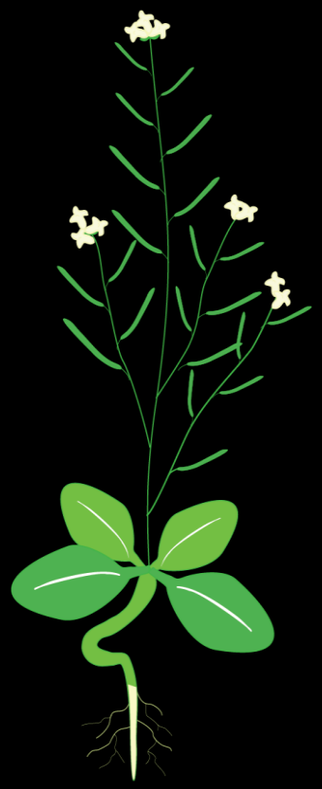
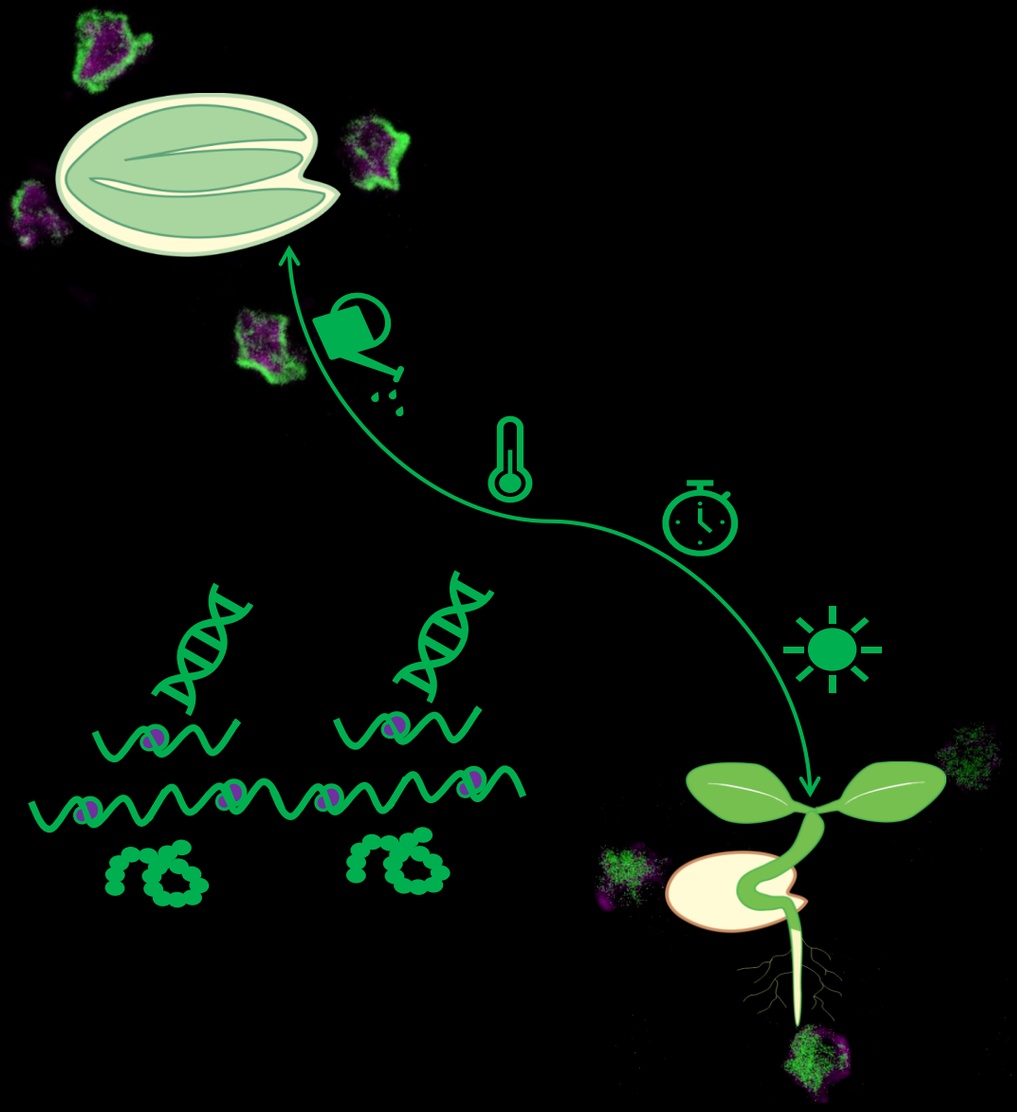


Transcriptional regulation of seed dormancy and seed longevity

Shuang Song

2021

Transcriptional regulation of seed dormancy and seed longevity



Shuang Song

Propositions

1. The analysis of 5’ monophosphorylated mRNAs with an intact 3’ Poly(A) tail isolated from dry seeds allows the identification of mRNAs that are important for seed longevity.
   (this thesis)
2. The artificial seed ageing technique elevated partial pressure of oxygen (EPPO) should replace natural ageing in seed research.
   (this thesis)
3. P-values should not be abandoned as it is the only fair approach to judge the reliability of research.
   (Inspired by Pike H. Statistical significance should be abandoned, say scientists. BMJ. 2019 Mar 25;364:l1374. doi: 10.1136/bmj.l1374. PMID: 30910774.)
4. A powerful hypothesis does not reduce the risks of failures in science.
5. The thin line between being impolite and funny is defined by language skill abilities.
6. Practice makes perfect is an excuse for repetition.

Propositions belonging to the thesis, entitled

Transcriptional regulation of seed dormancy and seed longevity

Shuang Song

Wageningen, 23 March 2021

Transcriptional regulation of seed dormancy and seed longevity

Shuang Song

Thesis committee

Promotor

Prof. Dr Leónie Bentsink

Personal Professor at the Laboratory of Plant Physiology

Wageningen University & Research

Other members

Prof. Dr GH Immink - Wageningen University & Research

Dr H.N. North - Institut Jean - Pierre Bourgin, INRAE - Centre de Versailles - Grignon, Versailles, France

Dr J Buitink - Institut National de la Recherche Agronomique (INRA), Angers, France

Dr F Lanfermeijer - Senior researcher - Bejo Zaden BV, Warmenhuizen

This thesis was conducted under the auspices of the Graduated School of Experimental Plant Sciences.

Transcriptional regulation of seed dormancy and seed longevity

Shuang Song

Thesis

submitted in fulfilment of the requirements for the degree of doctor

at Wageningen University

by the authority of the Rector Magnificus,

Prof. Dr A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Tuesday 23 March 2021

at 1.30 p.m. in the Aula

Shuang Song

Transcriptional regulation of seed dormancy and seed longevity

168 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2021)

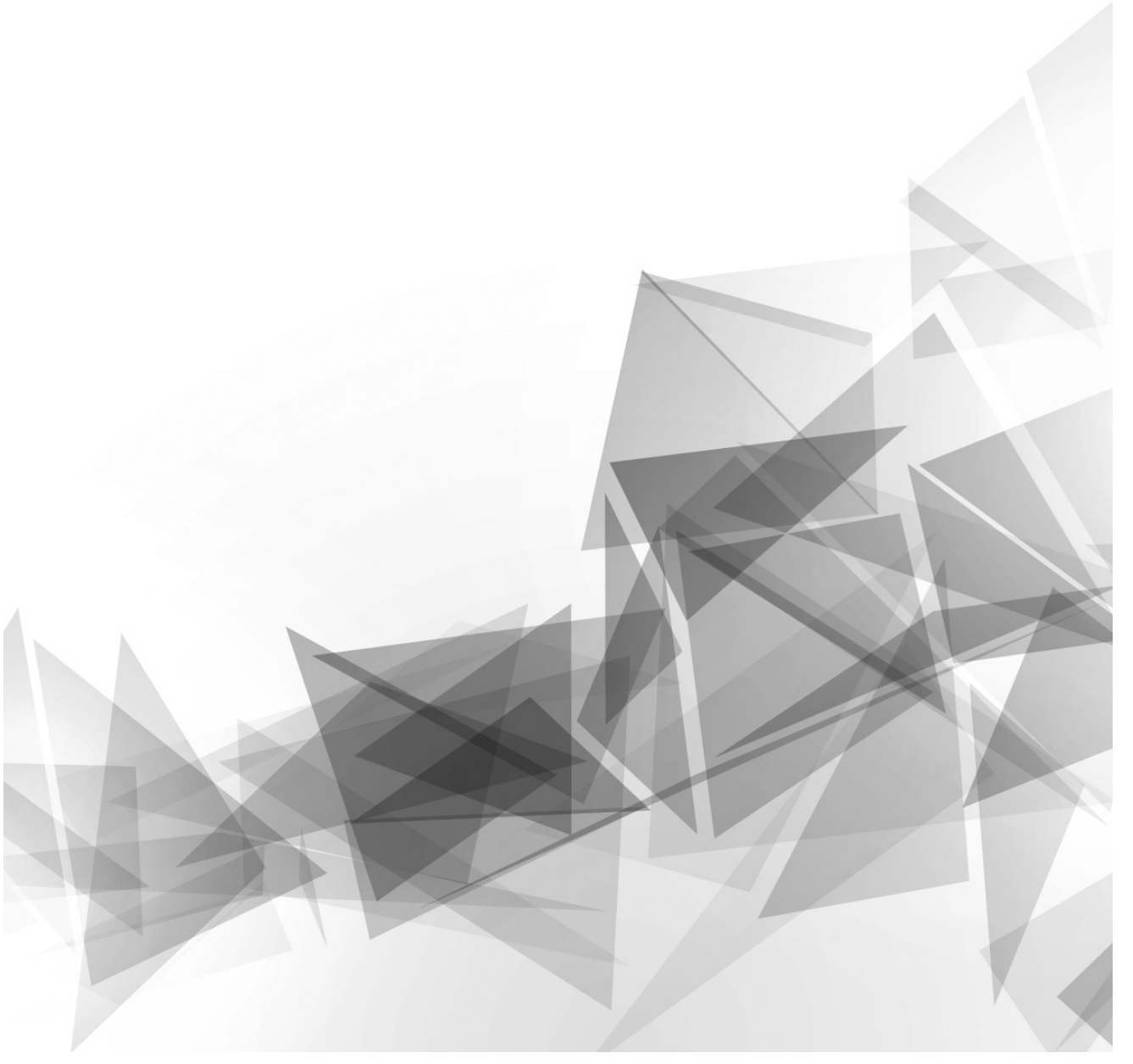
With references, with summary in English

ISBN: 978-94-6395-716-8
DOI: https://doi.org/10.18174/541890

Contents

Chapter 1

General Introduction



Seed and life

Seeds play critical roles for maintaining life on earth. Seeds are rich in proteins, carbohydrates, oil and minerals and as such important food resources. Rice, maize and wheat are the three main crops of which seeds supply human in their daily nutrition across the world. As raw materials, seeds are used in numerous ways, such as medicines, decorations, toys and fibers. Seeds have a unique structure, containing surrounding layers that protect the embryo. Moreover, seeds allow plants to survive unfavorable conditions by delaying germination until the establishment of the next generation can be guaranteed. As a sexual product, carrying recombinant genetic information from its parents, seeds contribute to the genetic diversity of species. This allows the offspring to inherit beneficial traits that contribute to the survival in unfriendly environments, which is an imperative solution to face the global climate changes (Bewley *et al.*, 2013; Boswell, 1961; Koornneef *et al.*, 2002; Vellvé, 2013). Overall, without seeds there will be no survival on earth.

The model plant *Arabidopsis thaliana*

To improve seed quality and yield, research focused on the timing of germination, seed dispersal, seed persistence and seedling establishment have become and will remain a hot topic in plant biology (Black, 1989; Saatkamp *et al.*, 2019). Arabidopsis is an ideal model system for this research (Koornneef *et al.*, 2010; North *et al.*, 2010; Wienkoop *et al.*, 2010). Its, mini plant size, short time to produce large amounts of seeds by self-pollination, five completely sequenced chromosomes and over 100.000 available T-DNA lines, make it a widely used model plant for physiological and molecular studies since Friedrich Laibach identified the chromosome numbers in Arabidopsis (Koornneef *et al.*, 2010; Laibach, 1907; Scholl *et al.*, 2000). Discoveries in Arabidopsis have greatly increased our understanding of plant biological processes, but may not have an immediate and direct societal impact (Stephenson *et al.*, 2019). However, recently whole genome sequences of several crops, including *Brassica napus* (Chalhoub *et al.*, 2014; Lawrenson *et al.*, 2015), wheat (Uauy, 2017), maize (Nannas *et al.*, 2015) and rice (Wing *et al.*, 2018) have become available. Combining this sequence information with genetic engineering

techniques such as CRISPR-Cas will make it feasible to transfer knowledge obtained in Arabidopsis to crops (Bevan *et al.*, 2005; Chang *et al.*, 2016; Knott *et al.*, 2018; Makarova *et al.*, 2011; Piquerez *et al.*, 2014).

