic PSAG12-*IPT* Petunias." Journal of the American Society for Horticultural Science **129**(1): 93-99.
Clark, J. W. and Donoghue, P. C. J. (2018). "Whole-Genome Duplication and Plant Macroevolution." Trends in Plant Science **23**(10): 933-945.
Clavijo, B. J., et al. (2017a). "W2RAP: a pipeline for high quality, robust assemblies of large complex genomes from short read data." bioRxiv: 110999.
Clavijo, B. J., et al. (2017b). "An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations." Genome Res **27**(5): 885-896.
Cockram, J. and Mackay, I. (2018). Genetic Mapping Populations for Conducting High-Resolution Trait Mapping in Plants. Plant Genetics and Molecular Biology. Varshney, R., Pandey, M. and Chitikineni, A., Springer, Cham. **164:** 109-138.
Collinge, M. and Boller, T. (2001). "Differential induction of two potato genes, *Stprx2* and *StNAC*, in response to infection by *Phytophthora infestans* and to wounding." Plant Mol Biol **46**(5): 521-529.
Connorton, J. M., et al. (2017). "Wheat Vacuolar Iron Transporter TaVIT2 Transports Fe and Mn and Is Effective for Biofortification." Plant Physiology **174**(4): 2434-2444.
Consortium, T. U. (2018). "UniProt: a worldwide hub of protein knowledge." Nucleic Acids Research **47**(D1): D506-D515.
Cormier, F., et al. (2015). "Detection of *NAM-A1* Natural Variants in Bread Wheat Reveals Differences in Haplotype Distribution between a Worldwide Core Collection and European Elite Germplasm." Agronomy **5**(2).
Crews, D. E. (2003). Human senescence: evolutionary and biocultural perspectives. Cambridge, Cambridge University Press.
Cunningham, B. C. and Wells, J. A. (1989). "High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis." Science **244**(4908): 1081-1085.
Dahlgren, J. P. and Roach, D. A. (2017). Demographic Senescence in Herbaceous Plants. The Evolution of Senescence in the Tree of Life. Jones, O. R., Shefferson, R. P. and Salguero-Gómez, R. Cambridge, Cambridge University Press**:** 303-319.
Dai, C., et al. (2018). "Calmodulin 1 Regulates Senescence and ABA Response in *Arabidopsis*." Frontiers in Plant Science **9**(803).
Danilova, M. N., et al. (2017). "Opposite roles of the *Arabidopsis* cytokinin receptors AHK2 and AHK3 in the expression of plastid genes and genes for the plastid transcriptional machinery during senescence." Plant Mol Biol **93**(4-5): 533-546.
Darwin, C. (1868). The Variation of Animals and Plants under Domestication. London, John Murray, Albemarle Street.
Davies, W. P. (2003). "An Historical Perspective from the Green Revolution to the Gene Revolution." Nutrition Reviews **61**(s6): S124-S134.
De Masi, F., et al. (2011). "Using a structural and logics systems approach to infer bHLH–DNA binding specificity determinants." Nucleic Acids Research **39**(11): 4553-4563.
Deevey, E. S. (1947). "Life Tables for Natural Populations of Animals." The Quarterly Review of Biology **22**(4): 283-314.

DeRocher, A. E., et al. (1991). "Expression of a Conserved Family of Cytoplasmic Low Molecular Weight Heat Shock Proteins during Heat Stress and Recovery." Plant Physiology **96**(4): 1038.
Diaz, C., et al. (2005). "Characterization of Markers to Determine the Extent and Variability of Leaf Senescence in Arabidopsis. A Metabolic Profiling Approach." Plant Physiology **138**(2): 898.
Distelfeld, A., et al. (2014). "Senescence, nutrient remobilization, and yield in wheat and barley." Journal of Experimental Botany **65**(14): 3783-3798.
Distelfeld, A., et al. (2012). "Divergent functions of orthologous NAC transcription factors in wheat and rice." Plant Mol Biol **78**(4-5): 515-524.
Dixon, L. E., et al. (2018). "*TEOSINTE BRANCHED1* Regulates Inflorescence Architecture and Development in Bread Wheat (*Triticum aestivum*)." The Plant Cell **30**(3): 563.
Dodsworth, S., et al. (2016). "Is post-polyploidization diploidization the key to the evolutionary success of angiosperms?" Botanical Journal of the Linnean Society **180**(1): 1-5.
Doležel, J., et al. (2012). "Chromosomes in the flow to simplify genome analysis." Functional & Integrative Genomics **12**(3): 397-416.
Domon, B. and Aebersold, R. (2006). "Mass Spectrometry and Protein Analysis." Science **312**(5771): 212.
Dowrick, G. J. (1957). "The influence of temperature on meiosis." Heredity **11**(1): 37-49.
Draeger, T. and Moore, G. (2017). "Short periods of high temperature during meiosis prevent normal meiotic progression and reduce grain number in hexaploid wheat (*Triticum aestivum* L.)." Theoretical and Applied Genetics **130**(9): 1785-1800.
Dubcovsky, J. and Dvorak, J. (2007). "Genome Plasticity a Key Factor in the Success of Polyploid Wheat Under Domestication." Science **316**(5833): 1862.
Dunker, A. K., et al. (2005). "Flexible nets." Febs j **272**(20): 5129-5148.
Duval, M., et al. (2002). "Molecular characterization of *AtNAM*: a member of the *Arabidopsis* NAC domain superfamily." Plant Mol Biol **50**(2): 237-248.
Ehness, R. and Roitsch, T. (1997). "Co-ordinated induction of mRNAs for extracellular invertase and a glucose transporter in *Chenopodium rubrum* by cytokinins." The Plant Journal **11**(3): 539-548.
Engler, C., et al. (2008). "A One Pot, One Step, Precision Cloning Method with High Throughput Capability." PLOS ONE **3**(11): e3647.
Ernst, H. A., et al. (2004). "Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors." EMBO Rep **5**(3): 297-303.
Eshaghi, S., et al. (2006). "Crystal Structure of a Divalent Metal Ion Transporter CorA at 2.9 Angstrom Resolution." Science **313**(5785): 354.
FAO (2019a, 05/09/2019). "FAO Cereal Supply and Demand Brief." Retrieved 20/09/2019, 2019, from http://www.fao.org/worldfoodsituation/csdb/en/.
FAO (2019b). "FAOSTAT Data." Retrieved 20/09/2019, 2019, from http://www.fao.org/faostat/en/#data.
Fasci, D., et al. (2018). "Histone Interaction Landscapes Visualized by Crosslinking Mass Spectrometry in Intact Cell Nuclei." Molecular & Cellular Proteomics **17**(10): 2018.
Feldman, M., et al. (1995). Wheats. Evolution of Crop Plants. Smartt, J. and Simmonds, N. W. London, Longman Scientific**:** 184-192.
Feng, H., et al. (2014). "The target gene of tae-miR164, a novel NAC transcription factor from the NAM subfamily, negatively regulates resistance of wheat to stripe rust." Mol Plant Pathol **15**(3): 284-296.
Fields, S. and Sternglanz, R. (1994). "The two-hybrid system: an assay for protein-protein interactions." Trends in Genetics **10**(8): 286-292.
Free, R. B., et al. (2009). "Identifying novel protein-protein interactions using co-immunoprecipitation and mass spectroscopy." Curr Protoc Neurosci **Chapter 5**: Unit 5.28.
Freeman, J., et al. (2011). "Temporal and spatial control of transgene expression using a heat-inducible promoter in transgenic wheat." Plant Biotechnology Journal **9**(7): 788-796.
Fu, D., et al. (2005). "Large deletions within the first intron in *VRN-1* are associated with spring growth habit in barley and wheat." Molecular Genetics and Genomics **273**(1): 54-65.
Gan, S. and Amasino, R. M. (1995). "Inhibition of Leaf Senescence by Autoregulated Production of Cytokinin." Science **270**(5244): 1986.

Gao, S., et al. (2016). "ABF2, ABF3, and ABF4 Promote ABA-Mediated Chlorophyll Degradation and Leaf Senescence by Transcriptional Activation of Chlorophyll Catabolic Genes and Senescence-Associated Genes in Arabidopsis." Molecular Plant **9**(9): 1272-1285.
Garapati, P., et al. (2015). "Transcription Factor ATAF1 in Arabidopsis Promotes Senescence by Direct Regulation of Key Chloroplast Maintenance and Senescence Transcriptional Cascades." Plant Physiology **168**(3): 1122.
Gardiner, L.-J., et al. (2019). "Integrating genomic resources to present full gene and putative promoter capture probe sets for bread wheat." GigaScience **8**(4).
Gardiner, L.-J., et al. (2018). "Hidden variation in polyploid wheat drives local adaptation." Genome Res **28**(9): 1319-1332.
Garrison, E. and Marth, G. (2012) Haplotype-based variant detection from short-read sequencing. arXiv e-prints
Gaujoux, R. and Seoighe, C. (2010). "A flexible R package for nonnegative matrix factorization." BMC Bioinformatics **11**(1): 367.
Gayon, J. and Zallen, D. T. (1998). "The Role of the Vilmorin Company in the Promotion and Diffusion of the Experimental Science of Heredity in France, 1840-1920." Journal of the History of Biology **31**(2): 241-262.
Gebert, M., et al. (2009). "A Root-Expressed Magnesium Transporter of the *MRS2/MGT* Gene Family in *Arabidopsis thaliana* Allows for Growth in Low-Mg^{2+} Environments." The Plant Cell **21**(12): 4018.
Ghanem, M. E., et al. (2008). "Hormonal changes during salinity-induced leaf senescence in tomato (*Solanum lycopersicum* L.)." Journal of Experimental Botany **59**(11): 3039-3050.
Ghosh, S., et al. (2018). "Speed breeding in growth chambers and glasshouses for crop breeding and model plant research." Nature Protocols **13**(12): 2944-2963.
Giorgi, D., et al. (2013). "FISHIS: Fluorescence In Situ Hybridization in Suspension and Chromosome Flow Sorting Made Easy." PLOS ONE **8**(2): e57994.
Goentoro, L., et al. (2009). "The incoherent feedforward loop can provide fold-change detection in gene regulation." Mol Cell **36**(5): 894-899.
Goto, C., et al. (2019). "Comprehensive nuclear proteome of *Arabidopsis* obtained by sequential extraction." Nucleus **10**(1): 81-92.
Gray, A. C., et al. (2016). "Animal-Friendly Affinity Reagents: Replacing the Needless in the Haystack." Trends in Biotechnology **34**(12): 960-969.
Greenwood, J. R., et al. (2017). "New alleles of the wheat domestication gene *Q* reveal multiple roles in growth and reproductive development." Development **144**(11): 1959.
Gregersen, P. L. (2011). Senescence and Nutrient Remobilization in Crop Plants. The Molecular and Physiological Basis of Nutrient Use Efficiency in Crops. Hawkesford, M. J. and Barraclough, P.**:** 83-102.
Greve, K., et al. (2003). "Interactions between plant RING-H2 and plant-specific NAC (NAM/ATAF1/2/CUC2) proteins: RING-H2 molecular specificity and cellular localization." Biochem J **371**(Pt 1): 97-108.
Grewal, S., et al. (2019). "Rapid identification of homozygosity and site of wild relative introgressions in wheat through chromosome-specific KASP genotyping assays." Plant Biotechnology Journal **0**(ja).
Grogan, S. M., et al. (2016). "Allelic Variation in Developmental Genes and Effects on Winter Wheat Heading Date in the U.S. Great Plains." PLOS ONE **11**(4): e0152852.
Grove, C. A., et al. (2009). "A Multiparameter Network Reveals Extensive Divergence between *C. elegans* bHLH Transcription Factors." Cell **138**(2): 314-327.
Guarda, G., et al. (2004). "Grain yield, nitrogen-use efficiency and baking quality of old and modern Italian bread-wheat cultivars grown at different nitrogen levels." European Journal of Agronomy **21**(2): 181-192.
Guo, J., et al. (2019). "Comparative proteomic analysis of multi-ovary wheat under heterogeneous cytoplasm suppression." BMC Plant Biology **19**(1): 175.
Guo, Y. and Gan, S. (2006). "*AtNAP*, a NAC family transcription factor, has an important role in leaf senescence." The Plant Journal **46**(4): 601-612.
Guttieri, M. J., et al. (2013). "Nutrient partitioning and grain yield of *TaNAM*-RNAi wheat under abiotic stress." Plant and Soil **371**(1): 573-591.

Hamilton, W. D. (1966). "The moulding of senescence by natural selection." Journal of Theoretical Biology **12**(1): 12-45.
Harari, Y. N. (2015). Sapiens : a brief history of humankind. New York, Harper.
Harlan, J. R., et al. (1973). "Comparative Evolution of Cereals." Evolution **27**(2): 311-325.
Harrington, S. A., et al. (2019a). "Validation and characterisation of a wheat GENIE3 network using an independent RNA-Seq dataset." bioRxiv: 684183.
Harrington, S. A., et al. (2019b). "Identification of a Dominant Chlorosis Phenotype Through a Forward Screen of the *Triticum turgidum* cv. Kronos TILLING Population." Frontiers in Plant Science **10**(963).
Harrington, S. A., et al. (2019c). "Conserved residues in the wheat (*Triticum aestivum*) NAM-A1 NAC domain are required for protein binding and when mutated lead to delayed peduncle and flag leaf senescence." BMC Plant Biology **19**(1): 407.
Harwood, W. A. (2011). "Advances and remaining challenges in the transformation of barley and wheat." Journal of Experimental Botany **63**(5): 1791-1798.
He, X. J., et al. (2005). "AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development." The Plant Journal **44**(6): 903-916.
Hedden, P. (2003). "The genes of the Green Revolution." Trends in Genetics **19**(1): 5-9.
Hegedus, D., et al. (2003). "Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress." Plant Mol Biol **53**(3): 383-397.
Hepworth, J. and Dean, C. (2015). "Flowering Locus C's Lessons: Conserved Chromatin Switches Underpinning Developmental Timing and Adaptation." Plant Physiology **168**(4): 1237.
Hickman, R., et al. (2013). "A local regulatory network around three NAC transcription factors in stress responses and senescence in Arabidopsis leaves." The Plant Journal **75**(1): 26-39.
Hirrlinger, J., et al. (2009). "Split-CreERT2: Temporal Control of DNA Recombination Mediated by Split-Cre Protein Fragment Complementation." PLOS ONE **4**(12): e8354.
Hoagland, D. R. and Arnon, D. I. (1950). "The water-culture method for growing plants without soil." Circular. California Agricultural Experiment Station **347**(2nd edit): 32 pp.
Horie, Y., et al. (2009). "Participation of chlorophyll *b* reductase in the initial step of the degradation of light-harvesting chlorophyll *a/b*-protein complexes in *Arabidopsis*." Journal of Biological Chemistry **284**(26): 17449-17456.
House, C. (2019). "resourcetrade.earth." Retrieved 20/09/2019, 2019, from https://resourcetrade.earth.
Hu, T., et al. (2003). "Agrobacterium-mediated large-scale transformation of wheat (*Triticum aestivum* L.) using glyphosate selection." Plant Cell Reports **21**(10): 1010-1019.
Huang, B. E., et al. (2015a). "MAGIC populations in crops: current status and future prospects." Theoretical and Applied Genetics **128**(6): 999-1017.
Huang, J., et al. (2018). "Distinct tissue-specific transcriptional regulation revealed by gene regulatory networks in maize." BMC Plant Biology **18**(1): 111.
Huang, Q., et al. (2015b). "*TaNAC29*, a NAC transcription factor from wheat, enhances salt and drought tolerance in transgenic Arabidopsis." BMC Plant Biology **15**: 268.
Huang, S., et al. (2002). "Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum*/*Aegilops* complex and the evolutionary history of polyploid wheat." PNAS **99**(12): 8133.
Huang, X., et al. (2009). "High-throughput genotyping by whole-genome resequencing." Genome Res **19**(6): 1068-1076.
Huynh-Thu, V. A. and Geurts, P. (2018). "dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data." Scientific Reports **8**(1): 3384.
Huynh-Thu, V. A., et al. (2010). "Inferring Regulatory Networks from Expression Data Using Tree-Based Methods." PLOS ONE **5**(9): e12776.
Huysmans, M., et al. (2018). "NAC Transcription Factors ANAC087 and ANAC046 Control Distinct Aspects of Programmed Cell Death in the Arabidopsis Columella and Lateral Root Cap." The Plant Cell **30**(9): 2197-2213.
Hwang, I., et al. (2012). "Cytokinin signaling networks." Annual Review of Plant Biology **63**: 353-380.

IWGSC (2014). "A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome." Science **345**(6194): 1251788.
IWGSC, et al. (2018). "Shifting the limits in wheat research and breeding using a fully annotated reference genome." Science **361**(6403): eaar7191.
Janack, B., et al. (2016). "Knockdown of WHIRLY1 Affects Drought Stress-Induced Leaf Senescence and Histone Modifications of the Senescence-Associated Gene *HvS40*." Plants (Basel) **5**(3).
Jasinski, S., et al. (2005). "KNOX Action in *Arabidopsis* Is Mediated by Coordinate Regulation of Cytokinin and Gibberellin Activities." Current Biology **15**(17): 1560-1565.
Jensen, M. K., et al. (2008). "Transcriptional regulation by an NAC (NAM-ATAF1,2-CUC2) transcription factor attenuates ABA signalling for efficient basal defence towards *Blumeria graminis* f. sp. *hordei* in Arabidopsis." The Plant Journal **56**(6): 867-880.
Jensen, M. K., et al. (2010). "The *Arabidopsis thaliana* NAC transcription factor family: structure-function relationships and determinants of ANAC019 stress signalling." Biochem J **426**(2): 183-196.
Jensen, M. K. and Skriver, K. (2014). "NAC transcription factor gene regulatory and protein-protein interaction networks in plant stress responses and senescence." IUBMB Life **66**(3): 156-166.
Jeong, J. S., et al. (2009). "Rice NAC proteins act as homodimers and heterodimers." Plant Biotechnology Reports **3**(2): 127-134.
Jiang, Y., et al. (2014). "*Arabidopsis* WRKY57 Functions as a Node of Convergence for Jasmonic Acid– and Auxin-Mediated Signaling in Jasmonic Acid–Induced Leaf Senescence." The Plant Cell **26**(1): 230.
Jimeno, D., et al. (2006). "Analysis of Kinesin-2 Function in Photoreceptor Cells Using Synchronous Cre-loxP Knockout of Kif3a with RHO-Cre." Investigative Ophthalmology & Visual Science **47**(11): 5039-5046.
Jing, H.-C., et al. (2005). "Ethylene-induced leaf senescence depends on age-related changes and *OLD* genes in *Arabidopsis*." Journal of Experimental Botany **56**(421): 2915-2923.
Jupe, F., et al. (2013). "Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations." The Plant Journal **76**(3): 530-544.
Kang, K., et al. (2009). "Senescence-Induced Serotonin Biosynthesis and Its Role in Delaying Senescence in Rice Leaves." Plant Physiology **150**(3): 1380.
Kang, M., et al. (2018). "The C-Domain of the NAC Transcription Factor ANAC019 Is Necessary for pH-Tuned DNA Binding through a Histidine Switch in the N-Domain." Cell Rep **22**(5): 1141-1150.
Kashino, Y., et al. (2002). "Proteomic Analysis of a Highly Active Photosystem II Preparation from the *Cyanobacterium Synechocystis* sp. PCC 6803 Reveals the Presence of Novel Polypeptides." Biochemistry **41**(25): 8004-8012.
Kassambara, A. (2019). "ggpubr: 'ggplot2' Based Publication Ready Plots.". from https://CRAN.R-project.org/package=ggpubr.
Kersey, P. J., et al. (2017). "Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species." Nucleic Acids Research **46**(D1): D802-D808.
Kim, D., et al. (2015). "HISAT: a fast spliced aligner with low memory requirements." Nature Methods **12**: 357.
Kim, H. J., et al. (2014). "Gene regulatory cascade of senescence-associated NAC transcription factors activated by ETHYLENE-INSENSITIVE2-mediated leaf senescence signalling in *Arabidopsis*." Journal of Experimental Botany **65**(14): 4023-4036.
Kim, H. J., et al. (2018a). "Time-evolving genetic networks reveal a NAC troika that negatively regulates leaf senescence in *Arabidopsis*." PNAS **115**(21): E4930.
Kim, J., et al. (2018b). "Comparative transcriptome analysis in Arabidopsis *ein2/ore3* and *ahk3/ore12* mutants during dark-induced leaf senescence." Journal of Experimental Botany **69**(12): 3023-3036.
Kim, J. H., et al. (2009a). "Trifurcate Feed-Forward Regulation of Age-Dependent Cell Death Involving *miR164* in *Arabidopsis*." Science **323**(5917): 1053.
Kim, S. H., et al. (2009b). "Aberrant mRNA transcripts and the nonsense-mediated decay proteins UPF2 and UPF3 are enriched in the Arabidopsis nucleolus." The Plant Cell **21**(7): 2045-2057.

Kim, W., et al. (2003). "Physicochemical Properties and End-use Quality of Wheat Starch as a Function of Waxy Protein Alleles." Journal of Cereal Science **37**(2): 195-204.
Kim, Y.-S., et al. (2013). "Mutation of the Arabidopsis NAC016 Transcription Factor Delays Leaf Senescence." Plant and Cell Physiology **54**(10): 1660-1672.
Kirch, T., et al. (2003). "The *DORNRÖSCHEN/ENHANCER OF SHOOT REGENERATION1* Gene of Arabidopsis Acts in the Control of Meristem Cell Fate and Lateral Organ Development." The Plant Cell **15**(3): 694.
Kirkwood, T. B. L. (1977). "Evolution of ageing." Nature **270**(5635): 301-304.
Kiszonas, A. M. and Morris, C. F. (2018). "Wheat breeding for quality: A historical review." Cereal Chemistry **95**(1): 17-34.
Kjaersgaard, T., et al. (2011). "Senescence-associated Barley NAC (NAM, ATAF1,2, CUC) Transcription Factor Interacts with Radical-induced Cell Death 1 through a Disordered Regulatory Domain." Journal of Biological Chemistry **286**(41): 35418-35429.
Klepikova, A. V., et al. (2016). "A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling." The Plant Journal **88**(6): 1058-1070.
Kovalchuk, N., et al. (2013). "Optimization of *TaDREB3* gene expression in transgenic barley using cold-inducible promoters." Plant Biotechnology Journal **11**(6): 659-670.
Kragelund, B. B., et al. (2012). "Order by disorder in plant signaling." Trends in Plant Science **17**(11): 625-632.
Krasileva, K. V., et al. (2017). "Uncovering hidden variation in polyploid wheat." PNAS **114**(6): E913.
Krzywinski, M., et al. (2009). "Circos: An information aesthetic for comparative genomics." Genome Res **19**(9): 1639-1645.
Kubaláková, M., et al. (2002). "Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.)." Theoretical and Applied Genetics **104**(8): 1362-1372.
Kucharewicz, W., et al. (2017). "Acceleration of leaf senescence is slowed down in transgenic barley plants deficient in the DNA/RNA-binding protein WHIRLY1." Journal of Experimental Botany **68**(5): 983-996.
Kumari, M., et al. (2007). Variation for Staygreen Trait and its Association with Canopy Temperature Depression and Yield Traits Under Terminal Heat Stress. Wheat Production in Stressed Environments. Developments in Plant Breeding. Buck, H. T., Nisi, J. E. and Salomón, N. Dordrecht, Springer. **12**.
Kunieda, T., et al. (2008). "NAC family proteins NARS1/NAC2 and NARS2/NAM in the outer integument regulate embryogenesis in *Arabidopsis*." The Plant Cell **20**(10): 2631-2642.
Lee, I. C., et al. (2011). "Age-Dependent Action of an ABA-Inducible Receptor Kinase, RPK1, as a Positive Regulator of Senescence in Arabidopsis Leaves." Plant and Cell Physiology **52**(4): 651-662.
Leopold, A. C. (1961). "Senescence in Plant Development." Science **134**(3492): 1727.
Leopold, A. C. (1980). Aging and senescence in plant developent. Boca Raton, CRC Press Inc.**:** 1-12.
Levy, A. A. and Feldman, M. (2002). "The Impact of Polyploidy on Grass Genome Evolution." Plant Physiology **130**(4): 1587.
Li, H. (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv e-prints
Li, S., et al. (2016). "The role of ANAC072 in the regulation of chlorophyll degradation during age- and dark-induced leaf senescence." Plant Cell Reports **35**(8): 1729-1741.
Li, W., et al. (2018). "NAC Family Transcription Factors in Tobacco and Their Potential Role in Regulating Leaf Senescence." Frontiers in Plant Science **9**: 1900.
Li, Z., et al. (2013). "ETHYLENE-INSENSITIVE3 is a senescence-associated gene that accelerates age-dependent leaf senescence by directly repressing *miR164* transcription in *Arabidopsis*." The Plant Cell **25**(9): 3311-3328.
Liang, C., et al. (2016). "GhABF2, a bZIP transcription factor, confers drought and salinity tolerance in cotton (*Gossypium hirsutum* L.)." Scientific Reports **6**: 35040.
Lim, J., et al. (2018). "Antagonistic Roles of PhyA and PhyB in Far-Red Light-Dependent Leaf Senescence in *Arabidopsis thaliana*." Plant and Cell Physiology **59**(9): 1753-1764.
Lim, P. O., et al. (2007a). "Leaf Senescence." Annual Review of Plant Biology **58**(1): 115-136.

Lim, P. O., et al. (2007b). "Overexpression of a chromatin architecture-controlling AT-hook protein extends leaf longevity and increases the post-harvest storage life of plants." The Plant Journal **52**(6): 1140-1153.
Lindemose, S., et al. (2014). "A DNA-binding-site landscape and regulatory network analysis for NAC transcription factors in Arabidopsis thaliana." Nucleic Acids Research **42**(12): 7681-7693.
Ling, H.-Q., et al. (2018). "Genome sequence of the progenitor of wheat A subgenome *Triticum urartu*." Nature **557**(7705): 424-428.
Liu, C. J., et al. (1992). "Nonhomoeologous translocations between group 4, 5 and 7 chromosomes within wheat and rye." Theoretical and Applied Genetics **83**(3): 305-312.
Liu, J., et al. (2016a). "*Ghd2*, a *CONSTANS*-like gene, confers drought sensitivity through regulation of senescence in rice." Journal of Experimental Botany **67**(19): 5785-5798.
Liu, T., et al. (2016b). "The *Arabidopsis* transcription factor ABIG1 relays ABA signaled growth inhibition and drought induced senescence." eLife **5**.
Liu, Y., et al. (2016c). "On the Dependency of Cellular Protein Levels on mRNA Abundance." Cell **165**(3): 535-550.
Lowe, K., et al. (2016). "Morphogenic Regulators *Baby boom* and *Wuschel* Improve Monocot Transformation." The Plant Cell **28**(9): 1998.
Lundström, M., et al. (2017). "Evolutionary history of the *NAM-B1* gene in wild and domesticated tetraploid wheat." BMC genetics **18**(1): 118-118.
Lunin, V. V., et al. (2006). "Crystal structure of the CorA Mg2+ transporter." Nature **440**(7085): 833-837.
Luo, M.-C., et al. (2017). "Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*." Nature **551**: 498.
Lupton, F. G. H. (1987). History of wheat breeding. Wheat Breeding: Its scientific basis. Lupton, F. G. H. Dordrecht, Springer Netherlands**:** 51-70.
Maccaferri, M., et al. (2019). "Durum wheat genome highlights past domestication signatures and future improvement targets." Nature Genetics.
Makita, A. (1992). "Survivorship of a monocarpic bamboo grass, *Sasa kurilensis*, during the early regeneration process after mass flowering." Ecological Research **7**(3): 245-254.
Malthus, T. R. (1798). An Essay on the Principle of Population. London, England, J. Johnson in St Paul's Church-yard.
Mamanova, L., et al. (2010). "Target-enrichment strategies for next-generation sequencing." Nature Methods **7**: 111.
Mao, C., et al. (2017). "A Rice NAC Transcription Factor Promotes Leaf Senescence via ABA Biosynthesis." Plant Physiology **174**(3): 1747-1763.
Mao, X., et al. (2014). "Novel NAC transcription factor *TaNAC67* confers enhanced multi-abiotic stress tolerances in Arabidopsis." PLOS ONE **9**(1): e84359.
Mao, X., et al. (2012). "*TaNAC2*, a NAC-type wheat transcription factor conferring enhanced multiple abiotic stress tolerances in *Arabidopsis*." Journal of Experimental Botany **63**(8): 2933-2946.
Maqbool, A., et al. (2015). "Structural basis of pathogen recognition by an integrated HMA domain in a plant NLR immune receptor." eLife **4**: e08709.
Marchand, G., et al. (2014). "Bridging physiological and evolutionary time-scales in a gene regulatory network." New Phytologist **203**(2): 685-696.
Martinez, A., et al. (1999). "Ecdysone agonist inducible transcription in transgenic tobacco plants." The Plant Journal **19**(1): 97-106.
Mascher, M., et al. (2017). "A chromosome conformation capture ordered sequence of the barley genome." Nature **544**(7651): 427-433.
Matallana-Ramirez, L. P., et al. (2013). "NAC Transcription Factor *ORE1* and Senescence-Induced *BIFUNCTIONAL NUCLEASE1* (*BFN1*) Constitute a Regulatory Cascade in *Arabidopsis*." Molecular Plant **6**(5): 1438-1452.
Mathew, I. E., et al. (2016). "Three Rice NAC Transcription Factors Heteromerize and Are Associated with Seed Size." Frontiers in Plant Science **7**: 1638-1638.
Matsuoka, D., et al. (2015). "An abscisic acid inducible Arabidopsis MAPKKK, MAPKKK18 regulates leaf senescence via its kinase activity." Plant Mol Biol **87**(6): 565-575.
Medawar, P. B. (1952). An Unsolved Problem of Biology. London, UK, H. K. Lewis.

Mendel, G. (1865). Versuche über Plflanzenhybriden. Verhandlungen des naturforschenden Vereines in Brünn, Bd. IV für das Jahr 1865.
Merewitz, E., et al. (2016). "Differentially Expressed Genes Associated with Improved Drought Tolerance in Creeping Bentgrass Overexpressing a Gene for Cytokinin Biosynthesis." PLOS ONE **11**(11): e0166676.
Merewitz, E. B., et al. (2011). "Protein accumulation in leaves and roots associated with improved drought tolerance in creeping bentgrass expressing an ipt gene for cytokinin synthesis." Journal of Experimental Botany **62**(15): 5311-5333.
Mett, V. L., et al. (1993). "Copper-controllable gene expression system for whole plants." PNAS **90**(10): 4567.
Miao, Y., et al. (2013). "The single-stranded DNA-binding protein WHIRLY1 represses *WRKY53* expression and delays leaf senescence in a developmental stage-dependent manner in Arabidopsis." Plant Physiology **163**(2): 746-756.
Miao, Y., et al. (2007). "Arabidopsis MEKK1 can take a short cut: it can directly interact with senescence-related WRKY53 transcription factor on the protein level and can bind to its promoter." Plant Mol Biol **65**(1-2): 63-76.
Miao, Y. and Zentgraf, U. (2007). "The antagonist function of Arabidopsis WRKY53 and ESR/ESP in leaf senescence is modulated by the jasmonic and salicylic acid equilibrium." The Plant Cell **19**(3): 819-830.
Miki, Y., et al. (2019). "Origin of wheat B-genome chromosomes inferred from RNA sequencing analysis of leaf transcripts from section Sitopsis species of *Aegilops*." DNA Research **26**(2): 171-182.
Miyazaki, Y., et al. (2009). "Genets of dwarf bamboo do not die after one flowering event: evidence from genetic structure and flowering pattern." J Plant Res **122**(5): 523-528.
Mo, Y., et al. (2018). "Mapping causal mutations by exome sequencing in a wheat TILLING population: a tall mutant case study." Molecular Genetics and Genomics **293**(2): 463-477.
Monfreda, C., et al. (2008). "Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000." Global Biogeochemical Cycles **22**(1).
Moore, I., et al. (1998). "A transcription activation system for regulated gene expression in transgenic plants." PNAS **95**(1): 376-381.
Morris, K., et al. (2000). "Salicylic acid has a role in regulating gene expression during leaf senescence." The Plant Journal **23**(5): 677-685.
Munné-Bosch, S. (2008). "Do perennials really senesce?" Trends in Plant Science **13**(5): 216-220.
Munné-Bosch, S. (2014). "Perennial roots to immortality." Plant Physiology **166**(2): 720-725.
Munné-Bosch, S. and Alegre, L. (2004). "Die and let live: leaf senescence contributes to plant survival under drought stress." Functional Plant Biology **31**(3): 203-216.
Murphy, K. M., et al. (2008). "Relationship between yield and mineral nutrient concentrations in historical and modern spring wheat cultivars." Euphytica **163**(3): 381-390.
Nakagawa, T., et al. (2007). "Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation." J Biosci Bioeng **104**(1): 34-41.
Ng, P. C. and Henikoff, S. (2003). "SIFT: Predicting amino acid changes that affect protein function." Nucleic Acids Research **31**(13): 3812-3814.
Niu, C. F., et al. (2012). "Wheat *WRKY* genes *TaWRKY2* and *TaWRKY19* regulate abiotic stress tolerance in transgenic *Arabidopsis* plants." Plant Cell and Environment **35**(6): 1156-1170.
NOAA National Centers for Environmental Information (2017). State of the Climate: National Climate Report for Annual 2017.
Oda-Yamamizo, C., et al. (2016). "The NAC transcription factor ANAC046 is a positive regulator of chlorophyll degradation and senescence in Arabidopsis leaves." Scientific Reports **6**(1): 23609.
Olmos, S., et al. (2003). "Precise mapping of a locus affecting grain protein content in durum wheat." Theoretical and Applied Genetics **107**(7): 1243-1251.
Olsen, A. N., et al. (2005a). "DNA-binding specificity and molecular functions of NAC transcription factors." Plant Science **169**(4): 785-797.
Olsen, A. N., et al. (2005b). "NAC transcription factors: structurally distinct, functionally diverse." Trends in Plant Science **10**(2): 79-87.

ONS (2014). Nutrient Intakes, Office of National Statistics.
Ooka, H., et al. (2003). "Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*." DNA Research **10**(6): 239-247.
Ortiz-Ospina, E., et al. (2014, October 2018). "Trade and Globalisation." Retrieved 20/09/2019, 2019, from https://ourworldindata.org/trade-and-globalization.
Padidam, M., et al. (2003). "Chemical-inducible, ecdysone receptor-based gene expression system for plants." Transgenic Res **12**(1): 101-109.
Pallotta, M., et al. (2003). Marker assisted wheat breeding in the southern region of Australia. Proceedings of the 10th International Wheat Genetics Symposium, Paestum, Italy, Istituto Sperimentale per la Cerealicultura Roma, Italy.
Park, P. J. (2009). "ChIP–seq: advantages and challenges of a maturing technology." Nature Reviews Genetics **10**(10): 669-680.
Park, S.-Y., et al. (2007). "The Senescence-Induced Staygreen Protein Regulates Chlorophyll Degradation." The Plant Cell **19**(5): 1649.
Paterson, A. H., et al. (2012). Ancient and Recent Polyploidy in Monocots. Polyploidy and Genome Evolution. P., S. and D., S. Berlin, Heidelberg, Springer.
Paux, E., et al. (2008). "A Physical Map of the 1-Gigabase Bread Wheat Chromosome 3B." Science **322**: 101-104.
Payandeh, J. and Pai, E. F. (2006). "A structural basis for Mg^{2+} homeostasis and the CorA translocation cycle." The EMBO Journal **25**(16): 3762-3773.
Pearce, S., et al. (2014). "Regulation of Zn and Fe transporters by the *GPC1* gene during early wheat monocarpic senescence." BMC Plant Biology **14**(1): 368.
Pearl, R. and Miner, J. R. (1935). "Experimental Studies on the Duration of Life. XIV. The Comparative Mortality of Certain Lower Organisms." The Quarterly Review of Biology **10**(1): 60-79.
Peng, J., et al. (1999). "'Green revolution' genes encode mutant gibberellin response modulators." Nature **400**(6741): 256-261.
Peng, J. H., et al. (2011). "Domestication evolution, genetics and genomics in wheat." Molecular Breeding **28**(3): 281.
Pepe, J. F. and Heiner, R. E. (1975). "Plant Height, Protein Percentage, and Yield Relationships in Spring Wheat1." Crop Science **15**(6): 793-797.
Pimentel, H., et al. (2017). "Differential analysis of RNA-seq incorporating quantification uncertainty." Nature Methods **14**: 687.
Pingali, P. L. (2012). "Green revolution: impacts, limits, and the path ahead." PNAS **109**(31): 12302-12308.
Pruzinska, A., et al. (2003). "Chlorophyll breakdown: pheophorbide a oxygenase is a Rieske-type iron-sulfur protein, encoded by the *accelerated cell death 1* gene." PNAS **100**(25): 15259-15264.
Qi, T., et al. (2015). "Regulation of Jasmonate-Induced Leaf Senescence by Antagonism between bHLH Subgroup IIIe and IIId Factors in Arabidopsis." The Plant Cell **27**(6): 1634-1649.
Qin, D., et al. (2008). "Heat stress-responsive transcriptome analysis in heat susceptible and tolerant wheat (*Triticum aestivum* L.) by using Wheat Genome Array." BMC Genomics **9**(1): 432.
Qiu, K., et al. (2015). "EIN3 and ORE1 Accelerate Degreening during Ethylene-Mediated Leaf Senescence by Directly Activating Chlorophyll Catabolic Genes in *Arabidopsis*." PLOS Genetics **11**(7): e1005399.
R Core Team (2018). R: A Language and Environment for Statistical Computing. Computing, R. F. f. S. Vienna, Austria.
Raab, S., et al. (2009). "Identification of a novel E3 ubiquitin ligase that is required for suppression of premature senescence in Arabidopsis." The Plant Journal **59**(1): 39-51.
Rackham, O. J., et al. (2016). "A predictive computational framework for direct reprogramming between human cell types." Nature Genetics **48**(3): 331-335.
Raines, T., et al. (2016). "The cytokinin response factors modulate root and shoot growth and promote leaf senescence in Arabidopsis." The Plant Journal **85**(1): 134-147.
Ramírez-González, R. H., et al. (2018). "The transcriptional landscape of polyploid wheat." Science **361**(6403): eaar6089.

Ramirez-Gonzalez, R. H., et al. (2015a). "RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat." Plant Biotechnology Journal **13**(5): 613-624.
Ramirez-Gonzalez, R. H., et al. (2015b). "PolyMarker: A fast polyploid primer design pipeline." Bioinformatics **31**(12): 2038-2039.
Rauf, M., et al. (2013). "ORE1 balances leaf senescence against maintenance by antagonizing G2-like-mediated transcription." EMBO Rep **14**(4): 382-388.
Reams, A. B. and Roth, J. R. (2015). "Mechanisms of gene duplication and amplification." Cold Spring Harb Perspect Biol **7**(2): a016592.
Rey, M.-D., et al. (2018). "Magnesium Increases Homoeologous Crossover Frequency During Meiosis in *ZIP4* (*Ph1* Gene) Mutant Wheat-Wild Relative Hybrids." Frontiers in Plant Science **9**: 509-509.
Rhee, S. Y. and Mutwil, M. (2014). "Towards revealing the functions of all genes in plants." Trends in Plant Science **19**(4): 212-221.
Richmond, A. E. and Lang, A. (1957). "Effect of Kinetin on Protein Content and Survival of Detached Xanthium Leaves." Science **125**(3249): 650.
Risacher, T., et al. (2009). Highly efficient *Agrobacterium*-mediated transformation of wheat via *in planta* inoculation. Transgenic Wheat, Barley and Oats: Production and Characterization Protocols. Jones, H. and Shewry, P., Humana Press. **478:** 115-124.
Ritchie, H. and Roser, M. (2017). "Micronutrient Deficiency." Retrieved 20/09/2019, 2019, from https://ourworldindata.org/micronutrient-deficiency.
Rivas-San Vicente, M. and Plasencia, J. (2011). "Salicylic acid beyond defence: its role in plant growth and development." Journal of Experimental Botany **62**(10): 3321-3338.
Rivero, R. M., et al. (2007). "Delayed leaf senescence induces extreme drought tolerance in a flowering plant." PNAS **104**(49): 19631-19636.
Rodríguez-Leal, D., et al. (2017). "Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing." Cell **171**(2): 470-480.e478.
Rogers, H. J. (2015). Senescence-Associated Programmed Cell Death. Plant Programmed Cell Death. McCabe, P., Springer, Cham.
Rolny, N., et al. (2011). "Is the electrolyte leakage assay an unequivocal test of membrane deterioration during leaf senescence?" Plant Physiology and Biochemistry **49**(10): 1220-1227.
Roser, M. and Ritchie, H. (2017, September 2019). "Crop Yields." Retrieved 20/09/2019, 2019, from https://ourworldindata.org/crop-yields.
Roslan, H. A., et al. (2001). "Characterization of the ethanol-inducible alc gene-expression system in *Arabidopsis thaliana*." The Plant Journal **28**(2): 225-235.
Rufo, R., et al. (2019). "From landraces to improved cultivars: Assessment of genetic diversity and population structure of Mediterranean wheat using SNP markers." PLOS ONE **14**(7): e0219867.
Sade, N., et al. (2017). "Stress-induced senescence and plant tolerance to abiotic stress." Journal of Experimental Botany **69**(4): 845-853.
Saitou, N. and Nei, M. (1987). "The neighbor-joining method: a new method for reconstructing phylogenetic trees." Molecular Biology and Evolution **4**(4): 406-425.
Sakuraba, Y., et al. (2016). "*Arabidopsis* NAC016 promotes chlorophyll breakdown by directly upregulating *STAYGREEN1* transcription." Plant Cell Reports **35**(1): 155-166.
Sakuraba, Y., et al. (2014a). "Phytochrome-interacting transcription factors PIF4 and PIF5 induce leaf senescence in *Arabidopsis*." Nature Communications **5**(1): 4636.
Sakuraba, Y., et al. (2014b). "*Arabidopsis* STAY-GREEN2 is a negative regulator of chlorophyll degradation during leaf senescence." Molecular Plant **7**(8): 1288-1302.
Salguero-Gómez, R., et al. (2013). "Plants do not count… or do they? New perspectives on the universality of senescence." Journal of Ecology **101**(3): 545-554.
Sannemann, W., et al. (2018). "Adaptive selection of founder segments and epistatic control of plant height in the MAGIC winter wheat population WM-800." BMC Genomics **19**(1): 559.
Santi, L., et al. (2008). "An efficient plant viral expression system generating orally immunogenic Norwalk virus-like particles." Vaccine **26**(15): 1846-1854.
Sauer, B. (1998). "Inducible Gene Targeting in Mice Using the Cre/*lox* System." Methods **14**(4): 381-392.