Seed development and its regulators

In Arabidopsis, seed development consists of two main stages. Stage I, the embryogenesis, starts with a single zygotic cell and an endosperm maternal cell after double fertilization and ends with a complete embryo architecture surrounded by an endosperm (Mansfield *et al.*, 1991; Mayer *et al.*, 1998; Mayer *et al.*, 1991). Stage II is the seed maturation stage. In the beginning of maturation the embryo expands through cell divisions to fill up the seed sac whereas the endosperm scales down to one cell layer, and starch, proteins and fatty acids accumulate (Goldberg *et al.*, 1994; Mansfield *et al.*, 1994). During late maturation, desiccation tolerance is induced, the seed water content decreases dramatically to become less than 10% and metabolic activities terminate. Furthermore, the accumulation of oligosaccharides still continues and the embryo acquires primary dormancy (Baud *et al.*, 2002; Baud *et al.*, 2008; Fait *et al.*, 2006; Goldberg *et al.*, 1989). Through genetic studies, four key seed maturation regulators, *ABSCISIC ACID-INSENSITIVE 3* (*ABI3*), *LEAFY COTYLEDON 1* (*LEC1*), *LEC2*, and *FUSCA 3* (*FUS3*) have been identified. In summary, mutants in these genes exhibit absence of seed dormancy and short seed longevity, indicating that seed dormancy and seed longevity are closely related to seed maturation (Bäumlein *et al.*, 1994; Clerkx *et al.*, 2004; Holdsworth *et al.*, 2008; Meinke, 1992; Meinke *et al.*, 1994; Nambara *et al.*, 1995; Ooms *et al.*, 1993; Santos-Mendoza *et al.*, 2008; To *et al.*, 2006). The key regulators of seed maturation are known to work synergistically to control seed development. *ABI3* cooperates with *LEC1* and *FUS3* to regulate elementary seed developmental programs, for instance, the amount of chlorophyll and anthocyanin is significantly higher in the embryos of the *fus3 abi3-1* and *lec1 abi3-1* double mutants than the single mutants *abi3-1*, *fus3* and *lec1* mutants. The abundance of ABI3 protein is also predicted to be controlled by *FUS3* and *LEC1* (Parcy *et al.*, 1997). *LEC2* activates *FUS3* during root initiation (Tang *et al.*, 2017). Moreover, these four genes have also been individually studied

and this showed that they are involved in controlling distinct biological processes: *ABI3* protects seeds during desiccation and negatively influences floral transition (Hong *et al.*, 2019; Khandelwal *et al.*, 2010). *LEC1* regulates the expression of fatty acid biosynthetic genes and thus influence oil yield in seeds (Mu *et al.*, 2008) and transcriptionally regulates photosynthesis during late seed development (Pelletier *et al.*, 2017). *LEC2* affects seed storage protein accumulation and oil body production in vegetative tissues (Stone *et al.*, 2008). *FUS3* triggers the accumulation of triacylglycerol in Arabidopsis seedlings (Zhang *et al.*, 2016). Ultimately, at the end of seed maturation the dry, dormant seed is ready to be harvested for future use.

Seed traits

Seed quality is important to meet the increasing demand for food and face the severe climate change. Vital seed traits that are valuable for agricultural benefit include, large seed size, high nutritional value, proper timing of germination and high seed longevity. These traits have been modified by intended or unintended selections during the domestication of crops (Godfray *et al.*, 2010; Martínez-Andújar *et al.*, 2012). Moreover, these traits are regulated by numerous regulators at the transcriptional, translational, hormonal and epigenetic levels, involving various pathways during seed development and maturation (Savadi, 2018).

Seed dormancy and its regulators

Seed dormancy has been defined as the inability of the seed to germinate even under favourable environmental conditions (Bewley, 1997; Wareing, 1971). Low seed dormancy might cause preharvest sprouting in damp conditions, whereas deep seed dormancy will distribute germination over time. Nonoptimal dormancy levels potentially decrease the crop yield in agriculture, thus studies on understanding the mechanisms of controlling seed dormancy could contribute to more efficient crop production. Usually, the term dormancy refers to primary dormancy that is induced during seed maturation, however there are other types of dormancy, like e.g., secondary dormancy that can be induced after harvest under natural or artificial conditions (Buijs *et al.*, 2020; Footitt *et al.*, 2017; Khan, 1980). Here I focus on

primary seed dormancy since this directly links to agriculture whereas secondary seed dormancy is more related to ecological systems.

Previous studies indicated that seed dormancy could be determined by diverse factors. Plant hormones including abscisic acid (ABA), gibberellins (GA), auxins, ethylene, brassinosteroids (BR) and cytokinins strongly influence plant development. ABA enhances seed dormancy whereas GA breaks seed dormancy (Bewley, 1997; Brocard-Gifford *et al.*, 2004; Finkelstein *et al.*, 2008; Koornneef *et al.*, 2002; Miransari *et al.*, 2014; Nambara *et al.*, 2010; Yamaguchi *et al.*, 2018). Auxin is indicated to be involved in the regulation of dormancy and preharvest sprouting in wheat (Ramaih *et al.*, 2003). Auxin overproducing transgenics displaying stronger seed dormancy than wild type Arabidopsis seeds (Liu *et al.*, 2013; Ramaih *et al.*, 2003). Dormancy induced by high temperatures is related to the reduction of ethylene in sunflower and lettuce (Corbineau *et al.*, 1988; Prusinski *et al.*, 1990), while dormancy release by nitric oxide causes the elevation of ethylene (Arc *et al.*, 2013). BR stimulate seed germination in Arabidopsis and tobacco (Leubner-Metzger, 2001; Steber *et al.*, 2001). Exogenous cytokinins promote *Lotus corniculatus* L. seed germination (Nikolić *et al.*, 2006). The interactions between the different plant hormones and their effects on seed dormancy have been extensively reviewed (Kucera *et al.*, 2005). ABA and GA antagonistically regulate seed dormancy and germination (Yamaguchi *et al.*, 2018). Auxin and ABA synergistically inhibit seed germination (Brady *et al.*, 2003; Liu *et al.*, 2013). Ethylene counteracts with ABA to promote endosperm rupture, both in seeds of *Arabidopsis thaliana* and *Lepidium sativum*. In addition, the rapid decrease of ABA levels in *Lepidium* seeds after the application of different concentrations of ethylene suggests that exogenous ethylene does not interfere with endogenous ABA contents (Linkies *et al.*, 2009). BR can break seed dormancy caused by ABA in Arabidopsis and promote seed germination by strengthening the embryo growth potential (Leubner-Metzger, 2001). Cytokinins play opposite roles to ABA since they promote seed germination in Arabidopsis (Wang *et al.*, 2011).

Mutant studies of genes regulating plant hormones also reveal the respective functions in seed dormancy and germination. The *ABA deficient 2* mutant (*aba2*) exhibits a yellower embryo and reduced seed dormancy compared with wild type (Nambara *et al.*, 1998). The GA deficient mutant, *ga1-3* could not germinate in the absence of exogenous GA application (Groot *et al.*, 1987). Arabidopsis auxin receptor and biosynthesis mutants have completely lost seed dormancy (Liu *et al.*, 2013). The ethylene insensitive mutant, *ethylene insensitive 2* displays an enhanced seed dormancy (Subbiah *et al.*, 2010). The BR insensitive mutant *bri1-1* and the biosynthetic mutant *det2-1* are more sensitive to ABA and thus have a lower germination potential (Steber *et al.*, 2001). Cytokinin receptor mutants in Arabidopsis germinate faster than wild type (Riefler *et al.*, 2006). Apart from these hormone related mutants, *dof affecting germination 2* (*dag2*) demonstrates a defective seed germination which is regulated by light and cold treatment (Gualberti *et al.*, 2002). That germination is strongly affected by light also becomes clear when investigating the germination of the phytochrome A and B mutants (*phya*, *phyb*). These two phytochromes control seed germination by sensing light within a broad spectral range. PhyA induces seed germination after irradiation at remarkably low fluence from the UV-A to the visible and the far-red light spectrum (300-780 nm). In contrast, PhyB regulates seed light sensing at much higher (10^4 times) fluences with a response in the red and far-red light (540-690 nm and 695-780 nm) (Shinomura *et al.*, 1996).

Seed dormancy is an adaptive trait that can be affected by temperature, relative humidity, altitudinal and climatic patterns both during seed development and maturation as well as during seed storage. When germinated in a controlled environment the dormancy of wheat seeds increases with elevated temperatures, from 10, 15, 20, until 30 °C, the effect is the strongest between 15 and 20 °C (Nyachiro *et al.*, 2002). During seed dry storage, dormancy is released, a process that is referred to as after ripening. Experiments on Arabidopsis seeds with different levels of seed dormancy show that a relative humidity of 50% at 25 °C is the most efficient in releasing dormancy across all genotypes within seven weeks. The critical

seed moisture content for dormancy release is calculated to be 0.06g H_2O/g dry weight, below this value the dormancy release is not possible above this value dormancy release is elevated with increasing temperatures (10, 15, 20, 25 °C) (Basbouss-Serhal *et al.*, 2016). During seed development, Arabidopsis accessions grown in environmental conditions with high temperature and high solar radiation and low rainfall in summer present high seed dormancy levels and thus low seed germination levels (Vidigal *et al.*, 2016). When the seeds are artificially buried in the field for two years and thus exposed to fluctuated soil temperature and humidity, they show dormancy cycling. Dormancy is induced at low temperatures whereas high temperatures release seed dormancy in the soil (Buijs *et al.*, 2020).

The environment is also a cue in providing natural genetic variation, resulting in accessions with different levels of seed dormancy, for example, the accession Landsberg *erecta* (L*er*) that is originally collected in Poland has low dormancy and the accession of the Cape Verde Islands (Cvi) has strong dormancy (Lawrence, 1976; Ratcliffe, 1976). This genetic variation allows Quantitative trait loci (QTL) analysis to identify seed dormancy loci. Several recombinant inbred line (RIL) populations have been developed by crossing accessions that have different dormancy levels. These RIL populations include L*er*/Cvi, L*er*/Antwerp-1, L*er*/Kashmir, L*er*/Shakdara, L*er*/Santa Maria da Feira-0 (Fei-0) and L*er*/Kondara. All populations segregate for seed dormancy indicating that both accessions contain loci that increase and decrease seed dormancy. The L*er*/Fei-0 and L*er*/Kashmir populations have the highest levels of seed dormancy, which was revealed by comparing the days of seed dry storage to reach 50% germination (DSDS50) at the population level (Bentsink *et al.*, 2010). Several of the strong QTL have been validated using near isogenic lines (NILs). The NILs can be used to further characterize the loci and identify the genes underlying the QTL (Alonso-Blanco *et al.*, 2000; Bentsink *et al.*, 2010). Using these genetic methods, a group of seed dormancy loci, *DELAY OF GERMINATION* (*DOG*), were identified (Fig. 1A) (Alonso-Blanco *et al.*, 2003; Bentsink *et al.*, 2010). Of these *DOG1* is the strongest and first cloned seed dormancy QTL. *DOG1* is characterized to encode AT5G45830 which is

a seed specific protein with controls seed dormancy (Fig. 1B) (Bentsink *et al.*, 2010; Bentsink *et al.*, 2006). *DOG1* expression increases at low temperatures during seed maturation, which leads to increased dormancy levels. Its expression also corelates with geographical position, seeds that are matured in south of Europe contain higher levels of DOG1 (transcript and protein) and are more dormant than seeds that are produced in the north of Europe (Chiang *et al.*, 2011). DOG1 protein accumulates and remains stable during seed maturation and storage (Nakabayashi *et al.*, 2012). The unknown nature of the DOG1 protein has severely hampered unrevealing its mechanism, however two recent publications show that the DOG1 protein binds to HYPERSENSITIVE GERMINATION 1 (AHG1) and AHG3, two members of the clade A PP2C phosphatases in seeds (Nee *et al.*, 2017, Nishimura *et al.*, 2018). Mutant analyses revealed that both AHG1 and AHG3 are required for DOG1 functioning. These phosphatases play an important role in ABA signaling and are therefore proposed to represent a converging point for ABA and DOG1 in the regulation of seed dormancy (Soppe *et al.*, 2020). The only other Arabidopsis dormancy QTL that is cloned so far is *DOG18,* this locus is identified as *REDUCED DORMANCY5* (*RDO5*) encoding a protein phosphatase 2C (PP2C), which positively regulates seed dormancy without influencing ABA and DOG1.The *rdo5* mutant displays an extreme low seed dormancy phenotype. Its expression is seed specific and highly abundant in dry seeds. Expression analyses also revealed that RDO5 reduces the expression of Arabidopsis the RNA binding proteins *PUMILIO9* (*APUM9*) and *APUM11*. Moreover, RDO5 acts as a pseudophosphatase and as such prohibits dephosphorylation during early seed germination (Xiang *et al.*, 2014; Xiang *et al.*, 2016).

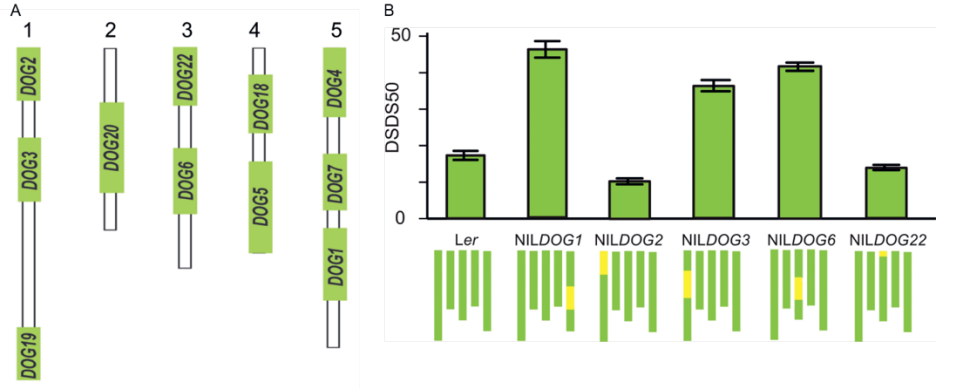


Figure 1. Seed dormancy QTL and their confirmation. A, Seed dormancy quantitative trait loci (QTL) identified using the Arabidopsis six recombinant inbred line populations. The QTL are named *DELAY OF GERMINATION* (*DOG*). The numbers, 1-5, indicate the chromosomes numbers. Green bars indicate the positions of the loci on the chromosomes. B, Dormancy phenotypes of L*er* and *DOG* near isogenic lines. DSDS50 stands for days after seed dry storage to reach 50% germination, the higher the value the deeper seed dormancy and *vice versa*. Data obtained from (Bentsink *et al.*, 2010).

QTL analyses to study seed dormancy regulators have not only been applied in Arabidopsis, but also on crops including wheat (Kato *et al.*, 2001; Torada *et al.*, 2005), wild oat (Fennimore *et al.*, 1999), rice (Guo *et al.*, 2004; Xie *et al.*, 2011), sunflower (Gandhi *et al.*, 2005), winter oilseed rape (Schatzki *et al.*, 2013) and two rowed barley (Li *et al.*, 2003). A QTL study in rice identified the *qsd-1-1* locus that encodes the DOG1 like protein OsDOG1L-3. OsDOG1L-3 affects seed dormancy by enhancing the seed ABA content, which is followed by an increased expression of ABA regulated genes (Wang *et al.*, 2020).

Seed longevity and its regulators

Seed longevity is the capacity of a seed to keep its vigor under dry storage (Rajjou *et al.*, 2008). Understanding the underlying mechanisms of seed longevity and using this knowledge to extend the seed life span, at low cost, will be beneficial for worldwide agriculture and plant diversity (Li, *et al.*, 2009; Long *et al.*, 2015).