Scarpeci, T. E., et al. (2017). "Overexpression of *AtERF019* delays plant growth and senescence, and improves drought tolerance in Arabidopsis." Journal of Experimental Botany **68**(3): 673-685.
Schena, M., et al. (1991). Vectors for constitutive and inducible gene expression in yeast. Methods in Enzymology, Academic Press. **194:** 389-398.
Schippers, J. H. M., et al. (2007). Developmental and Hormonal Control of Leaf Senescence. Annual Plant Reviews Volume 26: Senescence Processes in Plants. Gan, S.**:** 145-170.
Schippers, J. H. M., et al. (2015). "Living to Die and Dying to Live: The Survival Strategy behind Leaf Senescence." Plant Physiology **169**(2): 914-930.
Schommer, C., et al. (2008). "Control of jasmonate biosynthesis and senescence by miR319 targets." PLoS Biol **6**(9): e230.
Schwartz, R. and Dayhoff, M. (1978). Matrices for Detecting Distant Relationships. Atlas of Protein Sequences. Dayhoff, M., National Biomedical Research Foundation**:** 353-358.
Selth, L. A., et al. (2005). "A NAC domain protein interacts with tomato leaf curl virus replication accessory protein and enhances viral replication." The Plant Cell **17**(1): 311-325.
Seok, H. Y., et al. (2017). "*Arabidopsis AtNAP* functions as a negative regulator via repression of *AREB1* in salt stress response." Planta **245**(2): 329-341.
Shan, Q., et al. (2014). "Genome editing in rice and wheat using the CRISPR/Cas system." Nature Protocols **9**: 2395.
Shan, Q., et al. (2013). "Targeted genome modification of crop plants using a CRISPR-Cas system." Nature Biotechnology **31**: 686.
Shannon, P., et al. (2003). "Cytoscape: a software environment for integrated models of biomolecular interaction networks." Genome Res **13**(11): 2498-2504.
Shearer, J. (2012). The United Kingdom Cereals Industry, AHDB.
Shen, H., et al. (2009). "A Bioinformatic Analysis of NAC Genes for Plant Cell Wall Development in Relation to Lignocellulosic Bioenergy Production." BioEnergy Research **2**(4): 217.
Shen, J., et al. (2017). "The NAC-type transcription factor *OsNAC2* regulates ABA-dependent genes and abiotic stress tolerance in rice." Scientific Reports **7**: 40641-40641.
Shewry, P. R. (2009). "Wheat." Journal of Experimental Botany **60**(6): 1537-1553.
Shimatani, Z., et al. (2017). "Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion." Nature Biotechnology **35**: 441.
Shinmyo, A., et al. (1998). "Metabolic engineering of cultured tobacco cells." Biotechnol Bioeng **58**(2-3): 329-332.
Shorinola, O., et al. (2019). "Genetic Screening for Mutants with Altered Seminal Root Numbers in Hexaploid Wheat Using a High-Throughput Root Phenotyping Platform." G3: Genes|Genomes|Genetics **9**(9): 2799.
Sievers, F., et al. (2011). "Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega." Molecular systems biology **7**: 539-539.
Šimková, H., et al. (2008). "Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley." BMC Genomics **9**(1): 294.
Simons, K. J., et al. (2006). "Molecular Characterization of the Major Wheat Domestication Gene *Q*." Genetics **172**(1): 547.
Singh, S., et al. (2013). "Down-regulation of *OsSAG12-1* results in enhanced senescence and pathogen-induced cell death in transgenic rice plants." Journal of Biosciences **38**(3): 583-592.
Slade, A. J., et al. (2012). "Development of high amylose wheat through TILLING." BMC Plant Biology **12**: 69.
Snowball, K. and Robson, A. D. (1991). Nutrient Deficiencies and Toxicities in Wheat: A Guide for Field Identification. CIMMYT. Mexico, D.F.
Soltis, P. S. and Soltis, D. E. (2016). "Ancient WGD events as drivers of key innovations in angiosperms." Current Opinion in Plant Biology **30**: 159-165.
Song, Y., et al. (2014). "Age-Triggered and Dark-Induced Leaf Senescence Require the bHLH Transcription Factors PIF3, 4, and 5." Molecular Plant **7**(12): 1776-1787.
Souer, E., et al. (1996). "*The no apical meristem gene* of Petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries." Cell **85**(2): 159-170.
Sternberg, N. and Hamilton, D. (1981). "Bacteriophage P1 site-specific recombination: I. Recombination between loxP sites." Journal of Molecular Biology **150**(4): 467-486.

Steuernagel, B., et al. (2016). "Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture." Nature Biotechnology **34**: 652.
Sun, X., et al. (2013). "Multifarious Roles of Intrinsic Disorder in Proteins Illustrate Its Broad Impact on Plant Biology." The Plant Cell **25**(1): 38.
Sykorova, B., et al. (2008). "Senescence-induced ectopic expression of the *A. tumefaciens ipt* gene in wheat delays leaf senescence, increases cytokinin content, nitrate influx, and nitrate reductase activity, but does not affect grain yield." Journal of Experimental Botany **59**(2): 377-387.
Tabbita, F., et al. (2017). "Breeding for increased grain protein and micronutrient content in wheat: Ten years of the *GPC-B1* gene." Journal of Cereal Science **73**: 183-191.
Tang, Y., et al. (2012). "Molecular characterization of novel *TaNAC* genes in wheat and overexpression of *TaNAC2a* confers drought tolerance in tobacco." Physiol Plant **144**(3): 210-224.
The Gene Ontology Consortium (2018). "The Gene Ontology Resource: 20 years and still GOing strong." Nucleic Acids Research **47**(D1): D330-D338.
Tran, L. S., et al. (2004). "Isolation and functional analysis of Arabidopsis stress-inducible NAC transcription factors that bind to a drought-responsive *cis*-element in the *early responsive to dehydration stress 1* promoter." The Plant Cell **16**(9): 2481-2498.
Tulchinsky, T. H. (2010). "Micronutrient Deficiency Conditions: Global Health Issues." Public Health Reviews **32**(1): 243-255.
U.S. Department of Agriculture, A. R. S. (2019). "FoodData Central." from fdc.nal.usda.gov.
Uauy, C. (2017). "Wheat genomics comes of age." Current Opinion in Plant Biology **36**: 142-148.
Uauy, C., et al. (2006a). "The high grain protein content gene *Gpc-B1* accelerates senescence and has pleiotropic effects on protein content in wheat." Journal of Experimental Botany **57**(11): 2785-2794.
Uauy, C., et al. (2006b). "A NAC Gene regulating senescence improves grain protein, zinc, and iron content in wheat." Science **314**(5803): 1298-1301.
Uauy, C., et al. (2009). "A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat." BMC Plant Biology **9**(1): 115.
Uauy, C., et al. (2017). "Combining Traditional Mutagenesis with New High-Throughput Sequencing and Genome Editing to Reveal Hidden Variation in Polyploid Wheat." Annual Review of Genetics **51**(1): 435-454.
Ueda, H. and Kusaba, M. (2015). "Strigolactone Regulates Leaf Senescence in Concert with Ethylene in Arabidopsis." Plant Physiology **169**(1): 138.
Vacic, V., et al. (2007). "Characterization of molecular recognition features, MoRFs, and their binding partners." J Proteome Res **6**(6): 2351-2366.
van Doorn, W. G. (2008). "Is the onset of senescence in leaf cells of intact plants due to low or high sugar levels?" Journal of Experimental Botany **59**(8): 1963-1972.
van Doorn, W. G. and Woltering, E. J. (2004). "Senescence and programmed cell death: substance or semantics?" Journal of Experimental Botany **55**(406): 2147-2153.
Vanhaeren, H., et al. (2017). "Forever Young: The Role of Ubiquitin Receptor DA1 and E3 Ligase BIG BROTHER in Controlling Leaf Growth and Development." Plant Physiology **173**(2): 1269.
Vaupel, J. W., et al. (2004). "The case for negative senescence." Theoretical Population Biology **65**(4): 339-351.
Vergunst, A. C., et al. (2000). "Cre/*lox*-mediated recombination in *Arabidopsis*: evidence for transmission of a translocation and a deletion event." Chromosoma **109**(4): 287-297.
Vilella, A. J., et al. (2009). "EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates." Genome Res **19**(2): 327-335.
Vogelmann, K., et al. (2012). "Early Senescence and Cell Death in Arabidopsis *saul1* Mutants Involves the *PAD4*-Dependent Salicylic Acid Pathway." Plant Physiology **159**(4): 1477.
Vrána, J., et al. (2016). "Flow Analysis and Sorting of Plant Chromosomes." Current Protocols in Cytometry **78**(1): 5.3.1-5.3.43.
Vrána, J., et al. (2000). "Flow Sorting of Mitotic Chromosomes in Common Wheat (*Triticum aestivum* L.)." Genetics **156**(4): 2033.
Wachek, M., et al. (2006). "Oligomerization of the Mg^{2+}-transport proteins Alr1p and Alr2p in yeast plasma membrane." Febs j **273**(18): 4236-4249.

Wang, B., et al. (2018a). "A novel wheat NAC transcription factor, TaNAC30, negatively regulates resistance of wheat to stripe rust." Journal of Integrative Plant Biology **60**(5): 432-443.
Wang, F., et al. (2015). "*TaNAC1* acts as a negative regulator of stripe rust resistance in wheat, enhances susceptibility to *Pseudomonas syringae*, and promotes lateral root development in transgenic *Arabidopsis thaliana*." Frontiers in Plant Science **6**: 108-108.
Wang, S., et al. (2014). "Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array." Plant Biotechnology Journal **12**(6): 787-796.
Wang, T. L., et al. (2012). "TILLING in extremis." Plant Biotechnology Journal **10**(7): 761-772.
Wang, W., et al. (2018b). "Gene editing and mutagenesis reveal inter-cultivar differences and additivity in the contribution of *TaGW2* homoeologues to grain size and weight in wheat." Theoretical and Applied Genetics **131**(11): 2463-2475.
Wang, Y., et al. (2016). "Fine mapping of a dominant gene conferring chlorophyll-deficiency in *Brassica napus*." Scientific Reports **6**: 31419.
Warnes, G. R., et al. (2019). gplots: Various R Programming Tools for Plotting Data. R package version 3.0.1.1.
Watanabe, M., et al. (2013). "Comprehensive Dissection of Spatiotemporal Metabolic Shifts in Primary, Secondary, and Lipid Metabolism during Developmental Senescence in Arabidopsis." Plant Physiology **162**(3): 1290.
Waters, B. M., et al. (2009). "Wheat (*Triticum aestivum*) NAM proteins regulate the translocation of iron, zinc, and nitrogen compounds from vegetative tissues to grain." Journal of Experimental Botany **60**(15): 4263-4274.
Watson, A., et al. (2018). "Speed breeding is a powerful tool to accelerate crop research and breeding." Nature Plants **4**(1): 23-29.
Wehner, G., et al. (2016). "Expression profiling of genes involved in drought stress and leaf senescence in juvenile barley." BMC Plant Biology **16**: 3.
Wellburn, A. R. (1994). "The Spectral Determination of Chlorophylls *a* and *b*, as well as Total Carotenoids, Using Various Solvents with Spectrophotometers of Different Resolution." Journal of Plant Physiology **144**(3): 307-313.
Welner, D. H., et al. (2012). "DNA binding by the plant-specific NAC transcription factors in crystal and solution: a firm link to WRKY and GCM transcription factors." Biochem J **444**(3): 395-404.
Whittal, A., et al. (2018). "Allelic variation of vernalization and photoperiod response genes in a diverse set of North American high latitude winter wheat genotypes." PLOS ONE **13**(8): e0203068-e0203068.
Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. New York, Springer-Verlag.
Wickham, H., et al. (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.0.1. .
Wickham, H. and Henry, L. (2018). tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions. R package version 0.8.2.
Wilhelm, E. P., et al. (2009). "Photoperiod insensitive *Ppd-A1a* mutations in tetraploid wheat (*Triticum durum* Desf.)." Theoretical and Applied Genetics **118**(2): 285-294.
Williams, G. C. (1957). "Pleiotropy, Natural Selection, and the Evolution of Senescence." Evolution **11**(4): 398-411.
Wingen, L. U., et al. (2014). "Establishing the A. E. Watkins landrace cultivar collection as a resource for systematic gene discovery in bread wheat." Theoretical and Applied Genetics **127**(8): 1831-1842.
Wingen, L. U., et al. (2017). "Wheat Landrace Genome Diversity." Genetics **205**(4): 1657.
Wingler, A., et al. (2009). "Sugars, senescence, and ageing in plants and heterotrophic organisms." Journal of Experimental Botany **60**(4): 1063-1066.
Woo, H. R., et al. (2001). "ORE9, an F-Box Protein That Regulates Leaf Senescence in Arabidopsis." The Plant Cell **13**(8): 1779.
Woo, H. R., et al. (2019). "Leaf Senescence: Systems and Dynamics Aspects." Annual Review of Plant Biology **70**(1): 347-376.
Worland, A. J. (1996). "The influence of flowering time genes on environmental adaptability in European wheats." Euphytica **89**(1): 49-57.
Wu, H., et al. (2018). "Candidate Genes for Yellow Leaf Color in Common Wheat (*Triticum aestivum* L.) and Major Related Metabolic Pathways according to Transcriptome Profiling." International journal of molecular sciences **19**(6).

Wu, H., et al. (2003). "Factors influencing successful *Agrobacterium*-mediated genetic transformation of wheat." Plant Cell Reports **21**(7): 659-668.
Wu, K., et al. (2008). "HDA6 is required for jasmonate response, senescence and flowering in *Arabidopsis*." Journal of Experimental Botany **59**(2): 225-234.
Wu, S., et al. (2016). "NON-YELLOWING2 (NYE2), a Close Paralog of NYE1, Plays a Positive Role in Chlorophyll Degradation in *Arabidopsis*." Molecular Plant **9**(4): 624-627.
Wysoker, A., et al. (2009). "The Sequence Alignment/Map format and SAMtools." Bioinformatics **25**(16): 2078-2079.
Xia, N., et al. (2010a). "Characterization of a novel wheat NAC transcription factor gene involved in defense response against stripe rust pathogen infection and abiotic stresses." Mol Biol Rep **37**(8): 3703-3712.
Xia, N., et al. (2010b). "*TaNAC8*, a novel NAC transcription factor gene in wheat, responds to stripe rust pathogen infection and abiotic stresses." Physiological and Molecular Plant Pathology **74**(5): 394-402.
Xiao, D., et al. (2015). "SENESCENCE-SUPPRESSED PROTEIN PHOSPHATASE Directly Interacts with the Cytoplasmic Domain of SENESCENCE-ASSOCIATED RECEPTOR-LIKE KINASE and Negatively Regulates Leaf Senescence in Arabidopsis." Plant Physiology **169**(2): 1275-1291.
Xie, Q., et al. (2000). "Arabidopsis *NAC1* transduces auxin signal downstream of *TIR1* to promote lateral root development." Genes & Development **14**(23): 3024-3036.
Xie, Y., et al. (2014). "REVOLUTA and WRKY53 connect early and late leaf development in *Arabidopsis*." Development **141**(24): 4772.
Xu, F., et al. (2011). "A Soybean Dual-Specificity Kinase, GmSARK, and Its Arabidopsis Homolog, AtSARK, Regulate Leaf Senescence through Synergistic Actions of Auxin and Ethylene." Plant Physiology **157**(4): 2131.
Xu, Z., et al. (2015). "Wheat NAC transcription factor TaNAC29 is involved in response to salt stress." Plant Physiology and Biochemistry **96**: 356-363.
Xue, G.-P., et al. (2006). "*TaNAC69* from the NAC superfamily of transcription factors is up-regulated by abiotic stresses in wheat and recognises two consensus DNA-binding sequences." Functional Plant Biology **33**(1): 43-57.
Xue, G.-P., et al. (2011). "Overexpression of *TaNAC69* Leads to Enhanced Transcript Levels of Stress Up-Regulated Genes and Dehydration Tolerance in Bread Wheat." Molecular Plant **4**(4): 697-712.
Yamaguchi-Shinozaki, K. and Shinozaki, K. (1993). "Characterization of the expression of a desiccation-responsive *rd29* gene of *Arabidopsis thaliana* and analysis of its promoter in transgenic plants." Molecular & General Genetics **236**(2-3): 331-340.
Yamaguchi, M., et al. (2010). "VND-INTERACTING2, a NAC Domain Transcription Factor, Negatively Regulates Xylem Vessel Formation in *Arabidopsis*." The Plant Cell **22**(4): 1249.
Yan, L., et al. (2006). "The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*." PNAS **103**(51): 19581-19586.
Yan, L., et al. (2004). "Allelic variation at the *VRN-1* promoter region in polyploid wheat." Theoretical and Applied Genetics **109**(8): 1677-1686.
Yan, L., et al. (2003). "Positional cloning of the wheat vernalization gene *VRN1*." PNAS **100**(10): 6263.
Yang, D., et al. (2016). "Exogenous Cytokinins Increase Grain Yield of Winter Wheat Cultivars by Improving Stay-Green Characteristics under Heat Stress." PLOS ONE **11**(5): e0155437.
Yang, J., et al. (2014). "A NAP-AAO3 Regulatory Module Promotes Chlorophyll Degradation via ABA Biosynthesis in *Arabidopsis* Leaves." The Plant Cell **26**(12): 4862.
Yang, J. C., et al. (2003). "Involvement of abscisic acid and cytokinins in the senescence and remobilization of carbon reserves in wheat subjected to water stress during grain filling." Plant, Cell & Environment **26**(10): 1621-1631.
Yin, X. and Komatsu, S. (2016). "Nuclear Proteomics Reveals the Role of Protein Synthesis and Chromatin Structure in Root Tip of Soybean during the Initial Stage of Flooding Stress." J Proteome Res **15**(7): 2283-2298.
Yoo, S.-D., et al. (2008). "Dual control of nuclear EIN3 by bifurcate MAPK cascades in C2H4 signalling." Nature **451**(7180): 789-795.

Yoshimoto, K., et al. (2009). "Autophagy Negatively Regulates Cell Death by Controlling NPR1-Dependent Salicylic Acid Signaling during Senescence and the Innate Immune Response in *Arabidopsis*." The Plant Cell **21**(9): 2914.
Young, M. D., et al. (2010). "Gene ontology analysis for RNA-seq: accounting for selection bias." Genome Biology **11**(2): R14.
Yu, J., et al. (2015). "JAZ7 negatively regulates dark-induced leaf senescence in *Arabidopsis*." Journal of Experimental Botany **67**(3): 751-762.
Zadoks, J. C., et al. (1974). "A decimal code for the growth stages of cereals." Weed Research **14**(6): 415-421.
Zhang, H., et al. (2014). "Evolution of the BBAA Component of Bread Wheat during Its History at the Allohexaploid Level." The Plant Cell **26**(7): 2761.
Zhang, J., et al. (2019). "Identification and characterization of a novel stay-green QTL that increases yield in maize." Plant Biotechnology Journal **0**(0).
Zhang, S., et al. (2017). "The Arabidopsis Mitochondrial Protease FtSH4 Is Involved in Leaf Senescence via Regulation of WRKY-Dependent Salicylic Acid Accumulation and Signaling." Plant Physiology **173**(4): 2294-2307.
Zhang, X., et al. (2010). "The R-R-Type MYB-Like Transcription Factor, AtMYBL, is Involved in Promoting Leaf Senescence and Modulates an Abiotic Stress Response in Arabidopsis." Plant and Cell Physiology **52**(1): 138-148.
Zhang, Y., et al. (2013). "Protein analysis by shotgun/bottom-up proteomics." Chemical reviews **113**(4): 2343-2394.
Zhang, Y., et al. (2016). "Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA." Nature Communications **7**(1): 12617.
Zhang, Y., et al. (2015a). "Mitogen-activated protein kinase 6 mediates nuclear translocation of ORE3 to promote *ORE9* gene expression in methyl jasmonate-induced leaf senescence." Journal of Experimental Botany **67**(1): 83-94.
Zhang, Y., et al. (2015b). "PHYTOCHROME-INTERACTING FACTOR 5 (PIF5) positively regulates dark-induced senescence and chlorophyll degradation in *Arabidopsis*." Plant Science **237**: 57-68.
Zhao, D., et al. (2015). "Overexpression of a NAC transcription factor delays leaf senescence and increases grain nitrogen concentration in wheat." Plant Biol (Stuttg) **17**(4): 904-913.
Zhao, G., et al. (2017). "The Aegilops tauschii genome reveals multiple impacts of transposons." Nature Plants **3**(12): 946-955.
Zheng, C., et al. (2016). "Global Transcriptional Analysis Reveals the Complex Relationship between Tea Quality, Leaf Senescence and the Responses to Cold-Drought Combined Stress in *Camellia sinensis*." Frontiers in plant science **7**: 1858.
Zhou, C., et al. (2009). "An Arabidopsis Mitogen-Activated Protein Kinase Cascade, MKK9-MPK6, Plays a Role in Leaf Senescence." Plant Physiology **150**(1): 167.
Zhou, J., et al. (2015). "The dominant negative ARM domain uncovers multiple functions of PUB13 in Arabidopsis immunity, flowering, and senescence." Journal of Experimental Botany **66**(11): 3353-3366.
Zhou, X., et al. (2011). "WRKY22 transcription factor mediates dark-induced leaf senescence in *Arabidopsis*." Mol Cells **31**(4): 303-313.
Zhu, X., et al. (2015). "Jasmonic acid promotes degreening via MYC2/3/4- and ANAC019/055/072-mediated regulation of major chlorophyll catabolic genes." The Plant Journal **84**(3): 597-610.
Zimin, A. V., et al. (2017). "The first near-complete assembly of the hexaploid bread wheat genome, *Triticum aestivum*." GigaScience **6**(11).
Zong, Y., et al. (2017). "Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion." Nature Biotechnology **35**: 438.
Zuo, J., et al. (2000). "Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants." The Plant Journal **24**(2): 265-273.

8 Appendix

8.1 Sequences

8.1.1 CDS and protein sequences of NAM-A1 and NAM-B1.

Sequences were used as wild-type controls for yeast 2-hybrid and *N. benthamiana* cell death assays in Chapter 3 and Chapter 5. Subdomains i – v are highlighted in green on the protein sequences.

> TraesCS6A01G108300.2; NAM-A1

CDS:

ATGAGGTCCATGGGCAGCTCCGACTCATCCTCCGGCTCGGCGCAAAAAGCAGCGCGGCATCAGCAT
GAGCCGCCGCCTCCGCGGCAGCGGGGCTCGGCGCCGGAGCTCCCACCGGGCTTCCGGTTCCACCCG
ACGGACGAGGAGCTGGTCGTGCACTACCTCAAGAAGAAGGCCGCCAAGGTGCCGCTCCCCGTCACC
ATCATCGCCGAGGTGGATCTCTACAAGTTCGACCCATGGGAGCTCCCCGAGAAGGCGACCTTCGGG
GAGCAGGAGTGGTACTTCTTCAGCCCGCGCGACCGCAAGTACCCCAACGGCGCGCGGCCGAACCGG
GCGGCGACGTCGGGCTACTGGAAGGCCACCGGCACGGACAAACCTATCCTGGCCTCGGGGACGGGG
TGCGGCCTGGTCCGGGAGAAGCTCGGCGTCAAGAAGGCGCTCGTCTTCTACCGCGGGAAGCCGCCC
AAGGGCCTCAAAACCAACTGGATCATGCACGAGTACCGCCTCACCGACGCGTCTGGCTCCACCACC
ACCAGCCGGCCGCCGCCGCCTGTGACCGGCGGGAGCCGGGCTGCAGCCTCTCTGAGGTTGGACGAC
TGGGTGCTGTGCCGCATCTACAAGAAGATCAACAAGGCCGCGGCCGGAGATCAGCAGAGGAGCACG
GAGTGCGAGGACTCCGTGGAGGACGCGGTCACCGCGTACCCGCTCTATGCCACGGCGGGCATGGCC
GGTGCAGGTGCGCATGGCAGCAACTACGCTTCACCTTCACTGCTCCATCATCAGGACAGCCATTTC
CTGGAGGGCCTGTTCACAGCAGACGACGCCGGCCTCTCGGCGGGCGCCACCTCGCTGAGCCACCTG
GCCGCGGCGGCGAGGGCGAGCCCGGCTCCGACCAAACAGTTTCTCGCCCCGTCGTCTTCAACCCCG
TTCAACTGGCTCGATGCGTCACCCGCCGGCATCCTGCCACAGGCAAGGAATTTCCCTGGGTTTAAC
AGGAGCAGAAACGTCGGCAATATGTCGCTGTCATCGACGGCCGACATGGCTGGCGCGGCCGGCAAT
GCGGTGAACGCCATGTCCGCATTTATGAATCCTCTCCCCGTGCAAGACGGGACGTACCATCAACAC
CATGTCATCCTCGGCGCCCCACTGGCGCCAGAGGCTACCACAGGCGGCGCCACCTCTGGTTTCCAG
CATCCCGTCCAAGTATCCGGCGTGAACTGGAATCCCTGA

Protein:

MRSMGSSDSSSGSAQKAARHQHEPPPPRQRGSAPELPPGFRFHPTDEELVVHYLKKKAAKVPLPVT
IIAEVDLYKFDPWELPEKATFGEQEWYFFSPRDRKYPNGARPNRAATSGYWKATGTDKPILASGTG
CGLVREKLGVKKALVFYRGKPPKGLKTNWIMHEYRLTDASGSTTTSRPPPPVTGGSRAAASLRLDD
WVLCRIYKKINKAAAGDQQRSTECEDSVEDAVTAYPLYATAGMAGAGAHGSNYASPSLLHHQDSHF
LEGLFTADDAGLSAGATSLSHLAAAARASPAPTKQFLAPSSSTPFNWLDASPAGILPQARNFPGFN
RSRNVGNMSLSSTADMAGAAGNAVNAMSAFMNPLPVQDGTYHQHHVILGAPLAPEATTGGATSGFQ
HPVQVSGVNWNP*

> TraesCS6A01G108300.2; NAM-A1 (NAC domain, 1-217)

CDS:

ATGAGGTCCATGGGCAGCTCCGACTCATCCTCCGGCTCGGCGCAAAAAGCAGCGCGGCATCAGCAT
GAGCCGCCGCCTCCGCGGCAGCGGGGCTCGGCGCCGGAGCTCCCACCGGGCTTCCGGTTCCACCCG
ACGGACGAGGAGCTGGTCGTGCACTACCTCAAGAAGAAGGCCGCCAAGGTGCCGCTCCCCGTCACC
ATCATCGCCGAGGTGGATCTCTACAAGTTCGACCCATGGGAGCTCCCCGAGAAGGCGACCTTCGGG
GAGCAGGAGTGGTACTTCTTCAGCCCGCGCGACCGCAAGTACCCCAACGGCGCGCGGCCGAACCGG
GCGGCGACGTCGGGCTACTGGAAGGCCACCGGCACGGACAAACCTATCCTGGCCTCGGGGACGGGG
TGCGGCCTGGTCCGGGAGAAGCTCGGCGTCAAGAAGGCGCTCGTCTTCTACCGCGGGAAGCCGCCC
AAGGGCCTCAAAACCAACTGGATCATGCACGAGTACCGCCTCACCGACGCGTCTGGCTCCACCACC
ACCAGCCGGCCGCCGCCGCCTGTGACCGGCGGGAGCCGGGCTGCAGCCTCTCTGAGGTTGGACGAC
TGGGTGCTGTGCCGCATCTACAAGAAGATCAACAAGGCCGCGGCCGGAGATCAGCAG

Protein:

MRSMGSSDSSSGSAQKAARHQHEPPPPRQRGSAPELPPGFRFHPTDEELVVHYLKKKAAKVPLPVT
IIAEVDLYKFDPWELPEKATFGEQEWYFFSPRDRKYPNGARPNRAATSGYWKATGTDKPILASGTG
CGLVREKLGVKKALVFYRGKPPKGLKTNWIMHEYRLTDASGSTTTSRPPPPVTGGSRAAASLRLDD
WVLCRIYKKINKAAAGDQQ*

> DQ869673.1; NAM-B1

CDS:

ATGGGCAGCTCCGACTCATCTTCCGGCTCGGCGCAAAAAGCAACGCGGTATCACCATCAGCATCAG
CCGCCGCCTCCGCAGCGGGGCTCGGCGCCGGAGCTCCCGCCGGGCTTCCGGTTCCACCCGACGGAC
GAGGAGCTGGTGGTGCACTACCTCAAGAAGAAGGCCGACAAGGCGCCGCTCCCCGTCAACATCATC
GCCGAGGTGGATCTCTACAAGTTCGACCCATGGGAGCTCCCCGAGAAGGCGACCATCGGGGAGCAG
GAGTGGTACTTCTTCAGCCCGCGCGACCGCAAGTACCCCAACGGCGCGCGGCCGAACCGGGCGGCG
ACGTCGGGGTACTGGAAGGCCACCGGCACGGACAAGCCTATCCTGGCCTCGGGGACGGGGTGCGGC
CTGGTCCGGGAGAAGCTCGGCGTCAAGAAGGCGCTCGTGTTCTACCGCGGGAAGCCGCCCAAGGGC
CTCAAAACCAACTGGATCATGCACGAGTACCGCCTCACCGACGCATCTGGCTCCACCACCGCCACC
AACCGACCGCCGCCGGTGACCGGCGGGAGCAGGGCTGCTGCCTCTCTCAGGTTGGACGACTGGGTG
CTGTGCCGCATCTACAAGAAGATCAACAAGGCCGCGGCCGGCGATCAGCAGAGGAACACGGAGTGC
GAGGACTCCGTGGAGGACGCGGTCACCGCGTACCCGCTCTATGCCACGGCGGGCATGACCGGTGCA
GGTGCGCATGGCAGCAACTACGCTTCACCTTCACTGCTCCATCATCAGGACAGCCATTTCCTGGAC
GGCCTGTTCACAGCAGACGACGCCGGCCTCTCGGCGGGCGCCACCTCGCTGAGCCACCTAGCAGCG
GCGGCGAGGGCAAGCCCGGCTCCGACCAAACAGTTTCTCGCCCCGTCGTCTTCGACCCCGTTCAAC
TGGCTCGATGCGTCACCAGTCGGCATCCTCCCACAGGCAAGGAATTTTCCTGGGTTTAACAGGAGC
AGAAACGTCGGCAATATGTCGCTGTCATCGACGGCCGACATGGCTGGCGCAGTGGACAACGGTGGA
GGCAATGCGGTGAACGCCATGTCTACCTATCTTCCCGTGCAAGACGGGACGTACCATCAGCAGCAT

GTCATCCTCGGCGCTCCGCTGGTGCCAGAAGCCGCCGCCGCCACCTCTGGATTCCAGCATCCCGTT
CAAATATCCGGCGTGAACTGGAATCCCTGA

Protein:

MGSSDSSSGSAQKATRYHHQHQPPPPQRGSAPELPPGFRFHPTDEELVVHYLKKKADKAPLPVNII
AEVDLYKFDPWELPEKATIGEQEWYFFSPRDRKYPNGARPNRAATSGYWKATGTDKPILASGTGCG
LVREKLGVKKALVFYRGKPPKGLKTNWIMHEYRLTDASGSTTATNRPPPVTGGSRAAASLRLDDWV
LCRIYKKINKAAAGDQQRNTECEDSVEDAVTAYPLYATAGMTGAGAHGSNYASPSLLHHQDSHFLD
GLFTADDAGLSAGATSLSHLAAAARASPAPTKQFLAPSSSTPFNWLDASPVGILPQARNFPGFNRS
RNVGNMSLSSTADMAGAVDNGGGNAVNAMSTYLPVQDGTYHQQHVILGAPLVPEAAAATSGFQHPV
QISGVNWNP*

> DQ869673.1; NAM-B1 (NAC domain, 1-215)

CDS:

ATGGGCAGCTCCGACTCATCTTCCGGCTCGGCGCAAAAAGCAACGCGGTATCACCATCAGCATCAG
CCGCCGCCTCCGCAGCGGGGCTCGGCGCCGGAGCTCCCGCCGGGCTTCCGGTTCCACCCGACGGAC
GAGGAGCTGGTGGTGCACTACCTCAAGAAGAAGGCCGACAAGGCGCCGCTCCCCGTCAACATCATC
GCCGAGGTGGATCTCTACAAGTTCGACCCATGGGAGCTCCCCGAGAAGGCGACCATCGGGGAGCAG
GAGTGGTACTTCTTCAGCCCGCGCGACCGCAAGTACCCCAACGGCGCGCGGCCGAACCGGGCGGCG
ACGTCGGGGTACTGGAAGGCCACCGGCACGGACAAGCCTATCCTGGCCTCGGGGACGGGGTGCGGC
CTGGTCCGGGAGAAGCTCGGCGTCAAGAAGGCGCTCGTGTTCTACCGCGGGAAGCCGCCCAAGGGC
CTCAAAACCAACTGGATCATGCACGAGTACCGCCTCACCGACGCATCTGGCTCCACCACCGCCACC
AACCGACCGCCGCCGGTGACCGGCGGGAGCAGGGCTGCTGCCTCTCTCAGGTTGGACGACTGGGTG
CTGTGCCGCATCTACAAGAAGATCAACAAGGCCGCGGCCGGCGATCAGCAGTGA

Protein:

MGSSDSSSGSAQKATRYHHQHQPPPPQRGSAPELPPGFRFHPTDEELVVHYLKKKADKAPLPVNII
AEVDLYKFDPWELPEKATIGEQEWYFFSPRDRKYPNGARPNRAATSGYWKATGTDKPILASGTGCG
LVREKLGVKKALVFYRGKPPKGLKTNWIMHEYRLTDASGSTTATNRPPPVTGGSRAAASLRLDDWV
LCRIYKKINKAAAGDQQ*

8.1.2 Domesticated NAM-B1 sequence used in the transgenic constructs

The TtNAM-B1 gene sequence was domesticated to remove any *Bbs*I or *Bsa*I sites. One *Bbs*I site was identified (highlighted in yellow in the third exon), and was domesticated from “GTCTTC” to “ATCGTC”, removing the enzyme cut site but retaining the correct amino acid sequence. Exons are coloured, with the splice sites highlighted in red. The exonic sequence is based on the sequence of *NAM-B1* from *T. turgidum* ssp. *dicoccoides*, while the intronic sequence is taken from the non-functional copy of *NAM-B1* present in Chinese Spring.

>full-length domesticated *NAM-B1*

ATGGGCAGCTCCGACTCATCTTCCGGCTCGGCGCAAAAAGCAACGCGGTATCACCATCAGCATCAGCCGCCGCCTCC
GCAGCGGGGCTCGGCGCCGGAGCTCCCGCCGGGCTTCCGGTTCCACCCGACGGACGAGGAGCTGGTGGTGCACTACC
TCAAGAAGAAGGCCGACAAGGCGCCGCTCCCCGTCAACATCATCGCCGAGGTGGATCTCTACAAGTTCGACCCATGG
GAGCTCCCCGGTATGTTATGTCTATCTCGTCGGCCGGCCGTGCTTACTTTATCAAGCGCCGCAAATTTTAGGTGCAA
TTAAATAATCGAATAATCCATCCATCTCATGCTTATACTCCTGTGCACAAGTAGTATTTTTATATTCTTCCAGTACA
CATGTGTGTAGATGGTTTATGTATGTGATCCTGTCGTGCTTGTTCATGCGCTCGGGATCCGGATCCATCAGAGAAGG
CGACCATCGGGGAGCAGGAGTGGTACTTCTTCAGCCCGCGCGACCGCAAGTACCCCAACGGCGCGCGGCCGAACCGG
GCGGCGACGTCGGGGTACTGGAAGGCCACCGGCACGGACAAGCCTATCCTGGCCTCGGGGACGGGGTGCGGCCTGGT
CCGGGAGAAGCTCGGCGTCAAGAAGGCGCTCGTGTTCTACCGCGGGAAGCCGCCCAAGGGCCTCAAAACCAACTGGA
TCATGCACGAGTACCGCCTCACCGACGCATCTGGCTCCACCACCGCCACCAACCGACCGCCGCCGGTGACCGGCGGG
AGCAGGGCTGCTGCCTCTCTCAGGGTACGTACACGTGTCGATCGCACGGTCTAGCAGTATTTAATTGCTCTCCAGCT
TAATTAGGGTATTGTTGATGGTTGATGAAGTTGGACGACTGGGTGCTGTGCCGCATCTACAAGAAGATCAACAAGGC
CGCGGCCGGCGATCAGCAGAGGAACACGGAGTGCGAGGACTCCGTGGAGGACGCGGTCACCGCGTACCCGCTCTATG
CCACGGCGGGCATGACCGGTGCAGGTGCGCATGGCAGCAACTACGCTTCACCTTCACTGCTCCATCATCAGGACAGC
CATTTCCTGGACGGCCTGTTCACAGCAGACGACGCCGGCCTCTCGGCGGGCGCCACCTCGCTGAGCCACCTAGCAGC
GGCGGCGAGGGCAAGCCCGGCTCCGACCAAACAGTTTCTCGCCCCGTCATCGTCGACCCCGTTCAACTGGCTCGATG
CGTCACCAGTCGGCATCCTCCCACAGGCAAGGAATTTTCCTGGGTTTAACAGGAGCAGAAACGTCGGCAATATGTCG
CTGTCATCGACGGCCGACATGGCTGGCGCAGTGGACAACGGTGGAGGCAATGCGGTGAACGCCATGTCTACCTATCT
TCCCGTGCAAGACGGGACGTACCATCAGCAGCATGTCATCCTCGGCGCTCCGCTGGTGCCAGAAGCCGCCGCCGCCA
CCTCTGGATTCCAGCATCCCGTTCAAATATCCGGCGTGAACTGGAATCCC**TGA**

8.1.3 CDS and protein sequences of the yeast two-hybrid candidate genes

The CDS and protein sequences for all six candidate genes identified via the yeast two-hybrid screen in chapter 5 are presented below, with the exception of *NAM-A1* which is described in section 8.1.1. The position of the premature stop codon is indicated in the sequences in red.

> TraesCS2A02G101100 CDS

ATGTCGGACGTGACGGCGGTGATGGATCTGGCGCTTCCACCGGGGTTCCGGTTCCACC
CCACCGACGAGGAGGTGGTCACCCACTACCTCACCCGCAAGGTCCTCCGCGAATCCTT
CTCCTGCCAGGTGATCACCGACGTCGACCTCAACAAGAACGAGCCGTGGGAGCTCCCG
GGGAAGGCGAAGATGGGGGAGAAGGAGTGGTTCTTCTTCGTGCACAAGGGTCGGAAG
TACCCGACGGGAACGCGCACCAACCGGGCTACGAAGAAGGGGTACTGGAAGGCGACG
GGGAAGGACAAGGAGATCTTCCGCGGCAAGGGCCGGGACGCCGTCCTCGTCGGCATG
AAGAAGACGCTCGTCTTCTACACCGGCCGCGCCCCCAGCGGCGGGAAGACGCCGTGG
GTGATGCACGAGTACCGCCTCGAGGGCCAGCTGCCCCATCGTCTTCCCCACACCGCCA
AGGACGATTGGGCTGTTTGCCGGGTGTTCAACAAAGACTTGGCGGCGAAGAATGCGCC
GCCCCAGATGGCGCCGGCGGCTGACGGGGCCATGGAGGACCCGCTCGCCTTC[CTC/TG
A]GATGACTTGCTCATCGACACCAACCTCAACCTGTTCGACGACGCGGACCTGCCGATG
CTCATGGACTCTCCGTCTGGCGCTGACGACTTCGCCGGCGCTTCGAGCTCCACCTCCAG

CGCCGCCCTGCCGCTCGAGCCGGACGCGGAGCATTTGACTATCAAGATGGAGCCGCCG
CCGCAGCAGCAGCAGCAGATGCAGAGCCCAGACTACTTCATGCCGGCGACGGCCAAC
GGCAATCTTGGCGGTGCCGGGTACTCAACCTACCAGGCTGTGGTGGACCAGCAGGCCG
CGATCCGCATGTACTGCAAGCCGAAGGCGGAGGTAGCGTCTTCGTCGGCGCTGGGCTT
GGACATGGGGGCGCTCGCCGGCGCGGACACCTCGTTCCTGATGCCGTCGTCGCGGTCG
TACCTCGATCTGGAGGAGCTGTTCCGGGGCGAGGCTCTCATGGACTACTCCGACATGT
GGAAGATCTGA

> TraesCS2A02G101100 Protein

MSDVTAVMDLALPPGFRFHPTDEEVVTHYLTRKVLRESFSCQVITDVDLNKNEPWELPGK
AKMGEKEWFFFVHKGRKYPTGTRTNRATKKGYWKATGKDKEIFRGKGRDAVLVGMKK
TLVFYTGRAPSGGKTPWVMHEYRLEGQLPHRLPHTAKDDWAVCRVFNKDLAAKNAPPQ
MAPAADGAMEDPLAF[L/*]DDLLIDTNLNLFDDADLPMLMDSPSGADDFAGASSSTSSAA
LPLEPDAEHLTIKMEPPPQQQQQMQSPDYFMPATANGNLGGAGYSTYQAVVDQQAAIRM
YCKPKAEVASSSALGLDMGALAGADTSFLMPSSRSYLDLEELFRGEALMDYSDMWKI*