There are various factors influencing seed longevity. Firstly, the seed coat serves as a protective tissue to prevent the embryo and its seed reserves from detrimental environmental factors. Mutants with a defected seed coat exhibited reduced seed longevity compared to wild type (Clerkx *et al.*, 2004; Debeaujon *et al.*, 2000; Rajjou, *et al.*, 2008). Secondly, seed longevity is regulated by physiological factors. The vitamin E mutants, *vte1* disrupting the activity of tocopherol cyclase and *vte2* interrupting the activity of homogentisate phytyl transferase, both result in a lack of tocopherols and have a much shorter seed longevity than wild type seeds. This indicates the critical role of tocopherols in keeping seed vigor in dry seeds (Sattler *et al.*, 2004). Thirdly, environmental factors largely affect seed longevity. The effect of temperature during storage was tested on different plant species including groundnut, onion, sugar beet, barley, chickpea, wheat and cowpea and have shown a conserved quantitative relationship between storage temperature (from -13°C to 90°C) and seed longevity, where higher temperatures correlate with lower seed longevity (Dickie *et al.*, 1990). For *Phaseolus vulgaris* a negative logarithmic relationship between seed longevity and moisture content is determined, the optimal relative humidity in equilibrium with seed moisture content is suggested to be around 10% for keeping seed viability (Ellis *et al.*, 1989; Ellis *et al.*, 1990). A study examining the viability of red clover and alfalfa seeds stored at various moisture contents and temperatures during 14.5 years of hermetic storage shows that moisture content negatively correlates to seed longevity. Moreover, at low moisture content seed longevity is not affected by temperature, even 35°C at 2% relative humidity is not harmful for seed survival. (Ellis *et al.*, 2006). A combination experiment of six temperatures -170, -25, 5, 20, 35, 45°C with eight water contents from 0.21 to 0.01 $g \cdot H_2O \cdot g^{-1}$ DW in seven Brassicaceae species containing *Eruca vesicaria*, *Sisymbrium runcinatum*, *Rorippa nasturtium-aquaticum*, *Brassica repanda*, *Moricandia arvensis*, *Malcolmia littorea*, *Sinapis alba*, established that the critical water content for seed storage is between 0.02 to 0.03 $g \cdot H_2O \cdot g^{-1}$ at 45°C (Mira *et al.*, 2015).

Seed longevity is also an adaptive trait for which genetic variation is present in nature. Seed longevity has been defined as the capacity to remain a high germination ability after storage and can be examined by evaluating seed germination percentages during seed dry storage (Rajjou *et al.*, 2008). The time that seeds survive storage is species dependent. The longest survival time reported is that of the desiccation tolerant *Phoenix dactylifera* L seeds that survived for approximately 2000 years (Sallon *et al.*, 2008). Seeds of the model plant Arabidopsis can survive around 10 years in dry conditions (Bentsink *et al.*, 2000). The long survival time of dry seeds in dry conditions (lab bench storage) hamper the unravelling of mechanisms regulating seed longevity. Therefore, often artificial seed ageing techniques are applied to shorten the seed life span. The controlled deterioration test (CDT) and artificial ageing (AA) are based on applying high temperature and humidity, whereas the elevated partial pressure of oxygen (EPPO) method is based on increasing the absolute oxygen concentrations at ambient air conditions (Groot *et al.*, 2012; Rajjou *et al.*, 2008) (Tesnier *et al.*, 2002). In Arabidopsis, through QTL analysis, eleven loci which are related to seed storability *GERMINATION ABILITY AFTER STORAGE 1* (*GAAS1*) to *GAAS5* have been identified in six RIL populations after lab bench storage (Bentsink *et al.*, 2000; Nguyen *et al.*, 2012). These QTL only partly overlap with an earlier study using CDT, implying that CDT is not sufficient to mimic lab bench seed ageing (Buijs *et al.*, 2020; Nguyen *et al.*, 2012). A recent study, in which EPPO treatment was compared to lab bench storage, showed that EPPO is so far the best method to accelerate and mimic seed ageing in laboratory conditions (Buijs *et al.*, 2020; Groot *et al.*, 2012). Interestingly, the identified *GAAS* loci collocate with *DOG* loci (dormancy) and revealed a negative relationship between seed longevity and seed dormancy in natural variation (Alonso-Blanco *et al.*, 2003; Bentsink *et al.*, 2010; Nguyen *et al.*, 2012), a good reason to study these two traits simultaneously in seed life span experiments as we do in Chapter 4 of this thesis.

Seed dry storage and oxygen

Seeds of many species acquire a low moisture content during seed maturation and this contributes to extending their survival period (Buitink *et al.*, 2008; Sun *et al.*,

1994). An example of a remarkably long seed lifespan is that of Sacred Lotus seeds from China that were able to germinate after more then 1000 years of storage (Shen-Miller *et al.*, 1995). However, seeds eventually lose vigor even in dry and cool conditions, this deterioration is suggested to be caused by oxygen (Hendry, 1993). Oxygen interacts with seed components already during early embryogenesis, producing a mixture of molecules and free radicals that are called reactive oxygen species (ROS) (Bailly, 2004; Hendry, 1993). ROS accumulate over the seed life span and play bifold and crucial roles in seed biology. On one hand, suitable ROS contents are beneficial for the completion of seed germination, by for instance overcoming seed dormancy (Bazin *et al.*, 2011; Bouteau *et al.*, 2007; Oracz *et al.*, 2007). On the other hand, too high ROS levels are detrimental for seeds, for instance, causing mRNA and protein oxidation by randomly attacking RNA bases and producing protein carbonylation respectively that leads to loss of viability (Bailly, 2019; Bailly *et al.*, 2008; Bouteau *et al.*, 2007; Dalle-Donne *et al.*, 2006; Fleming *et al.*, 2018; Oracz *et al.*, 2007; Turrens, 2003). Seed stored mRNAs have been suggested to be potentially protected by monosome complexes, stress granules or processing bodies (Bai *et al.*, 2020; Sajeev *et al.*, 2019) and are sufficient to complete seed germination without *de novo* mRNAs synthesis. In contrast, seed stored proteins are not enough to promote seed germination, implying the stored mRNAs have to be translated to new proteins to initiate seed germination (Rajjou *et al.*, 2004). Further knowledge about the protection of mRNAs and proteins from ROS might allow the improvement of seed longevity. Moreover, identifying critical mRNAs and proteins for seed germination might be beneficial to improve seed conservation in the future.

Scope of the thesis

In this thesis, I used Arabidopsis seeds to study the transcriptional regulation of seed dormancy and longevity. Understanding the fundamental mechanisms of these important seed traits is crucial for crop improvement in the future.

In Chapter 2 I describe the functional analyses of the *DOG6* QTL (Fig. 1). By fine-mapping and complementation cloning I show that *DOG6* encodes the ANAC

transcription factor *ANAC060*. The clear seed dormancy phenotype of *anac060* mutants evidenced its inhibition of seed dormancy. Through sequencing analysis in different accessions, *ANAC060* was identified to have allelic diversities related to natural variation. A combination analysis from different sources of data revealed the potential downstream targets of *ANAC060*. Overall, how *ANAC060* regulates seed germination and its interaction with other regulators of seed germination has been described.

In **Chapter 3**, *ANAC060* and its closest homologs *ANAC040*, *ANAC089* were investigated in order to reveal possible functional redundancy. I have compared the sequences of the three homologous genes and search for motifs in their promoters. Furthermore, I have performed swapping experiments in which the promoters and coding sequences of *ANAC040* and *ANAC089* were exchanged with the ones of *ANAC60* and transformed to the *anac060* and *anac040* mutant backgrounds. The transgenic lines were phenotypically characterized to determine whether the homologs could cross complement each other's mutant phenotype.

In **Chapter 4**, the degradation profile of seed stored transcripts and proteins was explored in seed that were aged using the elevated partial pressure of oxygen (EPPO) artificial ageing system. Degradome sequencing was performed to determine whether the loss of viability could be explained by the degradation of specific mRNAs. Degradation patterns and differentially degraded transcripts were investigated and evaluated for their importance during seed germination.

In **Chapter 5**, the findings in the thesis and future perspectives are discussed.

Chapter 2

DELAY OF GERMINATION 6, encoding the *ANAC060* transcription factor, inhibits seed dormancy

S. Song, H. He, K. Gühl, M. van Bolderen, G. Buijs, L.A.J. Willems, L. Bentsink

Abstract

The timing of seed germination is regulated by seed dormancy. There is ample natural variation for seed dormancy among as well as within plant species. In Arabidopsis several *DELAY OF GERMINATION* quantitative trait loci have been identified, of which *DOG1* is best studied. Here we report the identification of *DOG6*, a quantitative trait locus with a similar strong effect on seed dormancy as *DOG*1. *DOG6* affects the timing of germination both in laboratory as well as in field conditions. Complementation cloning revealed that *DOG6* encodes the membrane bound transcription factor *ANAC060*. The absence of the ANAC060 protein or its sequestration outside the nucleus results in increased seed dormancy levels. The different natural variants of *ANAC060* differ for the presence of the membrane binding domain, either due to the fact that this domain is absent in the genomic sequence or because the cDNA is alternatively spliced. Our data indicates that ANAC060, when present in the nucleus, inhibits seed dormancy by attenuating the expression of protein phosphatases 2C class A proteins including *ABI 5 BINDING FACTOR 3*, *REDUCED DORMANCY5/DOG18* and *ABA insensitive 4*.

Significance Statement ANAC transcription factors are known to effect plant development as well as the response of plants to their environment. Here, we present the identification of *DELAY OF GERMINATION* 6 (*DOG6*), a seed dormancy quantitative trait locus that encodes the *ANAC060* transcription factor. We have identified different natural alleles of *ANAC060* and show that these genetic variants determine the localization of the protein. Only proteins encoded by alleles that lack the membrane binding domain end up in the nucleus. Here, they can affect transcription and as such attenuate seed dormancy.

Introduction

The timing of seed germination is essential for the survival of seed plants. This timing of germination is controlled by seed dormancy. Seed dormancy is an important life history trait for which ample genetic variation is present in nature. Studies employing this natural variation in *Arabidopsis thaliana* European accessions have revealed that low altitudes, high latitudes, high temperature, low summer precipitation and high radiation correlate with high seed dormancy levels (Debieu *et al.*, 2013; Vidigal *et al.*, 2016). The loci underlying the natural genetic variation in Arabidopsis have been identified using quantitative trait loci (QTL) mapping. In total eleven QTL have been revealed, of which two *DELAY OF GERMINATION 1* (*DOG1*) and *DOG18* (*RDO5*) have been cloned (Bentsink *et al.*, 2006; Xiang *et al.*, 2014). Mutant analyses have shown that *DOG1* functions down-stream of the ethylene receptor ETR1 and the ethylene response factor ERF12. The functioning of ERF12 depends on interaction with TOPLESS (TPL), together these proteins repress the expression of *DOG1* in the presence of ethylene (Li *et al.*, 2019). Moreover, *DOG1* expression is known to be increased by cold (Kendall *et al.*, 2011). Bryant *et al.* (2019) showed that LEUCINE ZIPPER TRANSCRIPTION FACTOR67 (bZIP67) protein increases in abundance when seeds perceive cold during maturation and that this transactivates *DOG1* and thereby helps to establish primary dormancy. In mature plants *DOG1* is down regulated by its antisense transcript (as*DOG1*) (Yatusevich *et al.*, 2017). Inhibition by the plant hormone abscisic acid (ABA) and drought inhibits *asDOG1* expression, and therefore leads to increased *DOG1* sense transcript. The depth of dormancy correlates with the amount of DOG1 protein which loses its function during seed dry storage (after-ripening), conditions that also lead to dormancy reduction (Nakabayashi *et al.*, 2012). Recently a model has been proposed in which DOG1 controls seed dormancy by its physical interaction with the phosphatases ABA-HYPERSENSITIVE GERMINATION 1 and 3. This interaction functionally blocks their roles in the release of seed dormancy (Nee *et al.*, 2017). DOG18/RDO5 is a member of the protein phosphatase 2C family of that does not

show phosphatase activity and is thought to act independently of DOG1 (Xiang *et al.*, 2014; Xiang *et al.*, 2016).

Here we describe the cloning of *DOG6*, the second strongest dormancy QTL in the genetic screen in which also *DOG1* and *DOG18*/*RDO5* have been identified (Alonso-Blanco *et al.*, 2003; Bentsink *et al.*, 2010). We have previously shown that several natural alleles of *DOG6* increase seed dormancy in the Landsberg *erecta* genetic background (Bentsink *et al.*, 2010). These include the Cape Verde Islands, Kashmir and Shahdara alleles, the latter two have been used for the fine mapping of the *DOG6* QTL. Complementation cloning revealed that *DOG6* encodes the membrane bound transcription factor *ANAC060* (*AT3G44290*). We show that the natural alleles of *ANAC060* determine the dormancy behaviour based on the fact of the presence of the membrane binding domain. The membrane binding domain retains the protein from the nucleus, genotypes in which ANAC060 is localized in the nucleus have a non-dormant phenotype. We show that *ANAC060* and *DOG1* regulate seed dormancy by additive pathways and propose a role for *ANAC060* at the point where ABA and *DOG1* diverge.

Results

DOG6 encodes the Arabidopsis NAC transcription factor, *ANAC060*

DOG6 has been identified as a major dormancy locus in five recombinant inbred populations that were made by crossing Landsberg *erecta* (L*er*) with other accessions. The presence of the QTL was verified using near isogenic lines (NILs) that contain introgression fragments of the Cape Verde Islands (Cvi), Santa Maria da Feira-0 (Fei-0), Kashmir-2 (Kas-2), Kondara (Kond) and Shakdara (Sha) accessions at the position of the *DOG6* locus in the L*er* genetic background. All these NILs conferred a more dormant phenotype than L*er* itself (Bentsink *et al.*, 2010).

To understand the molecular function of *DOG6*, the gene was cloned using positional cloning. Recombinants selected from crosses between L*er* and NIL*DOG6*-Sha and L*er* and NIL*DOG6*-Kas-2 positioned the *DOG6* locus between the markers T10D17-159697 and LK2 snp1 on chromosome 3 (Fig. 1, Supplemental table 1). This region of 84.5 Kb contained 15 ORFs (Fig. 1E). Dormancy phenotyping by means of after-ripening analyses (Soppe *et al.*, 2020) on knock-out mutants of these genes located in the region identified *AT3G44290* as the most likely candidate to encode for *DOG6* (Supplemental figure 1, Supplemental table 2). Two lines, SALK_127838C and SALK_012554C, with an insertion in *AT3G44290* had an increased seed dormancy level when compared to wild type Columbia-0 (Col-0) (Fig. 1F, 1G). The expression of *ANAC060* was confirmed to be significantly much lower in *anac060* mutants than in Col-0 (Fig. 1H). *AT3G44290* encodes the membrane bound transcription factor *ANAC060*. Complementation cloning performed by cloning the *AT3G44290* L*er* allele into NIL*DOG6*-Kas-2 using binary vector pKGW (pKGW-*DOG6*-L*er*) restored the reduced dormancy phenotype of L*er* (Fig. 1I), which confirmed that *DOG6* encodes the *ANAC060* transcription factor. In the following text we will refer to *ANAC060* since this is the official annotation of the gene.