> TraesCS2B02G118200 CDS

ATGTCGGACGTGACGGCGGTGATGGATCTGGAGGTGGAGGAGCCGCAGCTGGCGCTT
CCACCGGGGTTCCGGTTCCACCCCACCGACGAGGAGGTGGTCACCCACTACCTCACCC
GCAAGGTCCTCCGCGAATCCTTCTCCTGCCAAGTGATCACCGACGTCGACCTCAACAA
GAACGAGCCGTGGGAGCTCCCGGGCCTCGCGAAGATGGGCGAGAAGGAGTGGTTCTT
CTTCGTGCACAAGGGTCGGAAGTACCCGACGGGGACGCGCACCAACCGGGCGACGAA
GAAGGGGTACTGGAAGGCGACGGGGAAGGACAAGGAGATCTTCCGCGGCAAGGGCC
GGGACGCCGTCCTCGTCGGCATGAAGAAGACGCTCGTCTTCTACACCGGCCGCGCCCC
CAGCGGCGGGAAGACGCCGTGGGTGATGCACGAGTACCGCCTCGAGGGCGAGCTGCC
CCATCGCCTTCCCCGCACCGCCAAGGACGATTGGGCTGTTTGCCGGGTGTTCAACAAA
GACTTGGCGGCG[AGG/TGA]AATGCGCCCCAGATGGCGCCGGCGGCCGACGGTGGCAT
GGAGGACCCGCTCGCCTTCCTCGATGACTTGCTCATCGACACCGACCTGTTCGACGAC
GCGGACCTGCCGATGCTCATGGACTCTCCGTCTGGCGCTGACGACTTCGCCGGCGCTT
CGAGCTCCACCTGCAGCGCGGCCCTGCCGCTTGAGCCGGACGCGGAGCTACCGGTGCT
GCATCCGCAGCAGCAGCAGAGCCCCAACTACTTCTTCATGCCGGCGACGGCCAACGGC
AATCTTGGCGGCGCCGAGTACTCACCCTACCAGGCTATGGGGGACCAGCAGGCCGCG
ATCCGCAGGTACTGCAAGCCGAAGGCGGAGGTAGCGTCTTCGTCGGCGCTGCTGAGCC
CTTCGCTGGGCTTGGACACGGCGGCGCTTGCCGGCGCGGAGACCTCCTTCCTGATGCC
GTCATCGCGGTCGTACCTCGATCTGGAGGAGCTGTTCCGGGGCGAGCCTCTCATGGAC
TACTCCAACATGTGGAAGATCTGA

> TraesCS2B02G118200 Protein

MSDVTAVMDLEVEEPQLALPPGFRFHPTDEEVVTHYLTRKVLRESFSCQVITDVDLNKNE
PWELPGLAKMGEKEWFFFVHKGRKYPTGTRTNRATKKGYWKATGKDKEIFRGKGRDAV
LVGMKKTLVFYTGRAPSGGKTPWVMHEYRLEGELPHRLPRTAKDDWAVCRVFNKDLAA
[R/*]NAPQMAPAADGGMEDPLAFLDDLLIDTDLFDDADLPMLMDSPSGADDFAGASSSTC
SAALPLEPDAELPVLHPQQQQSPNYFFMPATANGNLGGAEYSPYQAMGDQQAAIRRYCK
PKAEVASSSALLSPSLGLDTAALAGAETSFLMPSSRSYLDLEELFRGEPLMDYSNMWKI*

> TraesCS2B02G118500 CDS

ATGTCGGACGTGACGGCGGTGATGGATCTGGAGGTGGAGGAGCCGCGGCTGTCGCTG
CCGCCGGGCTTCCGGTTCCACCCCACCGACGAGGAGGTGGTCACGCACTACCTCACCC
CCAAGGCCGTCAACAACGCCTTCTCCTGCCTCGTGATCGCCGACGTCGACCTCAACAA
GACGGAGCCGTGGGATCTCCCCGGGAAGGCGAAGATGGGGGAGAAGGAGTGGTACTT
CTTCGTGCACAAGGATCGCAAGTACCCGACGGGGACGCGCACCAACCGCGCCACGGA
GAAGGGGTACTGGAAGGCGACGGGGAAGGACAAGGAGATCTTCCGCGGCAAGGGCC
GGGACGCCGTGCTCGTCGGCATGAAGAAGACGCTCGTCTTCTACACCGGCCGCGCGCC
CCGCGGCGACAAGACGCCGTACGTCATGCACGAGTACCGCCTCGAGGGCCAGCTGCC
CCACCGCCTCCCCCGCTCCGCCAAGAACGATTGGGCTGTTTGCCGGGTGTTCGACAAG
GACTTAGCGGCGAAGAATGCGCCGCCC[CCG/TGA]ATGGCGCCGGCGGCCGTCGGGGT
CATGGAGGACCCGTACGCCTTCCTCGACGTCGACGACTTCCTCAACAACCCCGACCTG
CTCAACAACGCCGACCTGCCGATGCTCATGGACTCCCCGTCAGGCGCCGACGACTTCG
CCGGCGCCTCGAGCTCCACCTCCAGCGCCGCCGTGCCGCTCGAGCCGGACATGGAGCA
CCTTTCCATCAAGACGGAGCCACCGGTGCCGCAGCAGCAGATGCAGAGCCCCAACTAC
TTCTTCATGCCGGCGGCGACGCCCAACGGCAACCACGGCGGCGGCGGGTACTCACCCT
ACCAGGCCATGGGGGACCAGCAGACCGCGATCCGCAGGCACTGCAAGCCGGAGGCGG
CGTCTTCGTCGGCGCTGCTGAGCCCTTCGCTCGGCTTCGACGCGGGGGCGCTCGCCGG
CGCGGACACCTCGTTCCTGATGCCGTCGTCGCGGTCGTACCTCGATCTGGAGGCCATG
GACTACTCCAACATGTGGAAGATCTGA

> TraesCS2B02G118500 Protein

MSDVTAVMDLEVEEPRLSLPPGFRFHPTDEEVVTHYLTPKAVNNAFSCLVIADVDLNKTE
PWDLPGKAKMGEKEWYFFVHKDRKYPTGTRTNRATEKGYWKATGKDKEIFRGKGRDA
VLVGMKKTLVFYTGRAPRGDKTPYVMHEYRLEGQLPHRLPRSAKNDWAVCRVFDKDLA
AKNAPP[P/*]MAPAAVGVMEDPYAFLDVDDFLNNPDLLNNADLPMLMDSPSGADDFAGA
SSSTSSAAVPLEPDMEHLSIKTEPPVPQQQMQSPNYFFMPAATPNGNHGGGGYSPYQAMG
DQQTAIRRHCKPEAASSSALLSPSLGFDAGALAGADTSFLMPSSRSYLDLEAMDYSNMWK
I*

> TraesCS3B02G371200 CDS

ATGATGACGGCAATGGTGAAGGCCGAGCCGACGGCGGCGGAGGGGAGCAGCGGGCG
GAGGGACGCGGAGGCGGAGCTCAACCTGCCGCCGGGCTTCCGCTTCCACCCCACGGA
CGACGAGCTCGTCGTGCACTACCTCTGCAGGAAGGTGGCGGGCCAGCCGCAGCCCGTG
CCCATCATCGCCGAGGTCGATCTCTACAAGTTCAACCCCTGGGATCTGCCCGAGAGGG
CGCTGTTCGGGAGCAGGGAGTGGTACTTCTTCACGCCACGCGACCGCAAGTACCCCAA
CGGCTCGCGCCCCAACCGCTCCGCCGGCACCGGCTACTGGAAGGCCACCGGCGCCGAC
AAGCCCGTGGCGCCCAAGGAGAGCGGCGGCAGGACGGTCGGCATCAAGAAGGCGCTC
GTCTTCTACTCCGGCAGGGCGCCCAGGGGCGTCAAGACCGACTGGATCATGCACGAGT
ACCGCATCGCCGAAGCCGACCGCGCACCCGGCAAGAAGGGATCCCTCAAGCTGGACG
AATGGGTGCTGTGCCGGCTGTACAACAAGAAGAACAACTGGGAGAAGGTCAAGGTGG
AG[CAG/TGA]GACATGGCTGTGGAGGCCGGGCCAAACGGGGAGGTCATGGACGCGCT
GGCGACCGACGCCATATCCGACAGCTTCCAGACGCACGACTCCTCGGAGATCGACAGC
GCCTCCGGCCTGCAGCAGCGCGTCTTCACGGACATGGCGCAGGGGCAGCCGAGAGGC
GGCATGGTGACGGTGAAGGAGGACAGCGACTGGTTCACCGGCCTGAGCATGGACGAC
CTGCAGACTTGCTATATGAACATGGGGCAGATGGTGAACCCGGCGACGATGCCTGGGC
AGGACGGCAGCGGCTACCTACAGTCGATGAGCTCACCGCAGATGATGAGGCCGATGT
GGCAAACAATCTTGCCACCATTCTGA

> TraesCS3B02G371200 Protein

MMTAMVKAEPTAAEGSSGRRDAEAELNLPPGFRFHPTDDELVVHYLCRKVAGQPQPVPII
AEVDLYKFNPWDLPERALFGSREWYFFTPRDRKYPNGSRPNRSAGTGYWKATGADKPVA
PKESGGRTVGIKKALVFYSGRAPRGVKTDWIMHEYRIAEADRAPGKKGSLKLDEWVLCR
LYNKKNNWEKVKVE[Q/*]DMAVEAGPNGEVMDALATDAISDSFQTHDSSEIDSASGLQQR
VFTDMAQGQPRGGMVTVKEDSDWFTGLSMDDLQTCYMNMGQMVNPATMPGQDGSGY
LQSMSSPQMMRPMWQTILPPF*

> TraesCS5A02G049100 CDS

ATGTCGATGAGCTTCCTGAGCATGGTGGAGACGGAGCTGCCGCCGGGGTTCCGGTTCC
ACCCGAGAGACGACGAGCTCATCTGCGACTACCTCGCGCCCAAGGTCACCGGCAAGG
TCGGCTTCTCCGGCCGCCGCCCGCCCATGGTCGACGTCGACCTCAACAAGGTCGAGCC
ATGGGACCTCCCCGTGACTGCATCGGTGGGTGGCAAGGAGTGGTATTTCTATAGCCTA
AAGGATCGGAAATATGCGACGGGGCAGCGCACGAACCGAGCTACTGTGTCGGGGTAC
TGGAAGGCAACCGGGAAGGATAGGGTGGTGGCACGACGTGGTGCGCTAGTGGGGATG
AGGAAGACATTGGTGTTCTACCAAGGGAGGGCCCCTAAGGGGAGGAAGACGGAGTGG
GTGATGCACGAATACAGGTTGGAGGGTGCGCATGAGCAAGCTTCCAAGCAGGAGGAC
TGGGTCCTGTGCAGAGTCATCTGCAAGAAGAAATCAGGAGTTGGTGCCACCCCCAGGC
CAAGGAACCTCAGCAACATT[GTC/TGA]CATGGCACACCCACAGACACCTCCTCACCA
CCACTGCCACCTCTCATGGACACCACCCTAGCACAGCTCCAGGCCACCATGAACACCT

CCGCCGCCGCTGCCGCCGCCGCACTTGAGCAAGTGCCCTGCTTCTCCAGCTTCAGTAAT
AACAGTGCCAGCAACAGCAGCTACCTCCCCATGGTGACAGGCAACAGTAATGGCATG
GGCTACATGGACCATGGCCTGCCTGACTTTGGGAGCTACCTAGACCCCGCCATGAACT
GTGACAAGAAGGTCCTTAAGGCAGTGCTGAGCCAGCTTAGCTCCATGGGTGGTGAGGT
GGTGCCGAGCATGTCTCCTCAGATGGCTGCTGCAGTGAGCTCTACTTGGAACCACTTCT
AG

> TraesCS5A02G049100 Protein

MSMSFLSMVETELPPGFRFHPRDDELICDYLAPKVTGKVGFSGRRPPMVDVDLNKVEPWD
LPVTASVGGKEWYFYSLKDRKYATGQRTNRATVSGYWKATGKDRVVARRGALVGMRK
TLVFYQGRAPKGRKTEWVMHEYRLEGAHEQASKQEDWVLCRVICKKKSGVGATPRPRN
LSNI[V/*]HGTPTDTSSPPLPPLMDTTLAQLQATMNTSAAAAAAALEQVPCFSSFSNNSASN
SSYLPMVTGNSNGMGYMDHGLPDFGSYLDPAMNCDKKVLKAVLSQLSSMGGEVVPSMS
PQMAAAVSSTWNHF*

> TraesCS5B02G054200 CDS

ATGTCGATGAGCTTCCTGAGCATGGTGGAGACGGAGCTGCCGCCGGGGTTCCGGTTCC
ACCCGAGAGACGACGAGCTCATCTGCGACTACCTCGCGCCCAAGGTCACCGGCAAGG
TCGGCTTCTCCGGCCGCCGCCCGCCCATGGTCGACGTCGACCTCAACAAGGTCGAGCC
ATGGGACCTCCCTGTGACTGCATCGGTGGGCGGCAAGGAATGGTATTTCTATAGCCTA
AAGGATCGGAAATACGCGACGGGGCAGCGCACGAACCGAGCTACTGTGTCGGGATAC
TGGAAGGCGACCGGGAAGGATAGGGTGGTGGCACGACGAGGCGCCCTAGTGGGGATG
AGGAAGACATTGGTGTTCTACCAAGGGAGGGCCCCTAAAGGGAGGAAGACTGAGTGG
GTGATGCATGAATACAGGTTGGAGGGTGCACATGAGCAAGCTTCCAAGCAGGAGGAC
TGGGTCCTGTGCAGAGTCATCTGCAAGAAGAAATCAGGAGTTGGTGCCACCCCCAGGC
CAAGGAACCTCACCAACATT[GTC/TGA]CATGGCACATCCACAGACACCTCCTCACCGC
CACTGCCACCTCTCATGGACACCACCCTAGCACAGCTCCAGGCCACCATGAACACCTC
CTCAGCCGCCGCCGCCGCCGCAGCACTTGAGCAAGTGCCCTGCTTCTCCAGCTTCAGT
AATAACAGTGCCAGCAACAGCAGCTACCTCCCCATGGTGACAGGCAACAGTAATGGC
ATGAGCTACCTGGACCATGGCCTGCCTGACTTTGGGAGCTACCTAGACCCCGCCATGA
ACTGTGACAAGAAGGTCCTTAAGGCAGTGCTGAGCCAGCTGAGCTCCATGGGTGGTGA
GGTGGTGCCGAGCATGTCTCCTCAGATGGCTGCTGCAGTGAGCTCTACTTGGAACCAC
TTCTAG

> TraesCS5B02G054200 Protein

MSMSFLSMVETELPPGFRFHPRDDELICDYLAPKVTGKVGFSGRRPPMVDVDLNKVEPWD
LPVTASVGGKEWYFYSLKDRKYATGQRTNRATVSGYWKATGKDRVVARRGALVGMRK
TLVFYQGRAPKGRKTEWVMHEYRLEGAHEQASKQEDWVLCRVICKKKSGVGATPRPRN

LTNI[V/*]HGTSTDTSSPPLPPLMDTTLAQLQATMNTSSAAAAAAALEQVPCFSSFSNNSAS
NSSYLPMVTGNSNGMSYLDHGLPDFGSYLDPAMNCDKKVLKAVLSQLSSMGGEVVPSMS
PQMAAAVSSTWNHF*

8.2 Scripts

8.2.1 Coordinate conversion from CSS to RefSeqv1.0 genomes

Create Coordinate_Conversion_IWGSCtoWGAv1_Kronos.sh; available at
https://github.com/Uauy-Lab/K2282_scripts

```
#!/bin/sh
#SBATCH -p nbi-medium # partition (queue)
#SBATCH -N 1
#SBATCH -n 1
#SBATCH --mem=15000
#SBATCH -J BLAST_database_v1.0tetraploid
#SBATCH -o BLAST_database_v1.0tetraploid.%N.%j.out
#SBATCH -e BLAST_database_v1.0tetraploid.%N.%j.err
#SBATCH --mail-type=END,FAIL
#SBATCH --mail-user=sophie.harrington@jic.ac.uk

#Conversion of coordinates for the original MAPS-derived SNPs (in this case from line
K2282)
#the MAPS SNP file for each line can be downloaded from www.wheat-tilling.com

# To do this, we want to BLAST the SNP, including 150 bp of flanking sequence on either
side, against the NRGene assembly. This contig size should give us enough confidence in the
# position of the alignment and yet be small enough to avoid InDels that might disrupt the
alignment.

Input='K2282_maps_main_set.txt'
filename='K2282_maps_main_set_KronosRef'

#Make a directory to hold the coordinate conversion, and change into that directory
cd /nbi/group-data/ifs/NBI/Cristobal-Uauy/Sophie/ENQ2022/CoordinateConversion

#Construct a BED file with all SNP positions
#Extract first and second column, with Chrom/Scaffold and position; exclude first row for
header
sed -n '2,$ p' $Input | cut -f 1 /dev/stdin > $filename-Chrom.tab
sed -n '2,$ p' $Input | cut -f 2 /dev/stdin > $filename-Pos.tab

#create a third fle with every entry changed to the '_' symbol, same number of lines
sed 's/[0-9]*/_/ g' $filename-Pos.tab > $filename-underscore.tab

#Combine the three files together to give each SNP a unique ID (-d '\0' means no delimiter
included when pasting)
paste -d '\0' \
               $filename-Chrom.tab \
               $filename-underscore.tab \
               $filename-Pos.tab \
```

```
                > $filename-uniqueID.tab


#Now, to make coordinates base 0, need to subtract 1 from each entry, by creating one file
where each line entry is '1', then combine with the position file
sed 's/[0-9]*/1/ g' $filename-Pos.tab > $filename-1.tab
paste \
                $filename-Pos.tab \
                $filename-1.tab \
                > $filename-Pos-1.tab

#This allows the two columns to be subtracted from each other to give the new, base 0
position
awk '{ $3 = $1 - $2; print $3 }' $filename-Pos-1.tab > $filename-substracted.tab

#Now combine all of these files to make the correct BED file
paste \
                $filename-Chrom.tab \
                $filename-substracted.tab \
                $filename-Pos.tab \
                $filename-uniqueID.tab \
                > $filename.bed


#Next step: Extend coordinates to 150bp either side of the SNP, but use fasta index to
prevent extension past the start/end of scaffold.
#Use the command "slop" from BEDTools --> might be better just to have 300 bo

source bedtools-2.24.0; bedtools slop \
                -i $filename.bed \
                -g /nbi/group-data/ifs/NBI/Cristobal-Uauy/IWGSC_data/IWGSC_all/IWGSC-
all_UCW-Kronos_U.fa.fai \
                -l 300 \
                -r 0 \
                > $filename-slop-left-300.bed

#Using the modified BED file, extract SNP and flanking sequence using "getfasta" in
BEDTools; -name give the fasta file the unique ID from the BED file
source bedtools-2.24.0; bedtools getfasta \
                -fi /nbi/group-data/ifs/NBI/Cristobal-Uauy/IWGSC_data/IWGSC_all/IWGSC-
all_UCW-Kronos_U.fa \
                -bed $filename-slop-left-300.bed \
                -fo $filename-slop-left-300.fa \
                -name

#Finally, index the fasta file to get the length and name of each entry
source samtools-1.3; samtools faidx $filename-slop-left-300.fa

#Preparation is complete- now time to BLAST against the NRGene assembly

#To prevent odd mismatching that has been seen when BLASTing large numbers of fasta files
at once, separate each FASTA entry into a separate file (split every other lines to get the
two lines of the fasta entries out)
```

```
mkdir -p split-files-300
split -l 2 --suffix-length=4 --additional-suffix=.txt $filename-slop-left-300.fa ./split-
files-300/$filename-slop-left-300_split-

#Now cycle through all of the fasta files and BLAST against the NRGene, creating a separate
output file for each. Use stringent parameters (single hit and single HSP)

cd ./split-files-300/
mkdir ./split-BLAST-output-slop-left-300-Kronos
for file in $filename-slop-left-300_split-*; do
        source blast+-2.2.30; blastn \
                -task megablast \
                -db /nbi/group-data/ifs/NBI/Cristobal-
Uauy/WGAv1.0/161010_Chinese_Spring_v1.0_pseudomolecules_parts_tetraploid \
                -query $file \
                -max_target_seqs 1 \
                -max_hsps 1 \
                -outfmt "6 qseqid sseqid pident length mismatch gapopen qstart qend sstart
send evalue bitscore btop" \
                -parse_deflines \
                -out ./split-BLAST-output-slop-left-300-Kronos/WGAv1.0_parts-megablast-$file
        done

#Append all of the output files together; can't use cat due to number of files opened, so
use for-loop to append each file one at a time
cd ./split-BLAST-output-slop-left-300-Kronos
for file in WGAv1.0_parts-megablast-$filename-slop-left-300_split-*; do \
        cat $file >> /nbi/group-data/ifs/NBI/Cristobal-
Uauy/Sophie/ENQ2022/CoordinateConversion/$filename-slop-left-
300_vs_WGAv1.0_parts_tetraploid-megablast.txt \
        ; done

#for neatness, remove empty spaces from Bitscore column
sed 's/  // g' \
                /nbi/group-data/ifs/NBI/Cristobal-
Uauy/Sophie/ENQ2022/CoordinateConversion/$filename-slop-left-
300_vs_WGAv1.0_parts_tetraploid-megablast.txt \
                > /nbi/group-data/ifs/NBI/Cristobal-
Uauy/Sophie/ENQ2022/CoordinateConversion/$filename-slop-left-
300_vs_WGAv1.0_parts_tetraploid-megablast-sed.txt

#Now convert coordinates -> go into Excel and use "VLOOKUP" options as detailed below to
finalise the data:

## 1) Add columns for the positions flanking the SNP (i.e. the left and right positions
flanking the BLAST result; should be 150bp or less for each, based on the BLAST
requirements)
##              To do this, add the left and right positions from the -150.bed file created
earlier as the "left position" and "right position" columns. Then subtract the "left
position" from the SNP position to get "bp left of SNP", and similarly to obtain the right
side distance.
```

```
## 2) Subtract the send and sstart columns (from the BLAST output) to determine the
orientation of the BLAST sequence (i.e. a negative number is a (-) orientation region)

## 3) Get the correct SNP position for the RefSeq v1.0 assembly
##             3a) if the sequence is in the + orientation (See #2 above), use the
following formula: "subject-start" + "bp left of SNP" - "query-start" + 1
##             3b) if the sequence is in the - orientation (see #2 above), use the
following formula: "subject-start" - "bp left of SNP" + "query-start" - 1
#                      3.1)

## other columns can also be added to the file as desired from i.e. the original SNP set
(such as the WT and Mut allele etc.)
```

8.2.2 Comparison of NAM-A1 GENIE3 targets and nam-a1 DEGs

This R markdown script is available at https://github.com/Uauy-Lab/GENIE3_scripts/ and details how to extract genes and targets from the GENIE3 database and compare them against the *nam-a1* DEGs, which are available as a supplementary file to Harrington et al. (2019a).

```
---
title: "Random TF to Pearce et al KO data"
author: "Sophie Harrington"
date: "19 March 2019"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
library(dplyr)
library(ggplot2)
library(tidyr)
library(BSDA) ##for sign test
```

Compare overlap between random TFs downstream genes and the Pearce TILLING KO RNASeq data


```{r download data}
##GENIE3 data is available at: https://doi.ipk-gatersleben.de/DOI/53148abd-26a1-4ede-802b-
c2635af6a725/ef7a50bf-54bb-4041-b922-4e60af47dd14/0/1847940088
setwd("Z:/Lab Organisation/Resources/Transcriptome_networks/for_multiple_networks/")
load("top_1_million_genie3.Rda")
##get the TFs from the genie3 list
TF <- as.data.frame(distinct(link.list, from.gene))
TF <- TF[,"from.gene"]
##only keep genes not in D genome
TF <- grep("D01G", TF, value=TRUE, invert=TRUE)
##DEG output is available in Supplementary Table 2, in Harrington et al. 2019, biorXiv.
##get RNAseq output from NAM-A1 single mutants
NAMA1_vsWT_12DAANF <-
read.csv("Y://Sophie/Networks/Genie3_NAMcomparison/Anna_RNAseqoutput/12DAA_gpcA1_ref_WT_NOF
ILTERING.txt", sep=" ", header=TRUE)
```
```

```
NAMA1_vsWT_22DAA <-
read.csv("Y://Sophie/Networks/Genie3_NAMcomparison/Anna_RNAseqoutput/22DAA_gpcA1_ref_WT.txt
", sep=" ", header=TRUE)
##remove transcript ids from the RNA data
NAMA1_vsWT_12DAANF <- separate(NAMA1_vsWT_12DAANF, target_id, c("target_id","transcript"),
sep="\\.")
NAMA1_vsWT_12DAANF$name <- c("NAMA1_vsWT_12DAANF")
NAMA1_vsWT_22DAA <- separate(NAMA1_vsWT_22DAA, target_id, c("target_id","transcript"),
sep="\\.")
NAMA1_vsWT_22DAA$name <- c("NAMA1_vsWT_22DAA")
list_RNA <- list(NAMA1_vsWT_12DAANF,NAMA1_vsWT_22DAA)
```

Iterate through all of the transcription factors in the A/B genomes in the genie3 list and compare against each of the RNAseq datasets.

```{r iterate through the samples, warning=FALSE}
##want to carry out 1000 iterations of the following:
## 1)Randomly select 2 TFs from the list
## 2)Extract all downstream genes from genie3 database for each TF
## 3)Determine how many genes are in common
## 4)Store this value
##
## repeat this 1000 times, then plot the resulting distribution.
##create a matrix with all of the TFs and all of the RNA samples
TF_RNAdatasets <- expand.grid(TF,list_RNA)
##add column with the name of the RNA list
genie3_KO_comparison <- function(RNA, gene){
  ##randomly select 1 TF
  ##get all downstream genes
  gene1 <- link.list %>% filter(from.gene==gene)

  ##remove the 01G for 02G
  gene1$to.gene <- gsub("01G", "02G", gene1$to.gene)

  ##remove all D genome genes
  gene1 <- filter(gene1, !grepl("D02G", to.gene))

  ##get number of shared genes8
  shared <- length(intersect(gene1[,"to.gene"], RNA$target_id))
  num_gene1 <- nrow(gene1)
  num_gene2 <- nrow(RNA)

  ##get ratio
  if(num_gene1 < num_gene2 || num_gene1 == num_gene2){
    ratio <- shared/num_gene1
  } else if(num_gene2 < num_gene1){
    ratio <- shared/num_gene2
  }
  return(data.frame("RNA_Dataset"=RNA$name[1],"Genie3_Gene"=gene,"Ratio"=ratio, "RNA-Seq
genes"=num_gene2,"Genie3 DS Genes"=num_gene1, "Number shared"=shared))
}
```

```
output_rel <- mapply(genie3_KO_comparison, RNA=TF_RNAdatasets[,2], gene=TF_RNAdatasets[,1],
SIMPLIFY = FALSE)
output_rel <- bind_rows(output_rel)
write.csv(output_rel,
"Y://Sophie/Networks/Genie3_NAMcomparison/linklist_Genie3_NAMSharedDist.csv")
```

Now plot the distribution of the ratios, and include the intercept lines for the different NAM-A1 cases; This is Figure 1B.

```{r pretty graph}
output_rel <-
read.csv("Y://Sophie/Networks/Genie3_NAMcomparison/linklist_Genie3_NAMSharedDist.csv")
##subset output_rel
output_12_22 <- subset(output_rel, RNA_Dataset %in%
c("NAMA1_vsWT_12DAANF","NAMA1_vsWT_22DAA"))
plot.labels <- c("12 DAA","22 DAA")
names(plot.labels) <- c("NAMA1_vsWT_12DAANF","NAMA1_vsWT_22DAA")
##include the intercept lines for the NAM-A1 overlap with each RNA-Seq dataset.
annotation_df <- data.frame(xlab = c(12/79, 0.12), x1 = c(12/79, 4/79), ylab=c(30,50), y1 =
c(25,40), y2=c(0,0), lab=c("12/79","4/79"),
RNA_Dataset=c("NAMA1_vsWT_12DAANF","NAMA1_vsWT_22DAA"))
plot_homoeolog <- ggplot(output_12_22, aes(x=Ratio)) + geom_density(fill="darkorchid3",
alpha=0.7) + xlim(0,1) + ylab("Density") +   xlab("Ratio") + theme(axis.title =
element_text(size=rel(1))) + facet_wrap(~RNA_Dataset, ncol=1,
labeller=labeller(RNA_Dataset=plot.labels), scales="free") + theme_minimal() +
geom_text(data=annotation_df, aes(x = xlab, y =ylab, label=lab), size=3) +
geom_segment(data=annotation_df, aes(x=x1, xend=x1, y=y1, yend=y2), colour="black")
plot_homoeolog
svg(filename = "Z://Sophie/Manuscripts/Genie3/DraftFigures/PearceRandomDist.svg",width=2,
height=4)
plot_homoeolog
dev.off()
```

Do the sign test for stats:

```{r sign test}
print("Shared genes with NAM-A1")
print("12DAA nama1 vs WT")
output_rel %>% filter(RNA_Dataset %in% c("NAMA1_vsWT_12DAANF")) %>% SIGN.test(.$Ratio,
md=(12/79), alternative="less", conf.level = 0.95, data=.)
print("22DAA nama1 vs WT")
output_rel %>% filter(RNA_Dataset %in% c("NAMA1_vsWT_22DAA")) %>% SIGN.test(.$Ratio,
md=(4/79), alternative="less", conf.level = 0.95, data=.)
```

Can see that all of the NAM-A1 examples are significantly higher than the predicted median of the distribution for the relative comparisons.
What are the other genes that have high similarity with the Genie3 network?