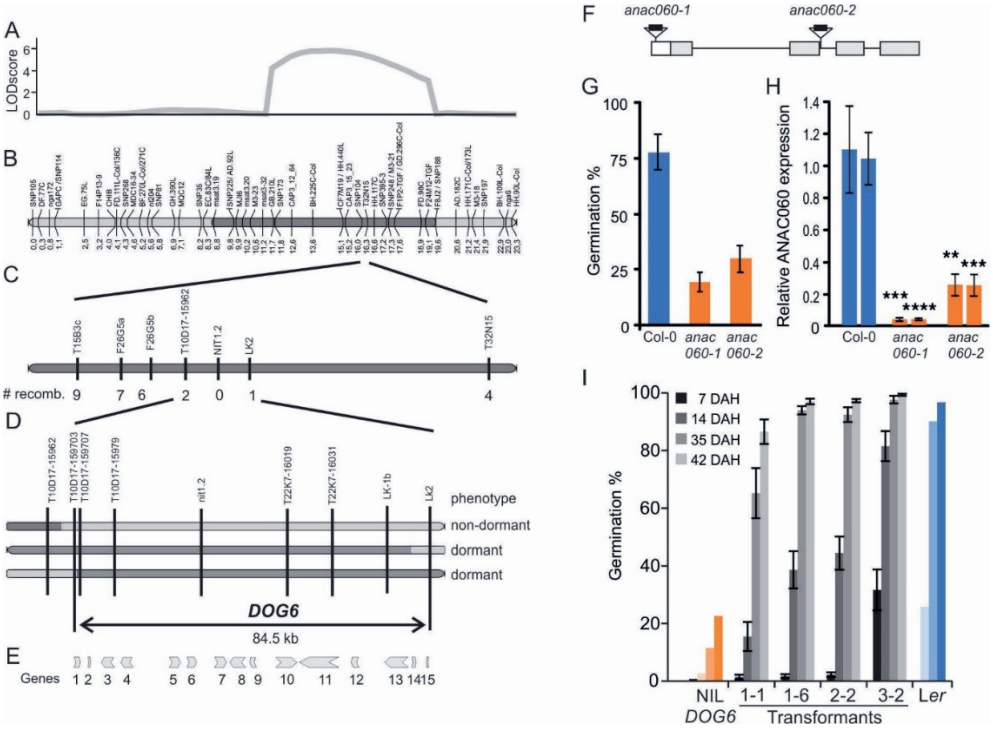


Figure 1. Genetic fine mapping of *DELAY OF GERMINATION 6* and the identification of the underlying gene. A. Position of the *DOG6* QTL in the genetic map of the L*er*/Kas-2 RIL population. B. Schematic representation of chromosome 3 of a near isogenic line that contains the Kas-2 introgression (dark grey bars) in the L*er* genetic background (pale grey bars). Junctions between pale and dark grey bars indicate crossover breakpoints. C. Zoom into the QTL region. D. Overview of the recombinants that are essential for determining the position of *DOG6*, at the right the dormancy phenotype of the recombinants is indicated. E. Genes that are located in the 84.5 kb fine-mapped region. As list containing the AGI numbers of the presented genes can be found in Supplemental table 2. F. schematic representation of the *ANAC060* gene model. 5'UTR in white, exons light grey and the T-DNA insertions are indicated by the black boxes. G. Germination phenotypes of freshly harvested seeds of two T-DNA knock-down lines that are disrupted in the *ANAC060* gene in comparison with wild type (Col-0). Averages and SE are presented, n=9. H. Relative *ANAC060* expression in Col-0 and the two knock-down lines. The expression values were normalized using two reference genes (*AT2G28390* and *AT4G12590*). Averages and SE are presented, n=9. I. Complementation cloning of *DOG6*/*ANAC060*. Germination behaviour of NIL*DOG6*-Kas-2, four independent transformants that contain the L*er* allele of *ANAC060* in the NIL*DOG6*-Kas-2 genetic background and L*er*. Asterisks indicate the significance determined by two-way ANOVA * P ≤ 0.05, ** P ≤ 0.01, *** P ≤ 0.001 and **** P ≤ 0.0001.

ANAC060 belongs to OsNAC8 subgroup of group I NAC transcription factors that includes *ANAC089* and *ANAC040* (Ooka *et al.*, 2003). *ANAC060* and *ANAC089* have earlier been identified to be important for sugar sensitivity in seedlings (Li *et al.*, 2011; Li *et al.*, 2014). We investigated the dormancy levels in down-out mutants of these homologs and revealed that the *anac089* mutants is slightly less dormant than its wild type, whereas *anac040* mutants do not show a dormancy phenotype (Supplemental figure 2). The group I NAC transcription factors are characterised by a N terminal NAC domain and a C-terminal alpha helix, which represents a transmembrane domain (TMD) that is known to retain the transcription factors to the cytoplasm (Li *et al.*, 2011; Li *et al.*, 2014).

Genetic variation for *ANAC060* determines its cellular localisation and thereby seed dormancy levels

Membrane bound transcription factors are often synthesized and stored in an inactive state (bound to the membrane) and become active (release to the nucleus) only in response to the appropriate environmental signal. This allows these transcription factors to quickly respond the changed environment. Sequence analyses of the different *ANAC060* alleles (Fei-0, Kas-2, Kond, Calver (Cal), Sha, Cvi, L*er*, Col-0, Antwerp-1 (An-1), Tacoma (Tac)) revealed several polymorphisms (Fig. 2A, Supplemental table 3). The major polymorphism is a 482 bp deletion in L*er* that because of this lacks the complete fourth exon which encodes the TMD. Col-0 and An-1 embrace a SNP in the third intron, which resulted in a 20 bp extension of the fourth exon that includes an in frame TAA stop codon. This early stop codon leads to the lack of the fourth exon that contains the TMD as was earlier reported for Col-0 (Li *et al.*, 2014). Five other sequence nucleotide polymorphisms (SNPs) caused amino acids changes. Fei-0 contains three SNPs in the second and third exon, namely Fei-1, Fei-2 and Fei-3 respectively. Fei-1 is located in the second exon, which is at the end of NAC domain. Fei-2 and Fei-3 are in the third exon. Tac and Cvi contain SNPs in the fourth exon, which are both located in TMD. The SNP in Tac resulted in early stop codon while the SNP in Cvi caused an amino acid change.

To investigate whether these polymorphisms affect the cellular localization of the ANAC060 protein, both transient and stable transformants were generated. *ANAC060* cDNA was fused with the green fluorescent protein (GFP) at the N-terminus with the 35S promoter for transient expression (p35S: GFP: *ANAC060*-L*er* and p35S: GFP: *ANAC060*-Kas-2) and the native Col-0 *ANAC060* promoter for stable transformed seeds (p*ANAC060*-Col-0: GFP: *ANAC060*-Ler and p*ANAC060*-Col-0: GFP: *ANAC060*-Kas-2) (Fig. 2, Supplemental figure 3). ANAC060-Kas-2 was located in membrane, including the ER, cell and nuclear membrane (Fig. 2B and 2C), whereas ANAC060-Ler was solely present in the nucleus (Fig. 2B and 2C). Based these and earlier findings we conclude that *ANAC060* is a transcription factor that results in a non-dormant phenotype when the protein is localized in the nucleus, a functional TMD retains the protein from the membrane which results in a higher seed dormancy. To test this hypothesis, we also investigated the localization of the other alleles (i.e., the *ANAC060* Col-0, Tac, Cvi and An-1 alleles), either transiently or by stable transformants using the At2S3 seed storage protein promoter (Kroj *et al.*, 2003). These analyses confirmed the relation between a missing or mutated TMD, localization in the nucleus and a non-dormant phenotype (Supplemental figure 3). Membrane bound transcription factors are often released from the membrane by cleavage of the TMD by membrane-associated proteases, a process referred to as intramembrane proteolysis (RIP) (Seo *et al.*, 2008). We hypothesized that dormancy breaking cues, like after-ripening, stratification or GA treatment would trigger RIP, however we were not able to identify any GFP signal in the nucleus for any of the ANAC060 GFP tagged dormant forms (data not shown). This might be caused by the fact that RIP does not occur as well as by the instability of the processed ANAC060 protein in the nucleus.

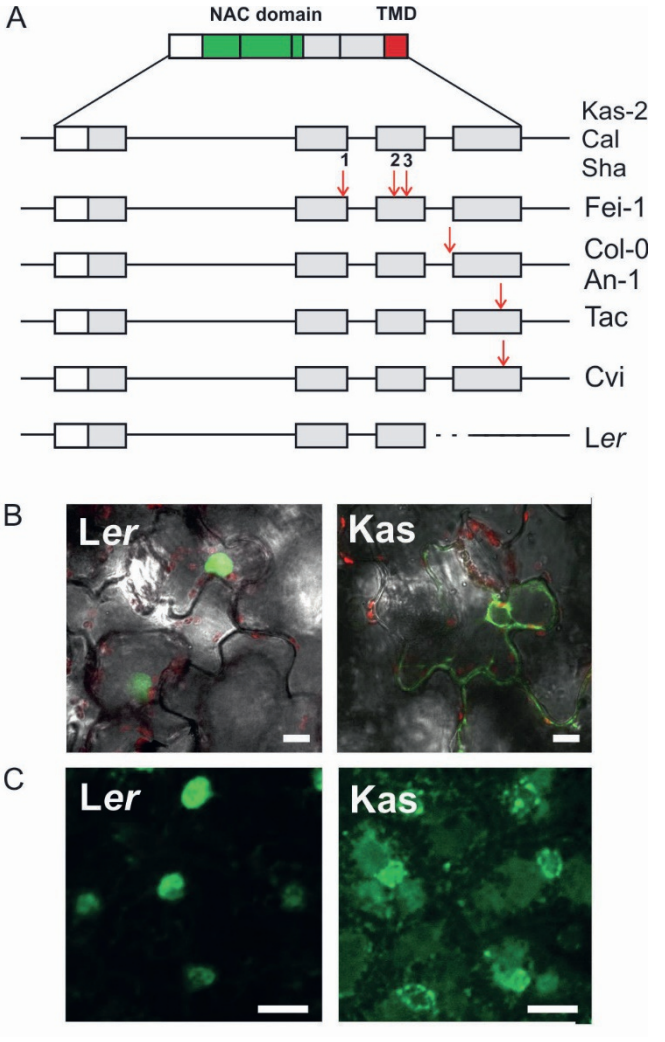


Figure 2. Sequence polymorphisms and cellular localisation of the ANAC060 protein. A. Schematic representation of *ANAC060*. The upper panel shows the *ANAC060* cDNA with the 5'UTR (open bar), the NAC domain (green bar) and the membrane binding domain (red bar) indicated. The vertical lines indicate the exon borders. In the lower panel the polymorphisms (red arrows; nature of the polymorphisms is indicated in supplemental table 1) of the different alleles are represented in the genomic sequence. B. Cellular localisation of the *ANAC060* L*er* and the *ANAC060* Kas-2 allele in *Nicotiana benthamiana* leaves. Fluorescence and bright-field merge of GFP fluorescence of leave cells expressing the p35S: GFP: *ANAC060* L*er* and p35S: GFP: *ANAC060*-Kas-2) fusion proteins. C. Cellular localisation of the *ANAC060* L*er* and *ANAC060* Kas-2 allele in seeds of the *anac060-1* knock-out mutant 18 days after pollination.

Fluorescence image of GFP fluorescence in seeds expressing the pAt2S3: GFP: *ANAC060*-L*er* and pAt2S3: GFP: *ANAC060*-Kas-2 fusion proteins. Scale bars, 10 µm.

To better understand the evolutionary pressure on the SNP for *ANAC060* the rareness of these SNPs was investigated in 94 accessions that were randomly selected from a set of 349 accessions of Arabidopsis HapMap population (Baxter *et al.*, 2010). The analyses with CAPS (Cleaved Amplified Polymorphic Sequences) markers that were designed to identify the SNPs revealed that the Fei-1, Tac and Cvi mutations are rare (Supplemental table 3). The Fei-2 and Fei-3 mutations are closely linked and remarkedly three genotypes were heterozygous for the L*er* mutation.

The Columbia allele of *ANAC060* is alternatively spliced in seeds

The expression of *ANAC060* under standard growth conditions is seed specific. The first transcripts are being detected in siliques from 8 days after pollination (DAP) onwards, expression peaks in dry mature seeds, remains high in dormant imbibed seeds and reduces in germinating seeds upon longer imbibition (Fig. 3A, 3B). As mentioned before, we confirmed the presence of a SNP in the third intron of the Col-0 allele that was identified before by Li *et al.* (2014). This SNP resulted in a protein lacking the TMD in seedlings grown on high sugar. However, we revealed two forms of the cDNA, one containing the extra 20 bp that results in a loss of the TMD in the protein and the other lacking the 20 bp resulting in the full-length protein containing the TMD. Based on this we hypothesised that in seeds *ANAC060* is alternatively spliced (Fig. 3B). The abundance of the two cDNAs was investigated to identify whether the different forms were related to the developmental states and/or treatments that would affect the dormancy status of the seeds, however no significant differences in mRNA abundance was detected in Col-0 (Fig. 3B). The low dormancy of Col-0 is explained by the fact that the short form which lacks the TMD is always present. The dominance of the short form is confirmed by the dominance of the *ANAC060* L*er* allele in reciprocal crosses between NIL*DOG6* and L*er* (Fig. 3C).