```{r overlap}
##get the genes that have high ratios in common
```

```
NAMA1_12DAA_higherratio <- output_rel %>% filter(RNA_Dataset %in% c("NAMA1_vsWT_12DAANF"))
%>% filter(Ratio > (12/79))
NAMA1_22DAA_higherratio <- output_rel %>% filter(RNA_Dataset %in% c("NAMA1_vsWT_22DAA"))
%>% filter(Ratio > (4/79))
higher_ratios <- rbind(NAMA1_12DAA_higherratio, NAMA1_22DAA_higherratio)
unique_genes <- unique(as.character(higher_ratios$Genie3_Gene)) ##20 genes in total
as.data.frame(higher_ratios)
write.csv(as.data.frame(unique_genes),"Y://Sophie/Networks/Genie3_NAMcomparison/HigherRatio
Genie3TFs.csv")
write.csv(higher_ratios,"Y://Sophie/Networks/Genie3_NAMcomparison/HigherRatioGenie3TFs_full
table.csv")
```

8.3 Additional Data

8.3.1 The 59 candidate genes within the 4.3 Mb YES-1 locus.

Annotated high-confidence genes located between markers SH044 and SH59985, as depicted in Figure 2-18 and Figure 2-19, are presented in Table 8-1.

8.3.2 Targets from the GENIE3 network shared between the NAC3-1 homoeologs.

Shared targets of the NAC3-1 triad (*NAC3-A1*, *NAC3-B1*, *NAC3-D1*) as predicted by the wheat GENIE3 network and depicted in Figure 5-25 are presented in Table 8-2.

Table 8-1: 59 High-confidence genes fall within the *YES-1* locus. Genes are listed in their physical order across the region encompassed by markers SH044 and SH59985. Expression data from the developmental timecourse of cv. Chinese Spring and cv. Azhurnaya (Ramírez-González et al. 2018) were used to identify genes expressed over 0.5 transcripts per million (TPM) in leaf and shoot tissue (column A). The number of developmental stages in which the gene is expressed above 0.5 TPM in the leaf/shoot is listed in column B; see Figure 2-19. Whether the gene is expressed in the leaves and shoots during both the vegetative and reproductive stages is listed in column C. The remaining columns list the associated *Arabidopsis thaliana* and *Oryza sativa* var. *japonica* orthologues, alongside their functions. Orthologues were obtained using Ensembl BioMart (Kersey et al. 2017). Associated references for ortholog functions are available in Harrington et al. (2019b).

Physical Order	**Gene**	**A**	**B**	**C**	***Arabidopsis thaliana* gene stable ID**	***Arabidopsis thaliana* gene function**	***Oryza sativa* var. *japonica* orthologue**	***Oryza sativa* var. *japonica* function**
1	TraesCS3A02G410700	y	2	n	-	-	Os08g0195900	F-box protein 418
2	TraesCS3A02G410800	n	0	N/A	AT4G28680, AT2G20340	TYRDC; Probable tyrosine decarboxylase 2; ELI5; Tyrosine decarboxylase 1; Catalyse H+ and L-tyrosine to CO2 and tyramine.	Os07g0437500	TRYPTOPHAN DECARBOXYLASE 2; overexpression leads to higher serotonin levels and delayed leaf senescence.
3	TraesCS3A02G410900	n	0	N/A	-	-	Os02g0248100	-
4	TraesCS3A02G411000	n	0	N/A	-	-	Os08g0195900	F-box protein 418
5	TraesCS3A02G411200	n	0	N/A	-	-	Os02g0248100	-
6	TraesCS3A02G411300	n	0	N/A	-	-	Os04g0193300, Os04g0194000, Os04g0191800, Os04g0191600	F-box protein 183, F-box protein 184, F-box protein 182
7	TraesCS3A02G411400	n	0	N/A	-	-	-	-
8	TraesCS3A02G411500	n	0	N/A	-	-	Os06g0329900	-
9	TraesCS3A02G411600	y	1	N/A	-	-	Os06g0329900	-
10	TraesCS3A02G411700	y	3	y	AT5G48430	-	Os01g0937100	-
11	TraesCS3A02G411900	y	2	n	AT5G48430	-	Os01g0937800, Os01g0937600, Os01g0937500	-

Physical Order	Gene	A	B	C	*Arabidopsis thaliana* gene stable ID	*Arabidopsis thaliana* gene function	*Oryza sativa* var. *japonica* orthologue	*Oryza sativa* var. *japonica* function
12	TraesCS3A02G412000	n	0	N/A	-	-	Os03g0821550	-
13	TraesCS3A02G412100	n	0	N/A	-	-	Os06g0329900	-
14	TraesCS3A02G412200	n	0	N/A	-	-	Os06g0329900	-
15	TraesCS3A02G412300	n	0	N/A	-	-	Os04g0193300, Os04g0194000, Os04g0191800, Os04g0191600	F-box protein 183, F-box protein 184, F-box protein 182
16	TraesCS3A02G412400	n	0	N/A	-	-	-	-
17	TraesCS3A02G412500	n	0	N/A	AT4G08780, AT3G49110, AT2G38380, AT5G06730, AT5G19880, AT4G08770, AT5G06720, AT3G49120, AT2G38390, AT3G32980, AT3G32980	PER38, PER33, PER22, PER54, PER58, PER37, PER53, PER34, PER23, PER32. Peroxidases, with possible functions in removal of hydrogen peroxide, responses to various stresses, etc.	Os10g0107000	-
18	TraesCS3A02G412600	n	0	N/A	AT3G07180	T1B9.16, GPI transamidase component PIG-S-like protein; predicted role in attaching GPI anchor to protein	Os01g0907300	-

Physical Order	Gene	A	B	C	*Arabidopsis thaliana* gene stable ID	*Arabidopsis thaliana* gene function	*Oryza sativa* var. *japonica* orthologue	*Oryza sativa* var. *japonica* function
19	TraesCS3A02G412700	n	0	N/A	AT3G48430	Lysine-specific demethylase REF6; demethylases the Lys-27 of histone H3, involved in regulation of 100s of genes involved in developmental patterning and stimulus response. Represses FLC xpression and regulates SOC1.	Os01g0907400	Jumonji JMJ705, biotic stress-responsive H3K27me2/3 demethylase acts in response to methyl-jasmonate for defense-related gene activation.
20	TraesCS3A02G412800	y	3	y	AT4G34420, AT3G50620, AT2G15730	P-loop containing nucleoside triphosphate hydrolases superfamily protein; Sulfotransferase	Os03g0394900	-
21	TraesCS3A02G412900	n	0	N/A	AT5G50260, AT3G48350, AT3G48340	KDEL-tailed cysteine endopeptidase CEP1, CEP2, CEP3. Invovled in the final stages of programmed cell death via extensin cleavage i.e. in biotrophic pathogen responses and tapetal PCD.	Os01g0907600	REP-1/OsEP3A/OsSAG12-1; senescence associated gene 12-1, induced during senescence and pathogen-induced cell death, downregulation enhances senescence. Also has a role in digestion of seed storage proteins to release nutrients to seedlings, induced by GA and repressed by ABA.
22	TraesCS3A02G413000	n	0	N/A	AT5G50260, AT3G48350, AT3G48340	KDEL-tailed cysteine endopeptidase CEP1, CEP2, CEP3. Invovled in the final stages of programmed cell death via extensin cleavage i.e. in biotrophic pathogen responses and tapetal PCD.	Os01g0907600	REP-1/OsEP3A/OsSAG12-1; senescence associated gene 12-1, induced during senescence and pathogen-induced cell death, downregulation enhances senescence. Also has a role in digestion of seed storage proteins to release nutrients to seedlings, induced by GA and repressed by ABA.

Physical Order	Gene	A	B	C	*Arabidopsis thaliana* gene stable ID	*Arabidopsis thaliana* gene function	*Oryza sativa* var. *japonica* orthologue	*Oryza sativa* var. *japonica* function
23	TraesCS3A02G413100	n	0	N/A	AT5G50260, AT3G48350, AT3G48340	KDEL-tailed cysteine endopeptidase CEP1, CEP2, CEP3. Invovled in the final stages of programmed cell death via extensin cleavage i.e. in biotrophic pathogen responses and tapetal PCD.	Os01g0907600	REP-1/OsEP3A/OsSAG12-1; senescence associated gene 12-1, induced during senescence and pathogen-induced cell death, downregulation enhances senescence. Also has a role in digestion of seed storage proteins to release nutrients to seedlings, induced by GA and repressed by ABA.
24	TraesCS3A02G413200	n	0	N/A	AT5G50260, AT3G48350, AT3G48340	KDEL-tailed cysteine endopeptidase CEP1, CEP2, CEP3. Invovled in the final stages of programmed cell death via extensin cleavage i.e. in biotrophic pathogen responses and tapetal PCD.	Os01g0907600	REP-1/OsEP3A/OsSAG12-1; senescence associated gene 12-1, induced during senescence and pathogen-induced cell death, downregulation enhances senescence. Also has a role in digestion of seed storage proteins to release nutrients to seedlings, induced by GA and repressed by ABA.
25	TraesCS3A02G413300	y	3	y	AT1G04830, AT4G13730	Ypt/Rab-GAP domain of gyp1p superfamily predicted protein	Os01g0908100	-
26	TraesCS3A02G413400	n	0	N/A	AT3G26120, AT1G67770	TEL1, TEL2; predicted role in mRNA and RNA binding	-	-
27	TraesCS3A02G413500	y	3	y	-	-	-	-
28	TraesCS3A02G413600	y	3	y	AT3G48360, AT5G63160	BT2, BT1. BTB/POZ and TAZ domain-containing protein 1 and 2. May be substrate-specific adaptor of the E3 ubiquitin protein ligase complex, with role in telomerase activation pathway and auxin responses, among others.	Os01g0908200	OsBT; mutation increases nitrogren use efficiency by 20% under low nitrogen conditions.

Physical Order	Gene	A	B	C	*Arabidopsis thaliana* gene stable ID	*Arabidopsis thaliana* gene function	*Oryza sativa* var. *japonica* orthologue	*Oryza sativa* var. *japonica* function
29	TraesCS3A02G413700	y	3	y	AT3G48360, AT5G63160	BT2, BT1. BTB/POZ and TAZ domain-containing protein 1 and 2. May be substrate-specific adaptor of the E3 ubiquitin protein ligase complex, with role in telomerase activation pathway and auxin responses, among others.	Os01g0908200	OsBT; mutation increases nitrogren use efficiency by 20% under low nitrogen conditions.
30	TraesCS3A02G413800	y	3	y	-	-	Os01g0908400	-
31	TraesCS3A02G413900	n	0	N/A	AT5G15870, AT1G18310	Glycosyl hydrolase family 81 proteins	Os09g0379900	-
32	TraesCS3A02G414000	y	3	y	AT3G19640	MRS2-3; magnesium transporter that can complement yeast mrs2 mutants	Os01g0908500	OsMRS2/MGT family member 3, part of the CorA-MRS2-ALR-Type Magnesium transporter family
33	TraesCS3A02G414100	n	0	N/A	AT1G72650, AT1G17460	TRFL6, TRFL3	Os02g0776700	CTMyb1
34	TraesCS3A02G414200	y	2	n	AT5G40600, AT3G27420	Bromodomain testis-specific protein	-	-
35	TraesCS3A02G414300	y	3	y	-	-	Os01g0908600	proline transporter 1
36	TraesCS3A02G414400	y	3	y	AT1G67800, AT5G14420, AT3G01650	RGLG2, RGLG1, E3 ubiquitin protein ligases. Both have roles# in regulating the abundance of auxin transport proteins, affecting apical dominance, and acts as a negative regulator of drought stress responses.	Os01g0908700	RING finger protein
37	TraesCS3A02G414500	y	1	N/A	-	-	-	-

Physical Order	Gene	A	B	C	*Arabidopsis thaliana* gene stable ID	*Arabidopsis thaliana* gene function	*Oryza sativa* var. *japonica* orthologue	*Oryza sativa* var. *japonica* function
38	TraesCS3A02G414600	n	0	N/A	AT4G08780, AT3G49110, AT2G38380, AT5G06730, AT5G19880, AT4G08770, AT5G06720, AT3G49120, AT2G38390, AT3G32980, AT3G32980	PER38, PER33, PER22, PER54, PER58, PER37, PER53, PER34, PER23, PER32. Peroxidases, with possible functions in removal of hydrogen peroxide, responses to various stresses, etc.	Os10g0107000	-
39	TraesCS3A02G414700	y	2	n	AT3G07530	Integrator complex subunit	-	-
40	TraesCS3A02G414800	n	0	N/A	AT2G34930	Disease resistance family protein / LRR family protein; predicted role in fungus and oomycete defense responses	-	-
41	TraesCS3A02G414900	y	1	N/A	-	-	-	-
42	TraesCS3A02G415000	n	0	N/A	AT2G34930	Disease resistance family protein / LRR family protein; predicted role in fungus and oomycete defense responses	-	-
43	TraesCS3A02G415100	n	0	N/A	-	-	-	-
44	TraesCS3A02G415200	n	0	N/A	-	-	Os01g0909500	C2H2 zinc finger protein
45	TraesCS3A02G415300	y	3	y	AT1G61065, AT3G15480, AT1G52910, AT4G27435	1,3-beta-glucan synthase component (DUF1218)	Os07g0462200	-

Physical Order	Gene	A	B	C	*Arabidopsis thaliana* gene stable ID	*Arabidopsis thaliana* gene function	*Oryza sativa* var. *japonica* orthologue	*Oryza sativa* var. *japonica* function
46	TraesCS3A02G415400	y	3	y	AT3G43920	DCL3, Endoribonuclease Dicer homolog 3. Role in post-transcriptional gene silencing. Promotes flowering via repression of FLC redundantly with DCL1.	Os01g0909200	OsDCL3a, DICER-LIKE 3A targets genes involved in giberellin and brassinosteroid homoeostasis, with deficiency leading to dwarfism and larger flag leaf angle.
47	TraesCS3A02G415500	y	2	y	-	-	-	-
48	TraesCS3A02G415600	n	0	N/A	-	-	-	-
49	TraesCS3A02G415700	n	0	N/A	AT3G16920, AT1G05850	CTL2, CTL1; chitinase-like proteins. CTL1 has essential role in plant growth and development, especially in cell expansion, root architecture, stress tolerance, etc.	Os08g0522500	-
50	TraesCS3A02G415800	n	0	N/A	-	-	Os01g0909300	-
51	TraesCS3A02G415900	y	3	y	-	-	Os01g0909400	-
52	TraesCS3A02G416000	n	0	N/A	-	-	-	-
53	TraesCS3A02G416100	y	3	y	AT1G08040, AT2G28310	Putative storage protein; Trimethylguanosine synthase (DUF707)	-	-
54	TraesCS3A02G416200	y	3	y	-	-	-	-
55	TraesCS3A02G416400	y	3	y	-	-	-	-
56	TraesCS3A02G416300	y	3	y	-	-	-	-
57	TraesCS3A02G416500	n	0	N/A	AT1G55750, AT3G61420	TFB1-1, TFB1-3; components of the transcription and DNA repair factor IIH core complex.	Os08g0338200	-
58	TraesCS3A02G416600	n	0	N/A	-	-	-	-
59	TraesCS3A02G416700	n	0	N/A	-	-	Os01g0965700	-

Table 8-2: Targets predicted by the wheat GENIE3 network which are shared in the NAC3 triad. Homoeologs are highlighted for each chromosome (different colours reflect different triads on each chromosome). The fold-change in senescing tissues and the calculation of a 2-fold increase in expression during senescence were calculated based on the Bayer developmental timecourse data (Ramírez-González et al. 2018). The timing of gene upregulation during senescence is reported for those genes included in the analysis of Borrill et al. (2019a), where DAA refers to days after anthesis. Arabidopsis orthologues were obtained from EnsemblPlants (Kersey et al. 2017), with their associated descriptions modified from the TAIR database (Berardini et al. 2015). The expression of the Arabidopsis orthologue during senescence was based on the Klepikova Arabidopsis Atlas (Klepikova et al. 2016), where genes were considered to have expression in senescence when the absolute TPM value in the senescent leaf was at least 75% of the maximum TPM.

Gene ID (v 1.1)	TF Family	Fold change in senescence	Upregulated 2x in senescence	Timing of upregulation	Arabidopsis Orthologue	Arabidopsis Orthologue Name	Arabidopsis Function	Arabidopsis Senescence Expression?
TraesCS1A02G145800		4.08	YES	UP 17 DAA	AT1G07040	NA	Chloroplast-located	YES
TraesCS1B02G364200		2.76	YES	UP 15 DAA	AT2G42750	NA	DNAJ heat shock N-terminal domain-containing protein	YES
TraesCS1D02G144400		4.16	YES	UP 17 DAA	AT1G07040	NA	Chloroplast-located	YES
TraesCS1D02G392300		2.16	YES	UP 15 DAA	AT4G13400	NA	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	YES
TraesCS1D02G403300		4.51	YES	NA	NA	NA	NA	NA
TraesCS2A02G102000	NAC	1.29	NO	UP 19 DAA	AT3G04070	NAC047	NAC domain containing protein 47; AKA SPEEDY HYPNOTASTIC GROWTH	YES
TraesCS2A02G505500		5.43	YES	UP 15 DAA	AT2G32270	ZIP3	A member of Zrt- and Irt-related protein (ZIP) family. Transcript is induced in response to zinc deficiency in the root and also in response to iron deficiency.	NO

Gene ID (v 1.1)	TF Family	Fold change in senescence	Upregulated 2x in senescence	Timing of upregulation	Arabidopsis Orthologue	Arabidopsis Orthologue Name	Arabidopsis Function	Arabidopsis Senescence Expression?
TraesCS2B02G048700		1.67	NO	UP 19 DAA	AT4G08870	ARGAH2	Encodes one of the two arginase in the genome. Gene expression is enhanced by methyl jasmonate treatment. It is involved in the defense response to B.cinerea.	NO
TraesCS2B02G118400	NAC	3.42	YES	UP 17 DAA	AT3G04060	anac046	NAC046 is a member of the NAC domain containing family of transcription factors. It was identified in a screen for regulators of chlorophyll protein gene expression. Mutants in NAC046 have delayed senescence and increased CHL content suggesting a role in regulation of senescence and chlorophyll degradation.	YES
TraesCS2B02G363800		1.27	NO	UP 15 DAA	AT1G67600	NA	Acid phosphatase/vanadium-dependent haloperoxidase-related protein	NO
TraesCS2B02G491000		0.08	NO	NA	NA	NA	NA	NA
TraesCS2D02G034900		1.82	NO	UP 17 DAA	AT4G08870	ARGAH2	Encodes one of the two arginase in the genome. Gene expression is enhanced by methyl jasmonate treatment. It is involved in the defense response to B.cinerea.	NO

Gene ID (v 1.1)	TF Family	Fold change in senescence	Upregulated 2x in senescence	Timing of upregulation	Arabidopsis Orthologue	Arabidopsis Orthologue Name	Arabidopsis Function	Arabidopsis Senescence Expression?
TraesCS2D02G163300		2.37	YES	UP 19 DAA	NA	NA	NA	NA
TraesCS2D02G407100		2.54	YES	UP 13 DAA	NA	NA	NA	NA
TraesCS2D02G527800		2.28	YES	UP 17 DAA	AT3G29636	NA	Transferase-like protein	NO
TraesCS3A02G078400	NAC	7.31	YES	UP 17 DAA	AT3G04070	NAC047	NAC domain containing protein 47; AKA SPEEDY HYPNOTASTIC GROWTH	YES
TraesCS3A02G283800		1.10	NO	NA	NA	NA	NA	NA
TraesCS3A02G406000	NAC	1.10	NO	UP 17 DAA	AT5G63790	ANAC102	Encodes a member of the NAC family of transcription factors. ANAC102 appears to have a role in mediating response to low oxygen stress (hypoxia) in germinating seedlings.	YES
TraesCS3B02G203700		5.27	YES	UP 13 DAA	AT1G02660	PLIP2	PLIP2 is a glycerolipid A1 lipase with substrate preference for monogalactosyldiacylglycerol. Expression is induced by ABA	YES
TraesCS3B02G439600	NAC	1.30	NO	UP 13 DAA	AT5G63790	ANAC102	Encodes a member of the NAC family of transcription factors. ANAC102 appears to have a role in mediating response to low oxygen stress (hypoxia) in germinating seedlings.	YES

Gene ID (v 1.1)	TF Family	Fold change in senescence	Upregulated 2x in senescence	Timing of upregulation	Arabidopsis Orthologue	Arabidopsis Orthologue Name	Arabidopsis Function	Arabidopsis Senescence Expression?
TraesCS3D02G078900	NAC	6.85	YES	UP 17 DAA	AT3G04070	NAC047	NAC domain containing protein 47; AKA SPEEDY HYPNOTASTIC GROWTH	YES
TraesCS3D02G283700		1.46	NO	NA	NA	NA	NA	NA
TraesCS3D02G401200	NAC	1.15	NO	UP 13 DAA	AT5G63790	ANAC102	Encodes a member of the NAC family of transcription factors. ANAC102 appears to have a role in mediating response to low oxygen stress (hypoxia) in germinating seedlings.	YES
TraesCS4A02G411000		4.41	YES	UP 17 DAA	AT3G44880	PAO/ACD1	Encodes a pheide a oxygenase (PAO). Accelerated cell death (acd1) mutants show rapid, spreading necrotic responses to both virulent and avirulent Pseudomonas syringae pv. maculicola or pv. tomato pathogens and to ethylene.	YES
TraesCS4B02G052200		2.60	YES	UP 15 DAA	NA	NA	NA	NA
TraesCS5A02G206600		5.91	YES	UP 17 DAA	AT3G25860	LTA2	Encodes the dihydrolipoamide S-acetyltransferase subunit of the plastid Pyruvate Dehydrogenase Complex (E2). Mutant has embryo defect.	YES
TraesCS5A02G318900		1.35	NO	NA	NA	NA	NA	NA

Gene ID (v 1.1)	TF Family	Fold change in senescence	Upregulated 2x in senescence	Timing of upregulation	Arabidopsis Orthologue	Arabidopsis Orthologue Name	Arabidopsis Function	Arabidopsis Senescence Expression?
TraesCS5A02G319900		2.97	YES	UP 19 DAA	AT4G22920	SGR1	Similar to the tomato senescence-inducible chloroplast stay-green protein 1. It is upregulated during maximal senescence in the Arabidopsis life cycle, especially in senescent leaves. Acts antagonistically with SGR2 to balance chlorophyll catabolism in chloroplasts with the dismantling and remobilizing of other cellular components in senescing leaf cells	YES
TraesCS5A02G375000		2.34	YES	UP 19 DAA	AT5G46180	DELTA-OAT	Ornithine delta-aminotransferase with role in arginine and ornithine catabolism	YES
TraesCS5A02G463700		3.46	YES	UP 17 DAA	AT5G65550	UGT91B1	UDP-Glycosyltransferase superfamily protein;(s	NO
TraesCS5B02G210100		1.03	NO	UP 19 DAA	AT1G54100	ALDH7B4	Aldehyde dehydrogenase	YES