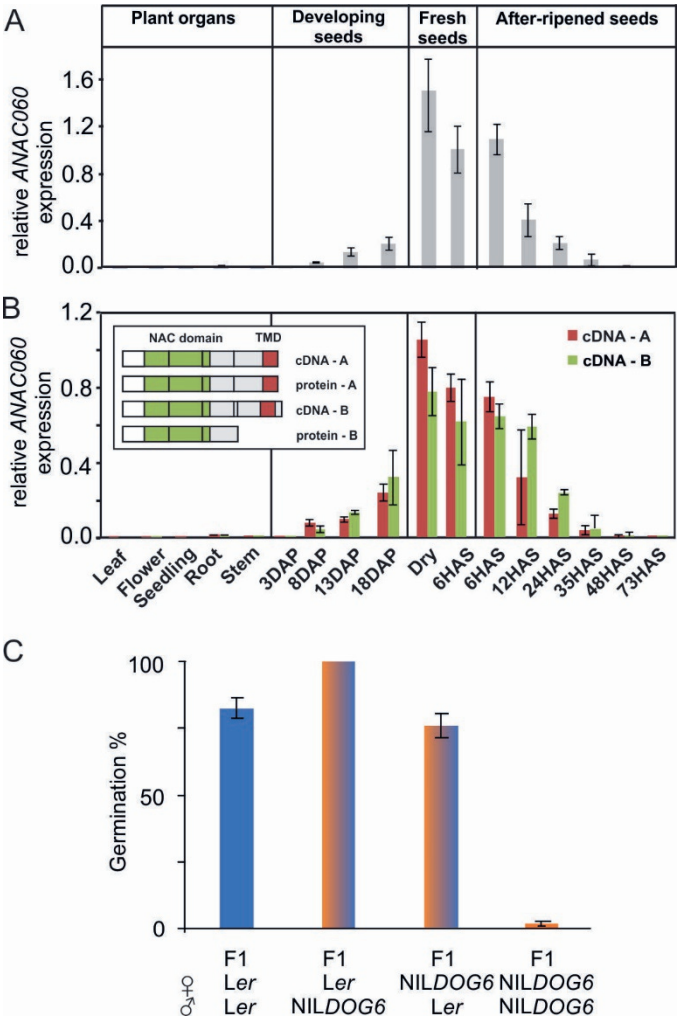


Figure 3. *ANAC060* alternative splicing and its germination characteristic. A. Relative *ANAC060* expression in plant organs including leaf, flower, seedling, root, stem, developing seeds, fresh seeds, after-ripened seeds. B. Two splicing forms of *ANAC060* is detected in different plant organs, leaf, flower, seedling, root, stem, siliques at 3, 8, 13, 18 days after pollination (DAP), freshly harvested seeds dry and 6 hours after imbibition (HAS) and after-ripened dry seeds; 6h, 12h, 24h, 35h, 48h, 73h: 6h, 12h, 24h, 35h, 48h, 73 HAS. Schematic presentation of the cDNAs and proteins of the two *ANAC060*-Col splicing forms is present at left top, cDNA – A refers to the annotated cDNA for which the protein contains the membrane binding domain, cDNA – B refers to the cDNA that contains 20 bp extra at the end of intron three, which results in a pre-mature stop codon and therefor a protein without the membrane binding domain. All the expression analyses have been performed on Col-0. The expression values were

normalized using two reference genes (*AT2G28390* and *AT4G12590*) and expressed as relative values of the freshly harvested dry seeds values. n = 3 biological replicates; error bars represent SE. C. Germination behavior of L*er*, NIL*DOG6*-Cvi and their reciprocal crosses at 68 days after seed harvest.

PROTEIN PHOSPHATASE 2C class A proteins are putative targets of *ANAC060*

ANAC060 was described to attenuate ABA signalling in seedlings when grown on high sugar (Li *et al.*, 2014) and recently it was revealed that *ANAC060* directly binds to the promoter of *ABI5* and represses the sugar-induced transcription of *ABI5* (Yu *et al.*, 2020). ChIP-Seq experiments expressing ANAC060 fused with GFP under the constitutive 35S promoter identified 500 direct targets of *ANAC060* in seedlings treated with glucose. A similar analysis in control conditions (1% sucrose) revealed 5091 putative targets (41.7% of these genes overlapped with the targets identified in the high fructose conditions). To identify targets of *ANAC060* in the regulation of seed dormancy we compared these putative targets to published gene expression data of dry and 24 imbibed dormant and fully after-ripened seeds (Bentsink *et al.*, 2010; Yazdanpanah *et al.*, 2017). We specifically addressed differences in expression between L*er* (ANAC060 present in the nucleus and thus expected to be active as a transcription factor) and NILDOG6-Kas (ANAC60 retained from the nucleus) seeds. We identify 108 genes as putative targets of *ANAC060* in the regulation of seed dormancy (Fig. 4A), of which only three genes overlap with the 500 direct targets identified in seedlings treated with glucose. This suggests that *ANAC060* regulates different genes when inhibiting seed dormancy in comparison to growth inhibition in high glucose. The three genes common between the two data sets are *AT2G27300* encoding *ANAC040* (*NTM1-like 8*), *AT3G23030* encoding *INDOLE-3-ACETIC ACID INDUCIBLE 2* (*IAA2*) and *AT4G18650* encoding *DELAY OF GERMINATION 1-LIKE 4* (*DOGL4*). *ANAC040* is a close homologue of *ANAC060* and to investigate possible redundancy and their genetic relation we have investigated this double mutant for seed dormancy behaviour. The *anac040* mutant does not show a dormancy phenotype that is different from wild type and the double mutant phenotype is indistinguishable from that of the *acac060* single mutant, suggesting that *ANAC060* is epistatic over *ANAC040* (Fig. 4B). DOGL4 belongs to

the DOGL family proteins but only shares a limited homology with DOG1 at the protein level. DOGL4 is induced by ABA and plays a role in mediating reserve accumulation in seeds. DOGL4 does not affect seed germination (Sall *et al.*, 2019). Overall, the 106 genes are rather diverse and no specific gene ontology terms have been identified as significant (Supplemental table 4).

Not surprisingly there are also ABA related genes among the putative targets, among which ABI5 BINDING PROTEIN 3 (ABF3) and PROTEIN PHOSPHATASE 2CA (PP2CA, AT3G11410). Seeds with mutations in either of these two genes are hypersensitive to ABA, suggesting a role for *ANAC060* in overcoming ABA sensitivity which is in agreement with the earlier reported findings (Kuhn *et al.*, 2006; Li *et al.*, 2014; Lopez-Molina *et al.*, 2003; Yoshida *et al.*, 2006). To further investigate the role of ABA in the regulation of seed dormancy by *ANAC060* we analysed seed dormancy in double mutants that contained the dormant Cvi allele of *ANAC060* in the *abscisic acid insensitive* 3-5 (*abi3-5*) and ABA deficient *aba1-1* genetic backgrounds. The seeds of both double mutants were completely non-dormant (Fig. 4C), confirming that the dormancy conferred by the lack of ANAC060 protein in the nucleus (*ANAC060*-Cvi allele, Supplemental figure 3) is ABA dependent. Moreover, the *ANAC060* knock-down mutants (*anac060-1* and *anac060-2*) are more sensitive to ABA when compared to their wild type Col-0 seeds (Fig. 4D).

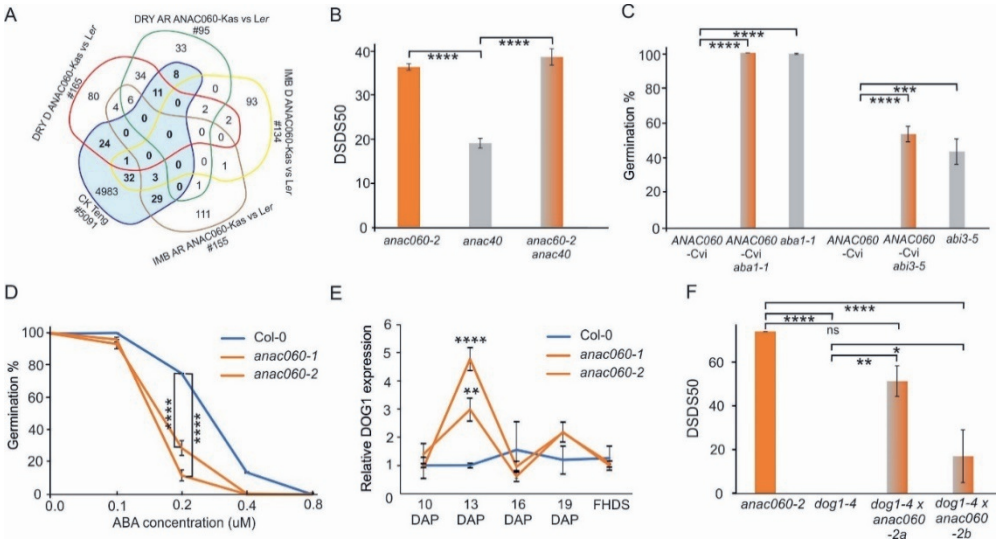


Figure 4. *ANAC060* putative targets. A. Venn diagram presenting the putative targets of *ANAC060*. Genes identified to directly bind to *ANAC060* in seedlings (CK Teng, (Yu *et al.*, 2020) and those that are differentially expressed in seeds that either contain the ANAC060 protein in the nucleus or in the cytoplasm (ANAC060-L*er* compared to ANAC060-Cvi) in four different physiological stages (dry dormant seeds; DRY D, dry after ripened seeds; DRY AR, imbibed dormant seeds; IMB D and imbibed after ripened seeds; IMB AR) have been compared. In bold the 108 putative *ANAC060* targets, the gene list is provided in Supplemental table 4. B. Dormancy levels expressed as DSDS50 (Days of Seed Dry Storage required to reach 50% of germination) for *anac060-2*, *anac040* and the *anac060-2 anac040* double mutants. C. *ANAC060* regulated dormancy depends on ABA. Germination percentage of freshly harvested seeds of NIL*DOG6*-Cvi (*ANAC060*-Cvi-*2*), *aba1-1* and the NIL*DOG6*-Cvi (*ANAC060*-Cvi-*2*), *aba1-1* double mutant and NIL*DOG6*-Cvi (*ANAC060*-Cvi-*2*), *abi3-5* and the NIL*DOG6*-Cvi (*ANAC060*-Cvi-*2*), *abi3-5* double mutant. D. *ANAC060* knock-down mutants are hypersensitive to ABA. Germination behaviour of Col-0 and *anac060-1* and *anac060-2* in different concentration of ABA. E. Relative *DOG1* expression during seed maturation in Col-0 and *anac060-1* and *anac060-2*. RNA was extracted at 10, 13, 16 and 19 days after pollination (DAP) and of freshly harvested dry seeds (FHDS). The expression values were normalized using two reference genes (*AT2G28390* and *AT4G12590*). Averages and SE of four biological replicates are presented. F. *ANAC060* and *DOG1* regulate seed dormancy by independent pathways. Dormancy levels expressed as DSDS50 (Days of Seed Dry Storage required to reach 50% of germination) for *anac060-2*, *dog1-4* and two independent double mutants. Averages and SE of three biological replicates are presented. Asterisks indicate the significance determined by one-way ANOVA in B, C, F, two-way ANOVA in D, E. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$ and **** $P \leq 0.0001$, ns is not significant.

ANAC060 and *DOG1* regulate dormancy by additive pathways

DOG1 is an important regulator of seed dormancy and although *DOG1* was not identified as a direct target of *ANAC060* we aimed at obtaining more insight into the relation between these dormancy regulators. RT-qPCR analyses during seed maturation revealed that the expression of *DOG1* is increased in the *anac060* mutant seeds at 13 days after pollination (DAP) suggesting that *ANAC060* is required for the repression of *DOG1* expression (Fig. 4E). In mature dry *anac060* seeds no differences in *DOG1* expression has been revealed when compared to Col-0 (Supplemental figure 4). Previously, *DOG1* and *DOG6/ANAC060* have been proposed to regulate seed dormancy by additive pathways (Bentsink *et al.*, 2010). To obtain more insight into the relation between these dormancy regulators the

anac060-2 mutant was crossed to the completely non-dormant *dog1-4* mutant (Bentsink *et al.*, 2006). The double mutants revealed a dormancy behaviour that was intermediate to that of the single mutants (Fig. 4F), which confirmed the additive effect that was reported earlier for the *DOG*NIL (NILDOG617), that contains the Cvi alleles of both *DOG1* and *DOG6*/*ANAC060*, compared to the single NILs (Alonso-Blanco *et al.*, 2003).

Discussion

The identification of genes underlying the *DOG* QTL has contributed to understanding how seed dormancy is regulated. *DOG1* was already cloned in 2006 (Bentsink *et al.*, 2006), however due to its lack of conserved domains only recently its mechanisms have been being revealed (Bryant *et al.*, 2019; Li *et al.*, 2019; Nakabayashi *et al.*, 2012; Nee *et al.*, 2017). With the cloning of *DOG6* the first transcription factor (*ANAC060*) underlying a *DOG* locus has been identified.

The presence of ANAC060 in the nucleus results in a non-dormant phenotype. Different genetic variants of *ANAC060* determine the localization of the protein. SNPs and deletions result in a loss of the TMD, directly or due to differential splicing. The nucleus localized ANAC060 protein, represented by its Tac and L*er* allele, is dominant. Col and An-1 alleles also result in a non-dormant phenotype due to insertion of a 22bp fragment of the 4th intron into the 4th exon which results in a stop codon just before the TMD. The genetic variants of *ANAC060* can be divided into two classes 1) structural variation already present in the genome sequence that determines whether the TMD is present or not, 2) alternative splicing, the different cDNAs in the Col and An-1 accession. The structural variants either represent point mutations that lead to early stop codons and because of that a loss of the TMD, as is the case for Tac, or a complete deletion of the TMD as is the case for the *ANAC060* L*er* allele. Also, for the *ANAC060* alleles that do contain the TMD there is genetic variation that might affect the localization of the protein in the cells (Supplemental figure 3).

ANAC060 has previously been identified as a fructose sensing QTL in a Col/C24 F2 population (Li *et al.*, 2014). The Col-0 *ANAC060* allele confers sugar insensitivity and was dominant over the sugar-sensitive C24 allele. The dominant nature of the Col allele is in agreement with our findings, however we do not reveal differences in *ABSCISIC ACID INSENSITIVE 4* (*ABI4*) expression between Col and the knock-down lines (Supplemental figure 4).