TraesCS5B02G228800		4.80	YES	UP 17 DAA	NA	NA	NA	NA
TraesCS5B02G321800		4.26	YES	UP 19 DAA	AT5G51070	CLPD	ATP-dependent Clp protease regulatory subunit; AKA Senescence Associated Gene 15 (SAG15)	YES
TraesCS5B02G368300		NA	NA	NA	NA	NA	NA	NA

Gene ID (v 1.1)	TF Family	Fold change in senescence	Upregulated 2x in senescence	Timing of upregulation	Arabidopsis Orthologue	Arabidopsis Orthologue Name	Arabidopsis Function	Arabidopsis Senescence Expression?
TraesCS5B02G376900		1.67	NO	UP 13 DAA	AT5G46180	DELTA-OAT	Ornithine delta-aminotransferase with role in arginine and ornithine catabolism	YES
TraesCS5D02G325900		2.51	YES	UP 19 DAA	AT4G22920	SGR1	Similar to the tomato senescence-inducible chloroplast stay-green protein 1. It is upregulated during maximal senescence in the Arabidopsis life cycle, especially in senescent leaves. Acts antagonistically with SGR2 to balance chlorophyll catabolism in chloroplasts with the dismantling and remobilizing of other cellular components in senescing leaf cells	YES
TraesCS6A02G141500		1.78	NO	UP 17 DAA	AT1G06570	HPD/PDS1	Mutation of the PDS1 locus disrupts the activity of p-hydroxyphenylpyruvate dioxygenase (HPPDase), the first committed step in the synthesis of both plastoquinone and tocopherols in plants.	YES

Gene ID (v 1.1)	TF Family	Fold change in senescence	Upregulated 2x in senescence	Timing of upregulation	Arabidopsis Orthologue	Arabidopsis Orthologue Name	Arabidopsis Function	Arabidopsis Senescence Expression?
TraesCS6A02G281000		2.63	YES	UP 17 DAA	AT5G48020	NA	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	YES
TraesCS6B02G169700		1.37	NO	UP 15 DAA	AT1G06570	HPD/PDS1	Mutation of the PDS1 locus disrupts the activity of p-hydroxyphenylpyruvate dioxygenase (HPPDase), the first committed step in the synthesis of both plastoquinone and tocopherols in plants.	YES
TraesCS6B02G432600		2.13	YES	UP 17 DAA	NA	NA	NA	NA
TraesCS6D02G063100		0.87	NO	NA	NA	NA	NA	NA
TraesCS7A02G127800		3.48	YES	NA	AT5G67140	NA	F-box/RNI-like superfamily protein	YES
TraesCS7A02G338900		6.96	YES	UP 19 DAA	AT5G13800	PPH	Pheophytinase that is involved in chlorophyll breakdown	YES
TraesCS7A02G537100		4.28	YES	NA	NA	NA	NA	NA
TraesCS7B02G448000		0.00	NO	NA	NA	NA	NA	NA
TraesCS7D02G518200		4.39	YES	NA	NA	NA	NA	NA
TraesCS7D02G518300		9.56	YES	NA	AT1G56240	PP2B13	phloem protein 2-B13	YES
TraesCS7D02G518400		28.29	YES	NA	AT1G56240	PP2B13	phloem protein 2-B13	YES
TraesCS7D02G524400		4.01	YES	NA	NA	NA	NA	NA
TraesCSU02G048500		24.23	YES	UP 3 DAA	NA	NA	NA	NA

8.3.3 Repeated phenotyping of the BC_1F_3 nac3-1 mutant lines

Two repeats were carried out of phenotyping the BC_1F_3 plants with mutations in the *NAC3-A1* and *NAC3-B1* genes, as described in Chapter 5. Each repeat was carried out under glasshouse conditions, though with different sowing dates (December 2018 for Figure 8-1 and February 2019 for Figure 8-2). The seedlings for each line were derived from the same BC_1F_2 parent for each experiment. Statistics for all graphs were carried out using the two-sample Wilcox test.

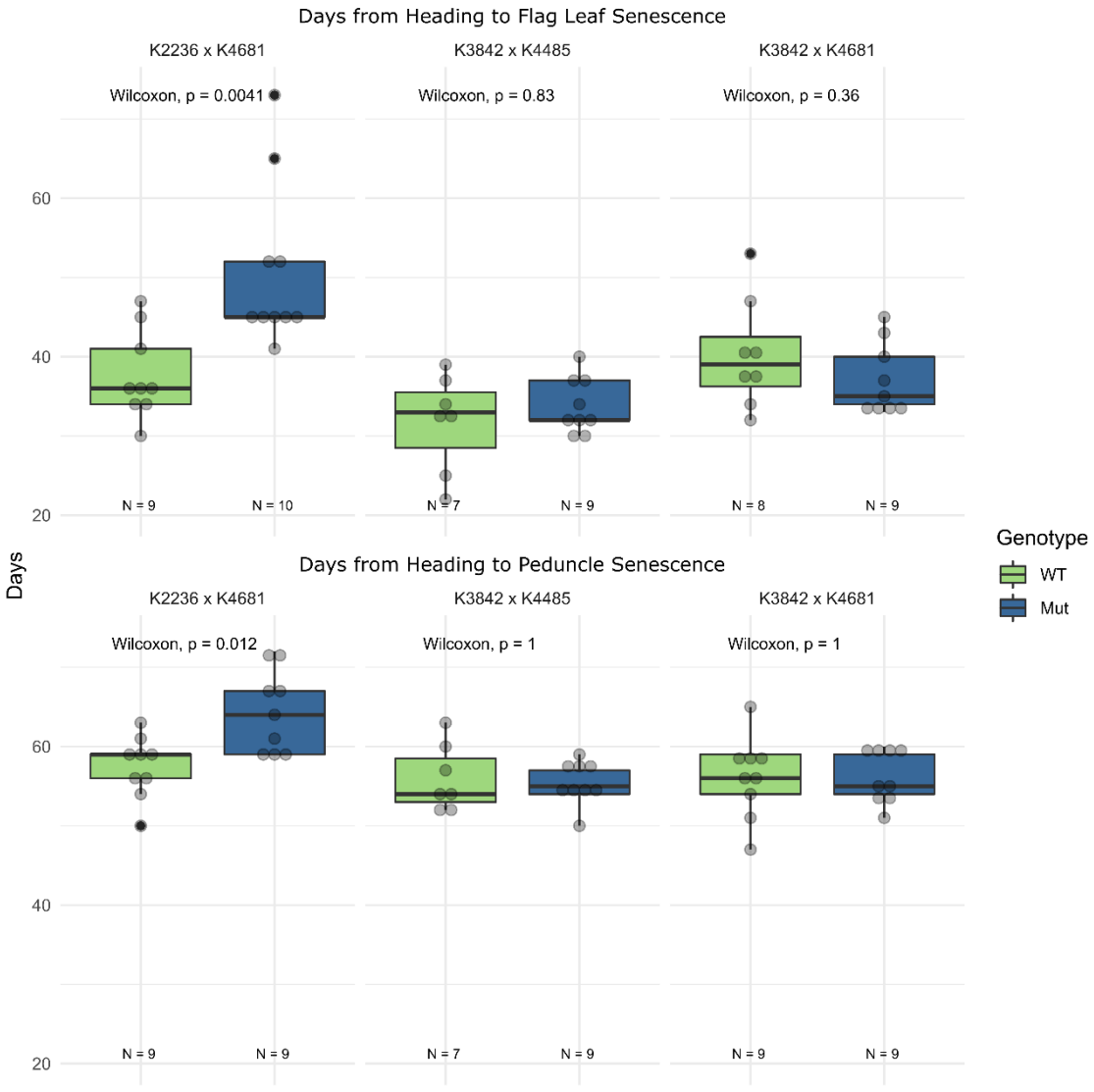

Figure 8-1: December 2018 repeat of the *nac3-1* BC_1F_3 mutant lines. Top panel depicts flag leaf senescence (shown as days from heading to onset of senescence); bottom panel depicts peduncle senescence (shown as days from heading to complete peduncle senescence).

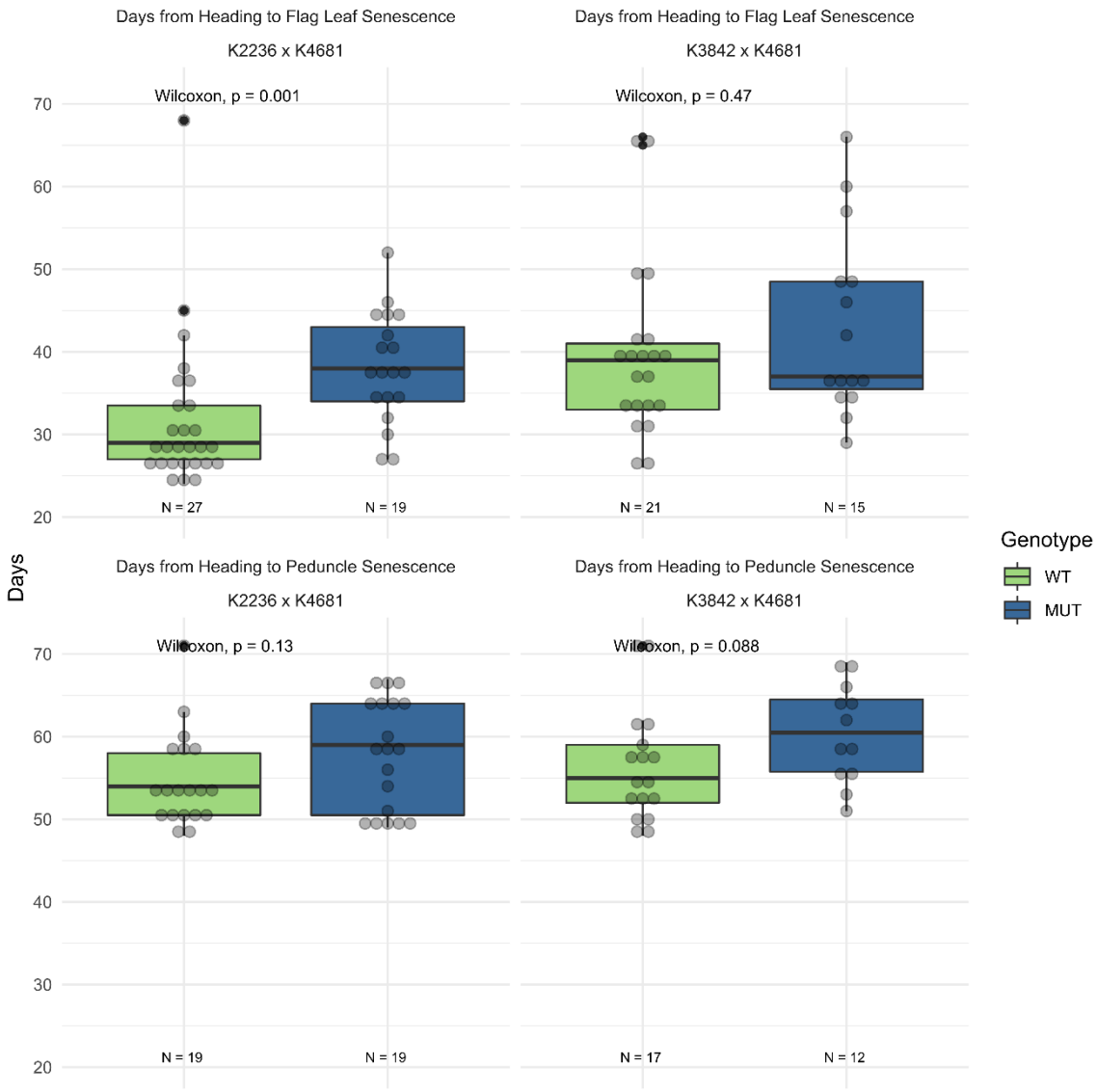

Figure 8-2: February 2018 repeat of the *nac3-1* BC_1F_3 mutant lines. Top panel depicts flag leaf senescence (shown as days from heading to onset of senescence); bottom panel depicts peduncle senescence (shown as days from heading to complete peduncle senescence). N varies between flag leaf and peduncle senescence as individual plants were destructively sampled in between these two timepoints, reducing sample size.

8.3.4 List of targets of NAC3-A1, NAC3-B1, and NAC3-D1 from the GENIE3 network.

GO term enrichment of the predicted targets of the NAC3-1 triad was presented in the following tables. The top 100 targets of each homoeolog are listed; the full set can be extracted from the GENIE3 network available at https://doi.org/10.5447/ipk/2018/7, using the scripts available at https://github.com/Uauy-Lab/GENIE3_scripts. The full list of targets for the three transcription factors are provided in the following tables, where "from.gene" is the transcription factor, "to.gene" is the predicted target, and "im" is the association weight. Larger "im" values correspond to better supported links.

8.3.4.1 NAC3-A1

from.gene	to.gene	im
TraesCS2A01G101400	TraesCS2B01G118500	0.028254
TraesCS2A01G101400	TraesCS2B01G548200	0.025726
TraesCS2A01G101400	TraesCS5B01G094000	0.024513
TraesCS2A01G101400	TraesCS2D01G420100	0.024195
TraesCS2A01G101400	TraesCS2D01G100900	0.024137
TraesCS2A01G101400	TraesCS5B01G524100	0.022925
TraesCS2A01G101400	TraesCS2D01G086900	0.022499
TraesCS2A01G101400	TraesCS7A01G445100	0.022359
TraesCS2A01G101400	TraesCS5B01G460300	0.022
TraesCS2A01G101400	TraesCS4D01G127900	0.021939
TraesCS2A01G101400	TraesCS3D01G383800	0.021769
TraesCS2A01G101400	TraesCS7B01G292800	0.021502
TraesCS2A01G101400	TraesCS3D01G352200	0.0213
TraesCS2A01G101400	TraesCS3B01G094200	0.02121
TraesCS2A01G101400	TraesCS3A01G454100	0.021107
TraesCS2A01G101400	TraesCS7B01G489600	0.021027
TraesCS2A01G101400	TraesCS3A01G469300	0.020411
TraesCS2A01G101400	TraesCS2A01G473500	0.020363
TraesCS2A01G101400	TraesCS1D01G206400	0.020262
TraesCS2A01G101400	TraesCS7A01G531900	0.020191
TraesCS2A01G101400	TraesCS7D01G539300	0.020104
TraesCS2A01G101400	TraesCS5A01G022400	0.020089
TraesCS2A01G101400	TraesCS7A01G523500	0.019868
TraesCS2A01G101400	TraesCS2D01G096400	0.019704
TraesCS2A01G101400	TraesCS6B01G134900	0.019349
TraesCS2A01G101400	TraesCS3D01G089300	0.019327
TraesCS2A01G101400	TraesCS5B01G264900	0.0193
TraesCS2A01G101400	TraesCS4A01G383200	0.019272
TraesCS2A01G101400	TraesCS7D01G447000	0.019263
TraesCS2A01G101400	TraesCS5D01G377500	0.019127
TraesCS2A01G101400	TraesCSU01G106000	0.019036
TraesCS2A01G101400	TraesCS7D01G042700	0.019011
TraesCS2A01G101400	TraesCS5D01G267900	0.018827
TraesCS2A01G101400	TraesCS1D01G373500	0.018812
TraesCS2A01G101400	TraesCS2B01G540600	0.018793
TraesCS2A01G101400	TraesCS5B01G093900	0.01879
TraesCS2A01G101400	TraesCS2B01G393500	0.018774
TraesCS2A01G101400	TraesCS7D01G144500	0.018771
TraesCS2A01G101400	TraesCS1B01G473000	0.018769
TraesCS2A01G101400	TraesCS7A01G428000	0.018744
TraesCS2A01G101400	TraesCS5B01G438600	0.018596
TraesCS2A01G101400	TraesCS5D01G349500	0.01858
TraesCS2A01G101400	TraesCS1D01G447600	0.018473
TraesCS2A01G101400	TraesCS5D01G380400	0.018369
TraesCS2A01G101400	TraesCS3B01G511300	0.018348
TraesCS2A01G101400	TraesCS5B01G434300	0.018336

TraesCS2A01G101400	TraesCS2A01G366500	0.018319
TraesCS2A01G101400	TraesCS7D01G548700	0.018259
TraesCS2A01G101400	TraesCS6D01G160800	0.018175
TraesCS2A01G101400	TraesCS2B01G100100	0.018153
TraesCS2A01G101400	TraesCS7A01G127800	0.018137
TraesCS2A01G101400	TraesCS7A01G437700	0.018114
TraesCS2A01G101400	TraesCS6B01G064400	0.018046
TraesCS2A01G101400	TraesCS1A01G438200	0.018037
TraesCS2A01G101400	TraesCS2B01G527500	0.01801
TraesCS2A01G101400	TraesCS5B01G273900	0.017981
TraesCS2A01G101400	TraesCS3A01G134700	0.017951
TraesCS2A01G101400	TraesCS3A01G134800	0.017894
TraesCS2A01G101400	TraesCS5A01G255400	0.017852
TraesCS2A01G101400	TraesCS4B01G091300	0.017802
TraesCS2A01G101400	TraesCS4B01G284000	0.017788
TraesCS2A01G101400	TraesCS5D01G461700	0.017775
TraesCS2A01G101400	TraesCS6A01G093400	0.01773
TraesCS2A01G101400	TraesCS7D01G519100	0.017729
TraesCS2A01G101400	TraesCS2D01G560300	0.017725
TraesCS2A01G101400	TraesCS5B01G524500	0.017679
TraesCS2A01G101400	TraesCS5D01G281500	0.017589
TraesCS2A01G101400	TraesCS3A01G020000	0.017584
TraesCS2A01G101400	TraesCS2B01G118400	0.017555
TraesCS2A01G101400	TraesCS2B01G461900	0.017554
TraesCS2A01G101400	TraesCS3B01G172100	0.017535
TraesCS2A01G101400	TraesCS6D01G311800	0.017527
TraesCS2A01G101400	TraesCS5A01G378500	0.017489
TraesCS2A01G101400	TraesCS4B01G399100	0.017308
TraesCS2A01G101400	TraesCSU01G001000	0.017276
TraesCS2A01G101400	TraesCS1B01G238800	0.017228
TraesCS2A01G101400	TraesCS7A01G165400	0.017211
TraesCS2A01G101400	TraesCS7D01G548800	0.017171
TraesCS2A01G101400	TraesCS5B01G445800	0.017143
TraesCS2A01G101400	TraesCS6D01G056200	0.017084
TraesCS2A01G101400	TraesCS2B01G405000	0.017056
TraesCS2A01G101400	TraesCS5A01G467300	0.017006
TraesCS2A01G101400	TraesCS4B01G294400	0.016998
TraesCS2A01G101400	TraesCS2B01G114400	0.016988
TraesCS2A01G101400	TraesCS5B01G024000	0.016967
TraesCS2A01G101400	TraesCS7D01G112600	0.016952
TraesCS2A01G101400	TraesCS6B01G326400	0.016896
TraesCS2A01G101400	TraesCS7D01G395500	0.016869
TraesCS2A01G101400	TraesCS2B01G495900	0.016834
TraesCS2A01G101400	TraesCS7B01G252100	0.016823
TraesCS2A01G101400	TraesCS2A01G125800	0.016818
TraesCS2A01G101400	TraesCS1A01G367400	0.016799
TraesCS2A01G101400	TraesCS5A01G462800	0.01678
TraesCS2A01G101400	TraesCS3A01G476400	0.016765

TraesCS2A01G101400	TraesCS6A01G105500	0.016744
TraesCS2A01G101400	TraesCS3D01G509500	0.016708
TraesCS2A01G101400	TraesCS1A01G053200	0.016692
TraesCS2A01G101400	TraesCS7D01G423100	0.016627
TraesCS2A01G101400	TraesCS5B01G199800	0.016578
TraesCS2A01G101400	TraesCS2A01G558900	0.016546

8.3.4.2 NAC3-B1

from.gene	to.gene	im
TraesCS2B01G118500	TraesCS2A01G101400	0.027663
TraesCS2B01G118500	TraesCS2D01G100900	0.026335
TraesCS2B01G118500	TraesCS2A01G505500	0.022434
TraesCS2B01G118500	TraesCS2B01G118400	0.021974
TraesCS2B01G118500	TraesCS5A01G292000	0.020949
TraesCS2B01G118500	TraesCS3B01G074300	0.020259
TraesCS2B01G118500	TraesCS6D01G384300	0.019564
TraesCS2B01G118500	TraesCS2D01G403800	0.017934
TraesCS2B01G118500	TraesCS3B01G447500	0.017182
TraesCS2B01G118500	TraesCS5A01G375000	0.016948
TraesCS2B01G118500	TraesCS5B01G228800	0.016354
TraesCS2B01G118500	TraesCS5D01G436100	0.016255
TraesCS2B01G118500	TraesCS3D01G401200	0.016253
TraesCS2B01G118500	TraesCS3B01G439600	0.015875
TraesCS2B01G118500	TraesCS3D01G078900	0.015682
TraesCS2B01G118500	TraesCS4B01G096800	0.015596
TraesCS2B01G118500	TraesCS7D01G518300	0.015545
TraesCS2B01G118500	TraesCS4B01G339800	0.015385
TraesCS2B01G118500	TraesCS7D01G137200	0.015303
TraesCS2B01G118500	TraesCS5A01G206600	0.015277
TraesCS2B01G118500	TraesCS2B01G491000	0.014892
TraesCS2B01G118500	TraesCS5A01G230300	0.014846
TraesCS2B01G118500	TraesCS2B01G363800	0.01483
TraesCS2B01G118500	TraesCS2D01G507800	0.01475
TraesCS2B01G118500	TraesCS3A01G283800	0.014731
TraesCS2B01G118500	TraesCS2D01G034900	0.01446
TraesCS2B01G118500	TraesCS3D01G283700	0.014433
TraesCS2B01G118500	TraesCS5D01G325900	0.014293
TraesCS2B01G118500	TraesCS1B01G288700	0.014276
TraesCS2B01G118500	TraesCS4B01G072100	0.014217
TraesCS2B01G118500	TraesCS4D01G057400	0.014074
TraesCS2B01G118500	TraesCS6B01G169700	0.014021
TraesCS2B01G118500	TraesCS6A01G329400	0.013993
TraesCS2B01G118500	TraesCS3A01G406000	0.013986
TraesCS2B01G118500	TraesCS4A01G336400	0.013865
TraesCS2B01G118500	TraesCS5D01G261400	0.013798
TraesCS2B01G118500	TraesCS7D01G413300	0.013631
TraesCS2B01G118500	TraesCS3A01G217400	0.01361

TraesCS2B01G118500	TraesCS7D01G518400	0.013592
TraesCS2B01G118500	TraesCS3D01G407700	0.013575
TraesCS2B01G118500	TraesCS4B01G292900	0.013566
TraesCS2B01G118500	TraesCS1D01G403200	0.013558
TraesCS2B01G118500	TraesCS7D01G346100	0.013403
TraesCS2B01G118500	TraesCS5A01G428100	0.013402
TraesCS2B01G118500	TraesCS2A01G350400	0.013397
TraesCS2B01G118500	TraesCS2A01G406900	0.01334
TraesCS2B01G118500	TraesCS1D01G403300	0.013177
TraesCS2B01G118500	TraesCS5D01G042300	0.013172
TraesCS2B01G118500	TraesCS2D01G313200	0.013097
TraesCS2B01G118500	TraesCS6B01G362700	0.013091
TraesCS2B01G118500	TraesCS1A01G145800	0.013032
TraesCS2B01G118500	TraesCS7D01G524400	0.01277
TraesCS2B01G118500	TraesCS2A01G273100	0.012765
TraesCS2B01G118500	TraesCS5D01G369400	0.012753
TraesCS2B01G118500	TraesCS4A01G411000	0.012679
TraesCS2B01G118500	TraesCS6B01G172000	0.012649
TraesCS2B01G118500	TraesCS5B01G503600	0.012645
TraesCS2B01G118500	TraesCS2D01G527800	0.012579
TraesCS2B01G118500	TraesCS4D01G282500	0.012553
TraesCS2B01G118500	TraesCS5D01G257300	0.01243
TraesCS2B01G118500	TraesCS4B01G052200	0.0123
TraesCS2B01G118500	TraesCS5B01G321800	0.012242
TraesCS2B01G118500	TraesCS5B01G303900	0.012145
TraesCS2B01G118500	TraesCS1A01G175500	0.012138
TraesCS2B01G118500	TraesCS5B01G291100	0.01209
TraesCS2B01G118500	TraesCS5B01G247800	0.012081
TraesCS2B01G118500	TraesCS1B01G364200	0.011997
TraesCS2B01G118500	TraesCS4B01G329300	0.01194
TraesCS2B01G118500	TraesCS5D01G275500	0.011936
TraesCS2B01G118500	TraesCSU01G048500	0.011914
TraesCS2B01G118500	TraesCS5A01G319900	0.011873
TraesCS2B01G118500	TraesCS6A01G141500	0.011873
TraesCS2B01G118500	TraesCS3B01G356000	0.01183
TraesCS2B01G118500	TraesCS3D01G179200	0.01182
TraesCS2B01G118500	TraesCS3D01G154300	0.011656
TraesCS2B01G118500	TraesCS4D01G328900	0.011609
TraesCS2B01G118500	TraesCS7B01G440000	0.011579
TraesCS2B01G118500	TraesCS3B01G094500	0.011567
TraesCS2B01G118500	TraesCS6A01G252900	0.011484
TraesCS2B01G118500	TraesCS1D01G144400	0.01145
TraesCS2B01G118500	TraesCS7D01G519100	0.011438
TraesCS2B01G118500	TraesCS1A01G261800	0.011399
TraesCS2B01G118500	TraesCS2D01G100700	0.011285
TraesCS2B01G118500	TraesCS4A01G123800	0.01125
TraesCS2B01G118500	TraesCS2A01G507400	0.011238
TraesCS2B01G118500	TraesCS7B01G448000	0.011214

TraesCS2B01G118500	TraesCS4D01G176700	0.011183
TraesCS2B01G118500	TraesCS5B01G355800	0.011178
TraesCS2B01G118500	TraesCS1D01G392300	0.011175
TraesCS2B01G118500	TraesCS6A01G400100	0.011132
TraesCS2B01G118500	TraesCS6D01G063100	0.011003
TraesCS2B01G118500	TraesCS1B01G164000	0.010998
TraesCS2B01G118500	TraesCS5B01G320200	0.010957
TraesCS2B01G118500	TraesCS4B01G099300	0.010915
TraesCS2B01G118500	TraesCS1B01G145000	0.010886
TraesCS2B01G118500	TraesCS6B01G273000	0.010849
TraesCS2B01G118500	TraesCS7A01G127800	0.010837
TraesCS2B01G118500	TraesCS3B01G322100	0.010837
TraesCS2B01G118500	TraesCS7D01G518200	0.010822
TraesCS2B01G118500	TraesCS2B01G436200	0.010794

8.3.4.3 NAC3-D1

from.gene	to.gene	im
TraesCS2D01G100900	TraesCS2B01G118500	0.033049
TraesCS2D01G100900	TraesCS5A01G316000	0.023281
TraesCS2D01G100900	TraesCS2A01G101400	0.022862
TraesCS2D01G100900	TraesCS4B01G052200	0.015921
TraesCS2D01G100900	TraesCS6A01G141500	0.015758
TraesCS2D01G100900	TraesCS6D01G341100	0.015167
TraesCS2D01G100900	TraesCS6B01G432600	0.015067
TraesCS2D01G100900	TraesCS5D01G384500	0.014484
TraesCS2D01G100900	TraesCS3D01G283700	0.014318
TraesCS2D01G100900	TraesCS3A01G078400	0.014253
TraesCS2D01G100900	TraesCS2B01G118400	0.013539
TraesCS2D01G100900	TraesCS1B01G066500	0.01325
TraesCS2D01G100900	TraesCS5A01G318900	0.013177
TraesCS2D01G100900	TraesCSU01G176600	0.013165
TraesCS2D01G100900	TraesCS2D01G100700	0.012974
TraesCS2D01G100900	TraesCS2B01G220900	0.012806
TraesCS2D01G100900	TraesCS2B01G491000	0.012687
TraesCS2D01G100900	TraesCS1A01G311700	0.012517
TraesCS2D01G100900	TraesCS6B01G169700	0.012386
TraesCS2D01G100900	TraesCS5B01G376900	0.012347
TraesCS2D01G100900	TraesCS5D01G325900	0.012317
TraesCS2D01G100900	TraesCS5A01G321300	0.012271
TraesCS2D01G100900	TraesCS3B01G439600	0.012114
TraesCS2D01G100900	TraesCS1A01G145800	0.012098
TraesCS2D01G100900	TraesCS3B01G609600	0.011964
TraesCS2D01G100900	TraesCS2A01G505500	0.011963
TraesCS2D01G100900	TraesCS6D01G063100	0.011872
TraesCS2D01G100900	TraesCS2B01G363800	0.011832
TraesCS2D01G100900	TraesCS3B01G093300	0.011493
TraesCS2D01G100900	TraesCS6A01G281000	0.011425

TraesCS2D01G100900	TraesCS3D01G090100	0.011413
TraesCS2D01G100900	TraesCS6D01G353400	0.011259
TraesCS2D01G100900	TraesCS7A01G114700	0.011238
TraesCS2D01G100900	TraesCS3B01G090400	0.011236
TraesCS2D01G100900	TraesCS4A01G495000	0.011183
TraesCS2D01G100900	TraesCS1D01G144400	0.011166
TraesCS2D01G100900	TraesCS2B01G048700	0.011069
TraesCS2D01G100900	TraesCS1B01G364200	0.01095
TraesCS2D01G100900	TraesCS4A01G411000	0.010931
TraesCS2D01G100900	TraesCS5A01G375000	0.010908
TraesCS2D01G100900	TraesCS3A01G406000	0.010888
TraesCS2D01G100900	TraesCS3A01G309500	0.010854
TraesCS2D01G100900	TraesCS5B01G291100	0.01072
TraesCS2D01G100900	TraesCS7A01G537100	0.010633
TraesCS2D01G100900	TraesCS2A01G318600	0.010622
TraesCS2D01G100900	TraesCS5A01G319900	0.010606
TraesCS2D01G100900	TraesCS5A01G206600	0.01058
TraesCS2D01G100900	TraesCS5A01G537500	0.010574
TraesCS2D01G100900	TraesCS7A01G199300	0.0104
TraesCS2D01G100900	TraesCSU01G176500	0.010399
TraesCS2D01G100900	TraesCS3B01G357200	0.010398
TraesCS2D01G100900	TraesCS1B01G135700	0.010395
TraesCS2D01G100900	TraesCS7D01G518200	0.01036
TraesCS2D01G100900	TraesCS2B01G382100	0.010212
TraesCS2D01G100900	TraesCS5B01G469400	0.01018
TraesCS2D01G100900	TraesCS2D01G034900	0.010172
TraesCS2D01G100900	TraesCS3D01G078900	0.010165
TraesCS2D01G100900	TraesCS1B01G444000	0.010111
TraesCS2D01G100900	TraesCS7A01G338900	0.010104
TraesCS2D01G100900	TraesCS2D01G581800	0.0101
TraesCS2D01G100900	TraesCS2A01G529900	0.010093
TraesCS2D01G100900	TraesCS3D01G401200	0.010069
TraesCS2D01G100900	TraesCS2D01G407100	0.010061
TraesCS2D01G100900	TraesCS5B01G368300	0.010054
TraesCS2D01G100900	TraesCS4B01G339800	0.010017
TraesCS2D01G100900	TraesCS3A01G283800	0.010013
TraesCS2D01G100900	TraesCS7A01G127800	0.00997
TraesCS2D01G100900	TraesCS3A01G137400	0.009965
TraesCS2D01G100900	TraesCS5A01G230200	0.009946
TraesCS2D01G100900	TraesCS5B01G321800	0.009935
TraesCS2D01G100900	TraesCS6D01G157700	0.00992
TraesCS2D01G100900	TraesCSU01G048500	0.009829
TraesCS2D01G100900	TraesCS2A01G102000	0.009799
TraesCS2D01G100900	TraesCS7D01G173300	0.0097
TraesCS2D01G100900	TraesCS5D01G443100	0.009683
TraesCS2D01G100900	TraesCS1D01G392300	0.009678
TraesCS2D01G100900	TraesCS5B01G204800	0.009621
TraesCS2D01G100900	TraesCS4B01G064800	0.009593

TraesCS2D01G100900	TraesCS5B01G210100	0.009588
TraesCS2D01G100900	TraesCS1D01G174800	0.009584
TraesCS2D01G100900	TraesCS3A01G520800	0.009582
TraesCS2D01G100900	TraesCS7D01G248800	0.009527
TraesCS2D01G100900	TraesCS7B01G170500	0.009497
TraesCS2D01G100900	TraesCS2B01G119100	0.009494
TraesCS2D01G100900	TraesCS3D01G059300	0.009457
TraesCS2D01G100900	TraesCS1D01G066300	0.009434
TraesCS2D01G100900	TraesCS5B01G228800	0.009431
TraesCS2D01G100900	TraesCS4D01G229900	0.009252
TraesCS2D01G100900	TraesCS4B01G377800	0.009239
TraesCS2D01G100900	TraesCS3D01G046100	0.009214
TraesCS2D01G100900	TraesCS1D01G088400	0.009143
TraesCS2D01G100900	TraesCS2B01G027700	0.009126
TraesCS2D01G100900	TraesCS3B01G278200	0.009113
TraesCS2D01G100900	TraesCS6D01G276900	0.009105
TraesCS2D01G100900	TraesCS2A01G520500	0.009101
TraesCS2D01G100900	TraesCS5D01G358600	0.009059
TraesCS2D01G100900	TraesCS4D01G012100	0.009014
TraesCS2D01G100900	TraesCS3A01G526400	0.008993
TraesCS2D01G100900	TraesCS7D01G175900	0.008992
TraesCS2D01G100900	TraesCS7D01G383500	0.00899