ANAC060 belongs to a family of three transcription factors, the other members (*ANAC040* and *ANAC089*) do not significantly affect seed germination or dormancy

(Fig. 4B, supplemental figure 2), but *ANAC040* has been identified as a putative *ANAC060* target involved in controlling both seed dormancy and sugar sensing. *ANAC040* lowers the GA threshold in response to salinity stress and thereby delays seed germination in these conditions (Kim, *et al.*, 2008). This suggests that both *ANAC040* and *ANAC060* provide adaptive mechanisms that delay seed germination in conditions that are not optimal for seedling establishment and further plant development. For *ANAC040* in vitro analyses suggest a role for paclobutrazol and cold in the processing of the protein, leading to the loss of the TMD by which the protein ends up in the nucleus (Kim *et al.*, 2008). The effect of cold on the release of the functional NAC-TF from the plasma membrane has also been reported for *ANAC062* (Seo *et al.*, 2010; Seo *et al.*, 2010). Cold stratification overcomes *ANAC060* dormancy however, we could not find any prove for post-translational regulation be means of membrane-associated proteases that lead to loss of the TMD and nuclear localization of the ANAC060 protein (data not shown). We can however not exclude processing of the ANAC060 protein. For *ANAC089*, a close homolog of *ANAC060*, reducing conditions promote the nuclear migration (Klein *et al.*, 2012).

NAC proteins constitute one of the largest groups of plant-specific transcription factors and are known to play essential roles in various developmental processes. They are also important in plant responses to stresses such as drought, soil salinity, cold, and heat, which adversely affect growth. Transcriptional control under environmental stresses plays a major role in plant adaptation. The importance of transcriptional regulation in plant adaptation is supported by large number of QTL which are identified based on SNPs in the promoter of the underlying gene, often resulting in mis-expression of the quantitative trait gene (Alonso-Blanco *et al.*, 2009). Quantitative trait loci analyses for seed dormancy in Arabidopsis revealed eleven *DOG* loci, for three of them the underlying gene has been identified, *DOG*1 (Bentsink *et al.*, 2006), *DOG18*/*RDO5* (Xiang *et al.*, 2016) and *DOG6*/*ANAC060* (current work). QTL and transcriptome analyses revealed that these QTL regulate seeds dormancy by independent pathways (Bentsink *et al.*, 2010). This finding is confirmed by molecular genetic analyses and based on this data we propose a model (Fig. 5).

This model is based on the earlier reported finding that both ABA and DOG1 are essential regulators of seed dormancy (Nee *et al.*, 2017). We propose that *ANAC060* acts at the point at which the roles of ABA and *DOG*1 diverge. Where the DOG1 protein interacts with ABA-HYPERSENSITIVE GERMINATION 1 (AHG1) and AHG3, *ANAC060* might interact with other PP2CAs, including PP2CA, ABF3 and RDO5/DOG18. This is supported by the identification of PP2CA and ABF3 as putative targets of *ANAC060* and by the reduced expression of *RDO5*/*DOG18* in the *ANAC060*-Cvi background (Xiang *et al.*, 2016).

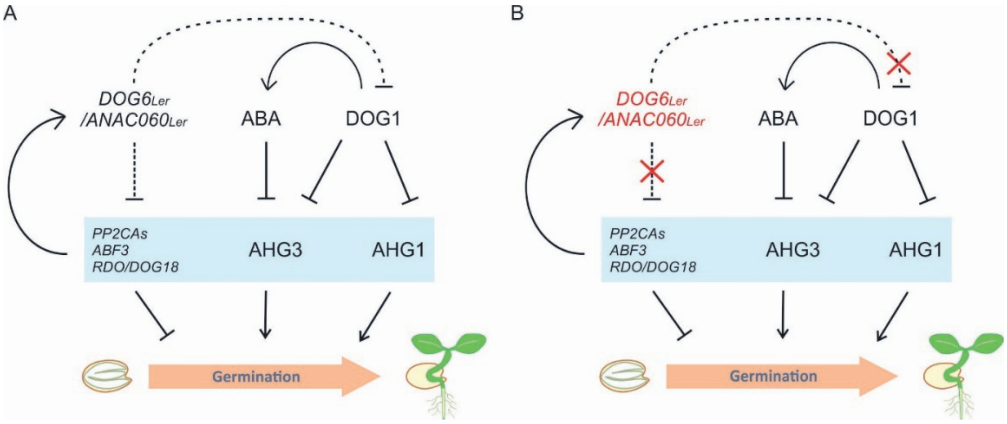


Figure 5. A model for the roles of *DOG6*/*ANAC060*, DOG1 and ABA in the control of seed dormancy. A. We propose that ABA promotes seed dormancy by inhibiting *DOG6*/*ANAC060* and the PP2CA class proteins. The model is based on the active form of the ANAC060/DOG6 protein, in which the protein is present in the nucleus. *DOG6*/*ANAC060* and DOG1 affect seed dormancy by additive pathways, it has been shown earlier DOG1 inhibits the action of AHG1 and AHG3. The expression of PP2CA, ABF3, ABI4 and RDO5/DOG18 is repressed in *ANAC060*-L*er*, PP2CA and ABF3 have been identified as putative targets is *ANAC060*. RDO5/DOG18 has been proposed function opposite compared to other clade A PP2Cs and might bind target proteins to prevent dephosphorylation by other protein phosphatases. B. Presents the situation in which the *ANAC60* gene is knocked-out or when the protein is associated to the membrane outside of the nucleus. (The model is largely based on the model presented Nee *et al*. (2017) has been extended with data obtained from literature (Xiang *et al*., 2014) and the current study).

Materials and Methods

Plant materials

Arabidopsis thaliana accessions Fei, Kas-2, Kond, Sha, Cvi, L*er,* Col-0, and An-1 and the T-DNA insertion lines *anac060-1*; SALK_012554C (N665285); *anac060-2*; SALK_127838C (N655936); three *anac040 alleles* (N587226, N613218, N106660) and the T-DNA lines listed in Supplemental table 2 were obtained from the Arabidopsis stock centres Arabidopsis Biological Resource Centre (ABRC) and Nottingham Arabidopsis Stock Centre (NASC). The *anac89* mutant (*Gt19255*) in Landsberg *erecta* (Ler) genetic background was a gift from Dr Sheng Teng (The Chinese Academy of Sciences, Shanghai, China). The accessions Antwerp-1, Cape Verde Islands were described before (el-Lithy *et al.*, 2006) and Calver and Tacoma were a gift from Dr Kathleen Donohue, Duke University, NC, USA. The near isogenic lines (NILs) NIL*DOG6*-Cvi, NIL*DOG6*-Kas-2, NIL*DOG6*-Sha were described before (Bentsink *et al.*, 2010). The *abscisic acid insensitive 3-5 (abi3-5) and* abscisic acid biosynthesis mutant *abscisic acid 1-1* (*aba1-1)* mutants were gifts from Dr Maarten Koornneef (Laboratory of Genetics, Wageningen University, The Netherlands). *dog1-4* (N105944; SM_3.20808) has earlier been described in (Bentsink *et al.*, 2006).

Plant growth

The plants were grown in controlled conditions 20°C/18°C (day/night) under a 16h photoperiod and 70% relative humidity using Rockwool blocks (4x4 cm) watered with Hyponex with three replicates per genotype.

Seed germination experiments

Germination experiments were performed as described previously (Joosen *et al.*, 2010). In brief, two layers of blue germination paper were equilibrated with 48 ml demineralized water in plastic trays (15 × 21 cm). Six samples of approximately 50 to 150 seeds were spread on wetted papers using a mask to ensure accurate spacing. Piled up trays were wrapped in a closed transparent plastic bag. The experiment was carried out in a 22°C incubator under continuous light. Pictures were taken twice a day for a period of 6 d. Germination was scored using the Germinator

package (Joosen *et al.*, 2010) and DSDS50 was quantified with the model described in He *et al.* (2014).

DNA and RNA isolation from seeds

DNA isolation: 10 mg of seeds were homogenized and mixed with 520 µl extraction buffer (0.6M NaCl, 100mM Tris, 40mM EDTA, 4% Sarkosyl and 1% SDS). Heat the samples for 10 min at 65°C and put it on ice for 5 min. After centrifuging, add 600 phenol/chloroform/isoamylacohol (25:24:1) to supernatant and mix. Centrifuge again and add 0.6 volume isopropanol, then leave it for at least 10 min at room temperature. Wash the pellet with ethanol and dry the pellet.

RNA isolation: Total RNA was extracted according to the hot borate protocol described by Maia *et al.* (2011). In brief, 7–10 mg of seeds for each genotype were homogenized and mixed with 800 ml of extraction buffer containing dithiothreitol (DTT) and PVP40 which had been heated to 80°C. Proteinase K was added and incubated for 15 min at 42°C. After adding 2M KCl, the samples were incubated on ice for 30 min and centrifuged. Ice-cold 8M LiCl was added to the supernatant and the tubes were incubated overnight on ice. After centrifugation, the pellets were washed with ice-cold 2M LiCl and centrifuged for 10 min. The pellets were re-suspended in 80 ml of water. The samples were phenol chloroform extracted, DNase treated and further purified. RNA quality and concentration were assessed by agarose gel electrophoresis and Nanodrop ND1000 spectrophotometry.

Complementation of *DOG6*/*ANAC060*

Binary constructs were prepared using the Gateway Technology (Invitrogen). A genomic DNA fragment of 4.3 kb of *DOG*6-Kas was amplified by PCR using primers attB1 *DOG*6/*ANAC060* 2141-F (GGGG ACA AGT TTG TAC AAA AAA GCA GGC TTC GCTACAAGCCTACAACAATCAGACG) and attB2 *DOG6*/*ANAC060* 3'UTR- 2-R (GGGG AC CAC TTT GTA CAA GAA AGC TGG GTC TCC TCC GTT AGG TTC CGT GA) and Phusion® High-Fidelity DNA Polymerase (New England Biolabs, Massachusetts, USA). This PCR product was cloned into cloned into pDONR221, resulting in pENTR-*DOG*6/*ANAC060*-Ler. pKWG *DOG6*/*ANAC060*-Ler was

produced from the above-mentioned entry clone and the destination vector pKGW red seed vector (Jan Verver Molbio; Invitrogen Life Technologies) by LR reaction. pKWG *DOG6*/*ANAC060*-L*er* was introduced by electroporation into Agrobacterium tumefaciens strain GV3101, which was subsequently used to transform L*er* and NIL*DOG6*-Kas plants by floral dipping (Clough *et al.*, 1998). Independent homozygous single insertion lines were selected based on their red fluorescence.

Sequence analyses

ANAC060 genetic variation was assessed by performing sequence analyses on overlapping PCR products that cover the *ANAC060* genomic region for the accessions Fei, Kas-2, Kond, Sha, Cvi, L*er,* Col-0, An-1, Cal and Tac.

Construction of expression vectors and generation of transgenic plants

35S: GFP: *ANAC060* constructs: The cDNA of Col, An, Kas-2, Cvi, Tac and Ler was utilized to build up constructs. The cDNA of *ANAC060* was amplified with 3' RACE-PCR (SMARTer™ RACE cDNA Amplification Kit from Clonetech) and cloned into gateway vector pDONR201. Then the cDNA was cloned into destination vector pH35NsGG which contains 35S promoter and GFP at N-terminal of *ANAC060*.

At2S3: GFP: *ANAC060* constructs: The cDNA of Kas-2, Cvi, Tac and L*er* was utilized to build up constructs. Firstly, the 353bp length promoter of At2S3 was cloned into gateway vector pDONR-P4P1. Then the fragment of GFP was added to the vector via pDONR207-P1P2. Then the cDNA of *ANAC060* was cloned into destination vector pDONR-P2rP3.

RT-qPCR expression analyses

In the *anac60* knock down mutants: RNA was isolated from dry mature seeds as described above. cDNA was synthesized from 750ng RNA using the iScript cDNA Synthesis Kit (Bio-Rad: 170-8890) according to the manufacturer's protocol. cDNA was diluted 10 times with sterile milliQ water. For each sample 2.5µL cDNA, 5µL iQ SYBR green supermix (Bio-Rad: 172-5125) and 0.5µL primer mix (10µL work solution) were added and supplemented with water to 10µL. RT-qPCR was performed on a CFX connect (Bio-Rad).

In different plant tissues: *ANAC060* expression analysis was performed using different plant tissues, including vegetative tissues (leaf (7th rosette leaf), flower (just opened flower), seedling (two weeks old seedling), root (two weeks old root), stem, developing siliques (3 days after pollination (DAP), 8DAP, 13DAP and 18DAP), dry dormant seeds and dormant seeds 6 hours after imbibition and imbibed after-ripened seeds (6h, 12h, 24h, 35h, 48h, 73h after imbibition).

In order to determine the two different *ANAC060* splice variants two different sets of primers have been used. Primer pair L4 only detects A form, S1 only detects B form to detect both forms we have used primer pair LS3.

Primers details can be found in Supplemental table 5. For all analyses the expression was normalized by the expression of two reference genes that are stably expressed in dry seeds: *At4G12590* and *At4G34270* (Dekkers ***et al.***, 2012).

Supporting Information

Supporting information can be downloaded by scanning the QR code or from http://www.wageningenseedlab.nl/thesis/ssong/SI/

Supplemental Figure 1. Germination phenotypes of T-DNA knockout lines that are candidates for the gene underlying the QTL.

Supplemental Figure 2. Germination phenotypes of mutants in the *ANAC060* homologs.

Supplemental Figure 3. Cellular localization of the different *DOG6* alleles.

Supplemental Figure 4. Expression analyses in the *anac060* knock-down mutant.

Supplemental Table 1. Markers for fine-mapping *DOG6*.

Supplemental Table 2. T-DNA lines for genes in the *DOG6* fine-mapped region and for genes homologous to *DOG6*.

Supplemental Table 3. Primers for the CAPS markers to detect the *ANAC060* alleles.

Supplemental Table 4. Gene list of 108 putative *ANAC060* targets.

Supplemental Table 5. Primers for the RT-qPCR analyses.

Chapter 3

ANAC060 and *ANAC040* are functionally redundant in the inhibition of seed dormancy

S. Song, L. A.J. Willems, A. Jiao, T. Zhao, M. Eric Schranz, L. Bentsink

Abstract

The NAC family of transcription factors is involved in plant development and various biotic and abiotic stresses. The *Arabidopsis thaliana* DELAY OF GERMINATION 6 (ANAC060) protein is inhibiting seed dormancy and as such determines the timing of seed germination. *ANAC040* and *ANAC089* are the two closest homologs of *ANAC060* based on protein and nucleotide sequence similarity. These three genes are predicted to be membrane bound transcription factors (MTFs) containing a conserved NAC domain but divergent C-terminal regions. Mutations in *ANAC040* lead to higher seed germination under salt stress. A premature stop codon in the *ANAC089* Cvi allele results in seeds exhibiting insensitivity to high concentrations of fructose. Thus, these three which genes are highly homologous at the protein/gene level confer distinct functions, although all germination related. In order to investigate whether the differences in function are caused by either a differential spatial or temporal regulation as a result of differences in the promoter regions or due to differences in their coding region (CDS) we have performed swapping experiments in which the promoter regions and the CDS of the three different NAC transcription factors have been exchanged. We examined seed dormancy and salt and fructose sensitivity of transgenic swapping lines in mutant backgrounds. We show that there is functional redundancy between *ANAC060* and *ANAC040*, but not between *ANAC060* and *ANAC089*.

Introduction

In *Arabidopsis thaliana*, 105 genes are predicted to be NAC proteins (Ooka *et al.*, 2003). The name NAC is based on a composition of three transcription factors (TFs): *NAM* (*no apical meristem*), *Arabidopsis thaliana ACTIVATING FACTOR1, 2* (*ATAF1, 2*) and *CUC2* (*cup-shaped cotyledon*) (Aida *et al.*, 1997; Souer *et al.*, 1996). NAC genes are described to contain a highly conserved N-terminal DNA binding domain also known as the NAC domain. Genome wide analysis revealed that more than 10% of the NAC TFs in Arabidopsis contain an α-helical transmembrane motif within their further varying C-terminal domain. These predicted membrane associated domains determine the transcriptional activity and localization of the NAC proteins (Ernst *et al.*, 2004; Hu *et al.*, 2006; Kim *et al.*, 2007). NAC proteins are functionally diverse, they are not only involved in various developmental processes, including embryo and flower formation, organ separation, lateral root development and shoot apical meristem formation (Aida *et al.*, 1997; Sablowski *et al.*, 1998; Souer *et al.*, 1996; Takada *et al.*, 2001; Vroemen *et al.*, 2003; Weir *et al.*, 2004; Xie *et al.*, 2000), but also in biotic and abiotic stresses, for instance, virus defence, wounding and microorganism defence, cold temperature sensitivity, drought responsiveness and ABA sensitivity (Fujita *et al.*, 2004; Hegedus *et al.*, 2003; Ren *et al.*, 2000; Tran *et al.*, 2004; Xie *et al.*, 1999).

ANAC060, *ANAC040* and *ANAC089* belong to the same NAC subgroup according to their protein sequence similarity (Ooka *et al.*, 2003). All three proteins are predicted to be membrane bound transcription factors (MTFs) since they contain a transmembrane domain (TMD) (Kim *et al.*, 2007; Li *et al.*, 2010). It has been shown that the full length ANAC060 protein, that contains the TMD, is associated to the nuclear membrane, whereas the truncated form lacking the TMD is localized in nucleus (Li *et al.*, 2014)(Chapter 2). Similarly, the full length ANAC040 and ANAC089 proteins were mainly detected on plasma or endoplasmic reticulum membranes and their truncated forms without the TMD in the nucleus (Kim *et al.*, 2007; Yang *et al.*, 2014). For several Arabidopsis MTFs, it has been shown that the presence or absence of the TMD affects the plant phenotype (Kim *et al.*, 2010).

ANAC060 inhibits seed dormancy when the protein is present in the nucleus (Chapter 2). There are various natural variants for *ANAC060* and sequence differences among them determine the presence of the TMD in the encoded proteins. For instance, the Kashmir allele of *ANAC060* contains an intact TMD, this full-length protein is localized at the nuclear membrane and the seeds display a dormant phenotype. The Landsberg *erecta* (L*er*) allele of *ANAC060*, lacking the TMD is localized in the nucleus, seeds containing this allele show a non-dormant phenotype (Chapter 2). *ANAC060* was also described to affect sugar sensing. The Columbia (Col) allele of *ANAC060*, that encodes a truncated protein, lacking the TMD, resulted in seedlings that are less sensitive to growth inhibition by high concentrations of sugar. In this process a role for abscisic acid (ABA) signalling via *ABSCISIC ACID INSENSITIVE 4* (*ABI4*) has been identified (Li *et al.*, 2014).

ANAC040 is involved in several physiological processes. Overexpressing a truncated ANAC040 protein resulted in a severe growth reduction and late flowering, the expression of *FLOWERING LOCUS T* was dramatically repressed in these lines (Kim *et al.*, 2007). Besides this, in the presence of high concentrations of salt, *anac040* mutant seeds could germinate to higher levels than wild type and the germination of an over-expresser, containing the truncated protein lacking the TMD, was severely reduced (Kim *et al.*, 2008). Moreover, the gain-of-function mutant *anac040-1D* with increased expression of *ANAC040* negatively regulated trichome formation by directly triggering the expression of *TRIPTYCHON* (*TRY*) and *TRICHOMELESS1* (*TCL1*), two genes that repress the formation of trichomes. Notably, the observed similar trichome phenotype of the truncated and full length ANAC040 protein in transgenic lines imply that in both transgenics the ANAC040 protein ends up in the nucleus. However, how ANAC040 re-localizes to nucleus remains elusive (Schnittger *et al.*, 1998; Tian *et al.*, 2017; Wang *et al.*, 2007).

ANAC089 is elevated by ER-stress and the truncated form of *ANAC089* lacking the TMD activates programmed cell death, this activity is controlled by bZIP28 and bZIP60 which are two known MTFs playing crucial roles in regulating cell viability during plant ER stress (Yang *et al.*, 2014). Moreover, the localisation of *ANAC089* is

also determined by natural genetic variation, like it was described for *ANAC060*. The Cape Verde Islands (Cvi) allele of *ANAC089* suppresses fructose signalling, due to a premature stop codon this protein lacks the TMD resulting in a nuclear localisation. Whereas the L*er* allele, that contains the TMD, is localized in cytoplasm and sensitive to fructose (Li *et al.*, 2011).

To investigate possible redundancy between these highly homologous NAC transcription factors Tian *et al.* (2017) compared the trichome phenotype in rosette leaves of the *anac060 anac040* double mutants to their single mutants. No differences in trichome formation on rosette leaves was found, when comparing *anac040*, *anac060* and *anac060anac040*. However, the double mutant did show more branched trichomes on the stems, which suggested a possible functional redundancy. It is well known that the function of eukaryotic genes is determined by its distinct functional constituents, for example, the enhancer/silencer, the promoter region, the coding sequence (CDS). Among these, the promoter and the CDS are the two central functional domains. The promoter region regulates the timing and pattern of a gene's expression. The CDS encodes a gene's protein which is in charge of the phenotype (Polyak *et al.*, 2003).

Here, we have performed promoter and CDS swapping experiments to disclose the functional redundancy and distinction of functions of *ANAC060*, *ANAC040* and *ANAC089*. These experiments revealed that there is functional redundancy between *ANAC060* and *ANAC040*, the different phenotypes of the native genes are likely the result of their distinct expression patterns. We did not detect functional overlap between *ANAC060* and *ANAC089*.

Results

The evolutionary relationships of *ANAC060*, *ANAC040* and *ANAC089*

Protein sequence comparisons were performed to identify genes with high homology to *ANAC060* in Arabidopsis. The 15 most homologous genes were selected based on an earlier comparison of the protein sequences of *ANAC* transcription factors (Schwacke *et al.*, 2003). ANAC040 (41% identity) and ANAC089 (64% identity) were found to be the two most homologous proteins to ANAC060 based on sequence similarity (Pearcon *et al.*, 1988; Pearson, 2000) (Fig. 1A). The homology is especially high at the N terminal part of the protein that contains the NAC domains (NAM) from 21-146 AA (Fig. 1B). In order to investigate the evolutionary history of the three NAC genes, we extracted the whole synteny network of the NAC gene family from the entire synteny network database constructed from 107 plant genomes (Zhao *et al.*, 2019). Interestingly, *ANAC060*, *ANAC040* and *ANAC089* were located in the same synteny cluster which indicated a common genomic origin, shared with other eudicots but lacking monocots (Fig. 2A). The constructed phylogenetic tree of this cluster included many Brassicaceae sequences and sequences from *Cleome gynandra* (cgy) and *Tarenaya hassleriana* (tha) which both belong to Cleomaceae family, which is the sister lineage to Brassicaceae. Using this information, we propose the following duplication history and evolution of the three genes. *ANAC040* and *ANAC060/ANAC0890* represent duplicates derived from the older At-Beta whole genome duplication event and *ANAC060* and *ANAC0890* were duplicated from At-Alpha whole genome duplication event (shared only by Brassicaceae species) (Fig. 2B).

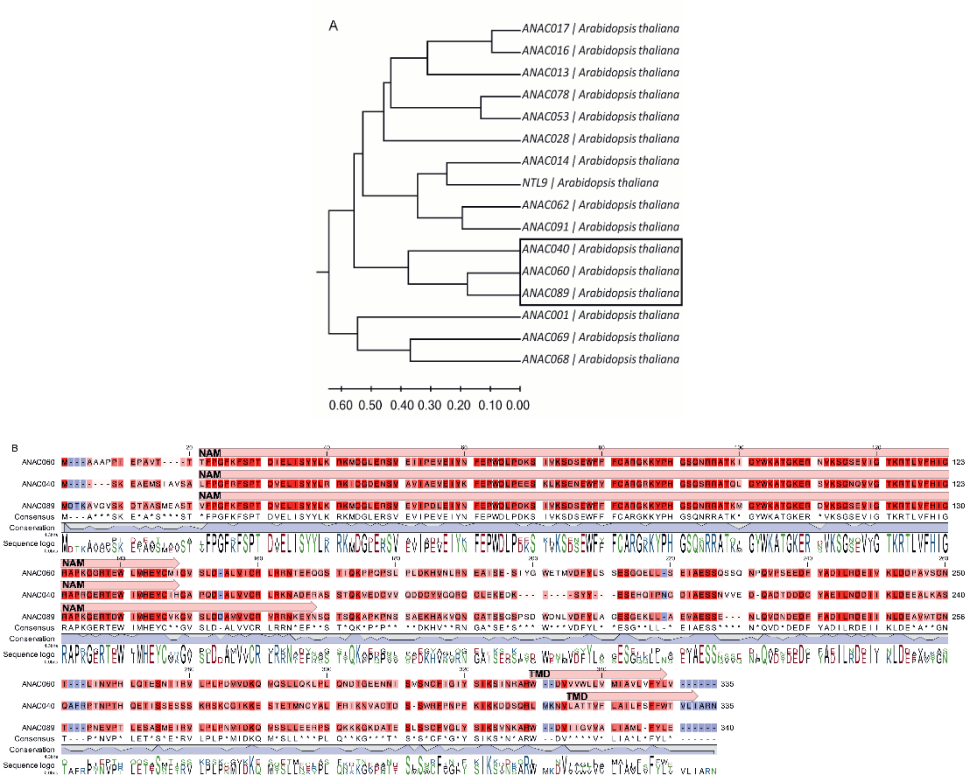


Figure 1. **Sequence comparisons of *ANAC060, ANAC040* and *ANAC089*.** A. The unrooted phylogenetic tree was built by using MEGA 7.0 (Kumar *et al*., 2016) and the UPGMA method (Sneath *et al*., 1973). The units of the branch lengths in the subgroup are comparable to the evolutionary distances used to interpret the phylogenetic tree. Poisson correction was used to compute the evolutionary distances (Zuckerkandl *et al*., 1965). The tree scale is shown below the tree. B. Protein sequences alignments for ANAC060, ANAC040 and ANAC089. Sequences are derived from the Columbia genome (TAIR). The NAC domain was predicted by plant transcription factor database and transmembrane domain was analysed by TMHMM Server v.2.0.

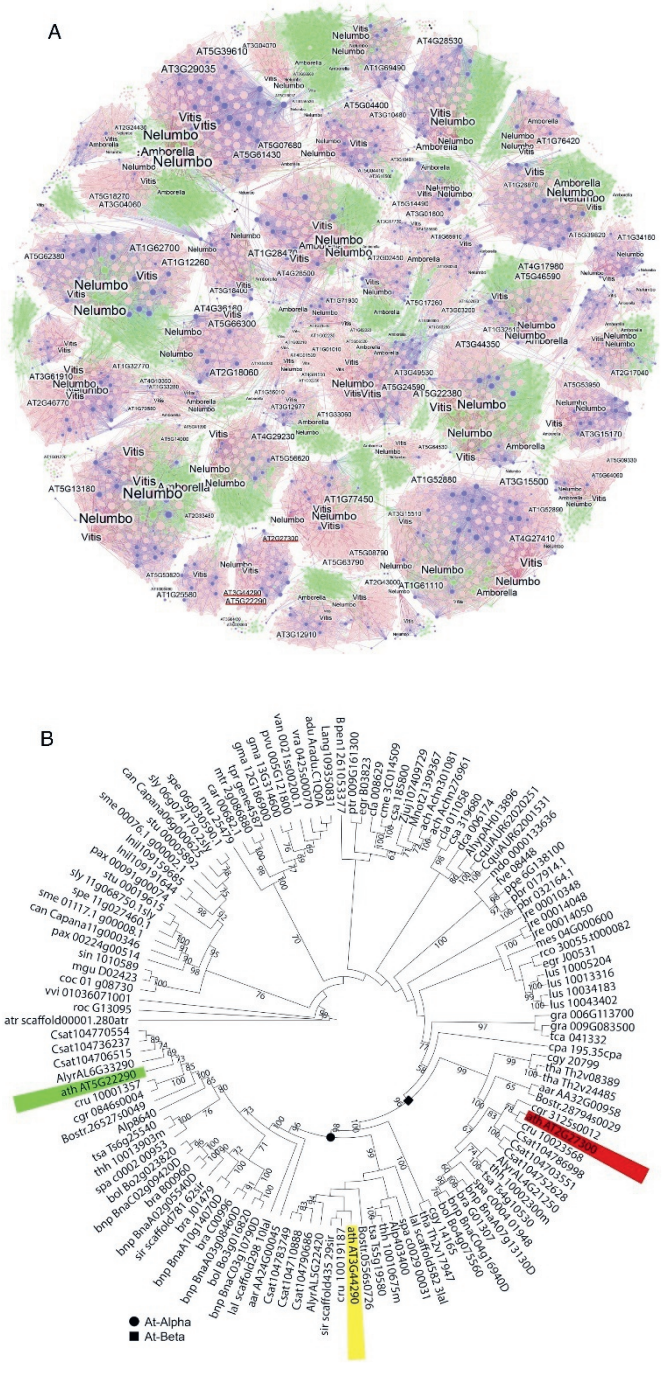

A
B
At-Alpha
At-Beta

Figure 2. Evolutionary analysis of NAC genes. A. Synteny network of NAC genes in 107 angiosperm genomes. The size of each node corresponds to the number of edges it has (node degree). Communities were labelled by the subfamilies/subfamily involved. The different colours in each cluster represent genes belonging to rosids (light pink), monocots (green), asterids (purple) (Zhao *et al.*, 2017; Zhao *et al.*, 2019). *ANAC060*, *ANAC040* and *ANAC089* were underlined in red. B. Phylogenetic analysis of *ANAC060*, *ANAC040* and *ANAC089*. *ANAC040* (*AT2G27300*, shaded in red) was retained from the older At-Beta whole genome duplication event. The species used to create this tree are listed Supplemental table 1. *ANAC060* (*AT3G44290*) and *ANAC089* (*AT5G22290*) are indicated in yellow and green boxes respectively. The black round and square nodes stand for At-Alpha and At-Beta polyploid events respectively. The bootstrap range displayed from 58 onwards.

To identify functional redundancy between *ANAC060* and its homologs *ANAC040* and *ANAC089,* the mutants were investigated for the phenotypes that have been reported earlier. Seeds of *anac060-1* (SALK_127838C), *anac060-2* (SALK_012554C) both showed deeper primary dormancy than wild type Col-0. The primary dormancy levels of *anac040* and *anac089* did not significantly differ from that of their respective wild types Col-0 and L*er* after one week after ripening (Supplemental figure 2, Chapter 2). The *anac060-2 anac040* double mutant phenotype did not significantly differ from the single *anac060* mutant, indicating that *ANAC060* is epistatic over *ANAC040* (Fig. 4B in Chapter 2). Due to the difference in genetic background between the available *anac060* and *anac089* mutants, Col-0 and L*er* respectively, we did not make the *anac060 anac089* double mutant. Phenotypic analyses of such a cross would be complicated due to genetic segregation of seed dormancy loci that are present in L*er* and Col-0 (van der Schaar *et al.*, 1997).

Anac040 mutant seeds were more resistant to salt than wild type Col-0 (Kim *et al.*, 2008). In order to confirm this phenotype and compare the sensitivity of *anac040* together with *anac060-1* in our laboratory conditions, seeds were sown on a series of salt concentrations. After three days, *anac040* seeds displayed a dramatically lower sensitivity to 150 mM of salt than Col-0, whereas *anac060-1* showed a salt sensitivity that was similar to Col-0 (Fig. 3C).

The fructose sensitivity of *anac089*, L*er* and Col-0 has earlier been investigated by Li *et al.* (2011). We confirm these results and show that *anac060-1* was also highly sensitive to fructose (Supplemental figure 1).

The *ANAC040* coding sequence rescues the *anac060* dormancy phenotype

To understand why high protein identity scores can result in distinct gene functions and to investigate whether the CDS could lead to functional redundancy, promoters and CDS swapping experiments were performed. For the swapping experiments, all promoters were cloned from Col-0. The active CDSs, thus the alleles that result in a lack of the TMD, have been selected and cloned from *ANAC060*, *ANAC40* and *ANAC089*. Seeds of *anac060-1*, Col-0 and transgenic lines containing the P*ANAC060*::Col-0_*ANAC040ΔC* (ΔC referrers to the allele that lacks the TMD), P*ANAC040*::L*er*_*ANAC060* and P*ANAC060*::L*er*_*ANAC060* in *anac060-1* background were examined for their dormancy level (measured as DSDS50; Days of Seed Dry Storage required to reach 50% of germination). The transgenic lines containing the P*ANAC060*::L*er*_*ANAC060* and P*ANAC060*::Col-0_*ANAC040ΔC* constructs both complemented the *anac060-1* dormancy phenotype. This was not the case for the recombinant lines expressing the P*ANAC040*::Ler_*ANAC060* construct (Fig. 3A).

To further investigate the redundancy between *ANAC060* and *ANAC040* the same constructs were transformed to the *anac40* mutant background. As described above, the Col-0_*ANAC040* contains a TMD and there is no dormancy difference between *anac040* and Col-0 in our standard dormancy testing conditions (22°C, continuous light). In order to be able to identify small differences in seed dormancy, germination experiments were also performed at 25°C. During seed storage (after-ripening) the germination-window widens, meaning that seeds can germinate better in less optimal conditions, e.g., higher temperatures (Alvarado *et al.*, 2002). This phenomenon allowed us to investigate whether the truncated ANAC040 protein under the *ANAC060* promoter could release seed dormancy in the *anac40* mutant background. There is no difference in dormancy between *anac040*, Col-0 and the complementation line in the mutant background. However, the transgenic line

P*ANAC060*::Col-0_*ANAC040ΔC* was less dormant than its background line *anac040*, showing again that ANAC040 when present in the nucleus can overcome seed dormancy (Fig. 3B).

Moreover, the single and double mutants and the transgenic seeds were also tested for salt sensitivity according to the study by Kim *et al.* (2008). *Anac040* and the double mutant *anac060-1 anac040* were resistant to 150mM salt, whereas Col-0 and *anac060-1* were not (Fig. 3C). Indicating that *ANAC040* is epistatic over *ANAC060* in regulating salt sensitivity. The complementation line P*ANAC040*::Col-0_*ANAC040ΔC* only partly complemented the mutant phenotype. Both of swapped transgenic lines P*ANAC060*::Col-0_*ANAC040ΔC* and P*ANAC040*::L*er*_*ANAC060* displayed significant slower germination rate than *anac040* under high concentration of salt (Fig. 3D).

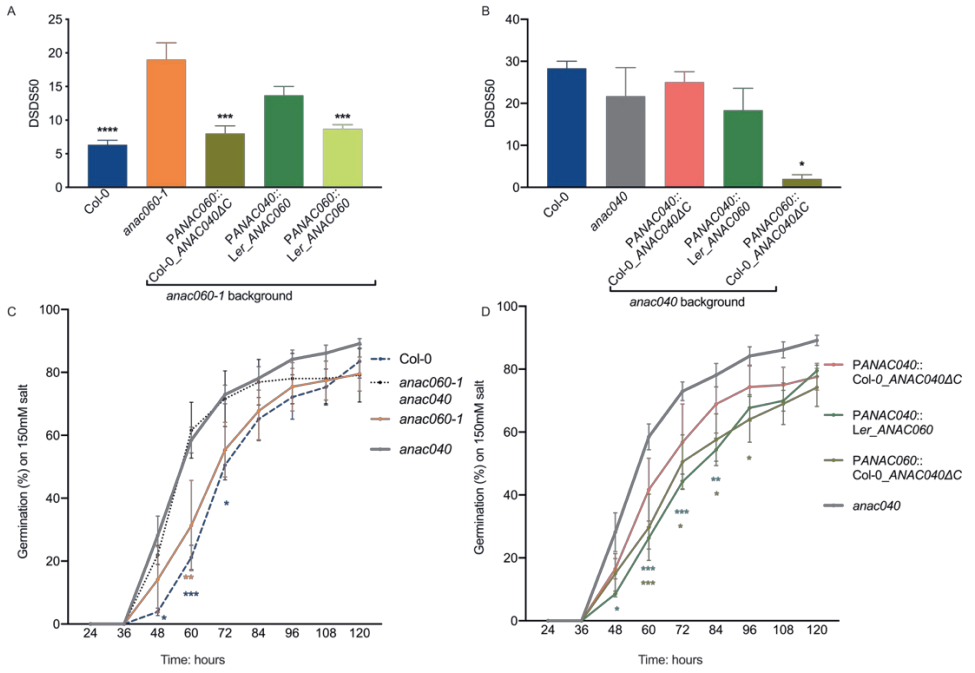


Figure 3. Dormancy and salt phenotypes of swapped transgenic lines. A, Dormancy phenotype of transgenic lines in the *anac060-1* genetic background. Seed dormancy is displayed as DSDS50 (days of seed dry storage required to reach 50% germination) levels. B, Germination percentages of transgenic

lines in the *anac040* background at 7 days after harvest in 25°C. C and D, Salt sensitivity of mutants and transgenic lines in the *anac040* genetic background. Bars indicate the mean value of three replicates and the standard errors. Statistical significances were calculated using one-way ANOVA in A, B and two-way ANOVA in C and D. Asterisks indicate significant differences with the respective backgrounds *anac060-1* or *anac040* (*: P ≤ 0.05, **: P ≤ 0.01 and ***: P ≤ 0.001).

The results described above show that there is functional redundancy between the *ANAC040* and *ANAC060* CDS, however there are distinct phenotypes in the different test conditions. For example, the complementation line containing the P*ANAC040*::Col-0_*ANAC040ΔC* construct did not overcome seed dormancy in the *anac040* mutant background, whereas the P*ANAC060*::Col-0_*ANAC040ΔC* construct did. Similarly, the P*ANAC040*::L*er*_*ANAC060* led to a full complementation whereas P*ANAC040*::Col-0_*ANAC040ΔC* only partially complemented the *anac040* mutant when germination was tested in salt. Expression analyses were performed to investigate if these phenotypes could be explained by the expression level of the transgenes, thus the activity of the promoters. First the expression in dry seeds was analysed. *ANAC060* expression is down-regulated in the *anac060* mutant dry seeds and shows a wild type expression in Col-0 and the *anac040* mutant (Fig. 4A). *ANAC040* is not expressed in dry seeds but is significantly induced under the *ANAC060* promoter in the *anac040* mutant background (Fig. 4B). As reported by Kim, Sang-Gyu *et al.* (2008), *ANAC040* is strongly induced in 3 days cold-imbibed seeds. We confirmed the promoter activity of P*ANAC040* after three days of cold stratification by investigating the *ANAC060* expression in the P*ANAC04*0::L*er*_*ANAC060* transgenic line (*anac060-1* background) (Fig. 4C). Also, *ANAC040* expression was induced in Col-0 and, although to a lower level, in the transgenic lines containing the *ANAC040* promoter (Fig. 4D). The partial complementation of germination in salt by the transgenic lines containing the *ANAC040* promoter (Fig. 3D) might be explained by the fact the *anac040* expression in the complementation line is lower than that in Col-0. Moreover, the *ANAC040* expression is significantly reduced in the *anac060* mutant background (Fig. 4D), this also partly rescued by the *ANAC040* complementation construct.

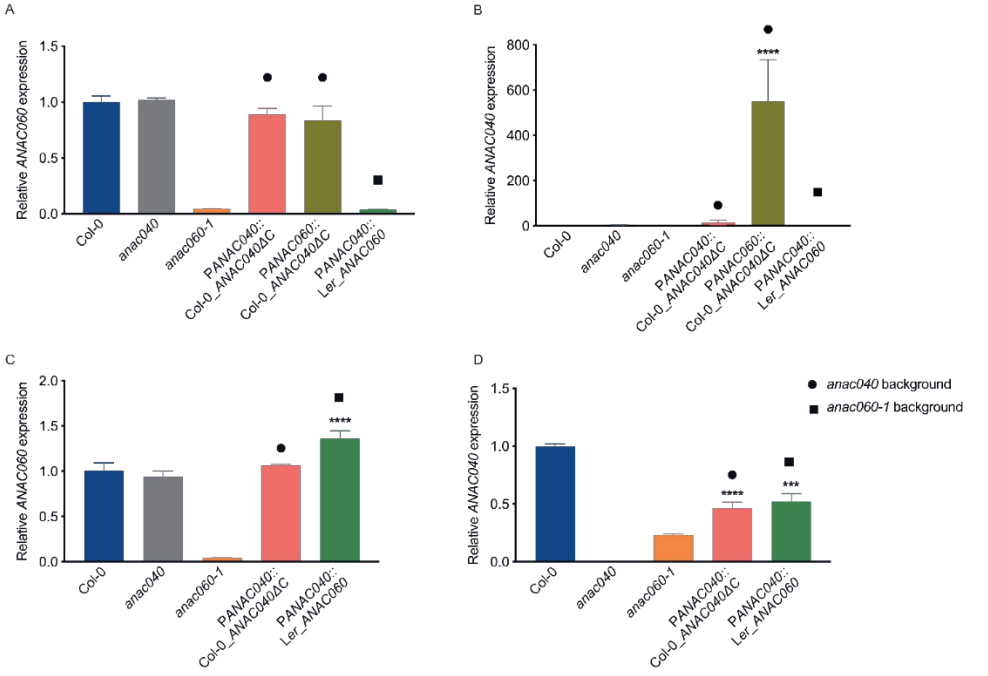


Figure 4. Expression of *ANAC060* and *ANAC040* in respective lines determined by quantitative RT-qPCR. A and B, Expression level of *ANAC060* and *ANAC040* in dry seeds. C and D, Expression level of *ANAC060* and *ANAC040* in three days cold-imbibed seeds. Col-0 as a control and its value was accommodated to be 1. Solid rounds and squares indicate the respective mutant backgrounds *anac040* and *anac060-1*. The expression level of all other line was relative to 1. Statistical significances were calculated using one-way ANOVA. Significances were compared between each transgenic line and its background mutant. Bars indicate the mean value of three replicates and the standard error. Asterisks indicate significant differences between the transgenic lines and their respective mutant backgrounds (***: $P \leq 0.001$, ****: $P \leq 0.0001$).

The redundancy between *ANAC060* and *ANAC040* is also assessed by double mutant analyses. The experiment performed in Chapter 2 shows that *ANAC060* is epistatic over *ANAC040* in regulating seed dormancy, which is probably explained by the lack of *ANAC040* expression in dry seeds (Fig. 4B). Moreover, the *anac060-2 anac040* double mutant shows a significant higher germination than single *anac060-1* mutants in the presence of 150 mM salt, which indicates that when the *ANAC040* is highly induced it would be epistatic over *ANAC060* (Fig. 3C). Moreover,

ANAC040 is a putative downstream target of *ANAC060*, this is supported by the reduced expression of *ANAC040* in *anac060-1* mutant seeds after three days of cold stratification (Fig. 4D) and the up-regulation of *ANAC040* when *ANAC060* is present in the nucleus (Chapter 2).

ANAC060 and *ANAC040* both regulate seed germination and although both are expressed in seeds their expression patterns are rather different. *ANAC060* is mainly induced during seed maturation and its expression remains high in dry seeds but significantly reduces during seed imbibition (Fig. 3B Chapter 2). *ANAC040* is induced during seed imbibition especially in response to cold (Fig. 4D) and it can overcome seed dormancy (Fig. 3B). Based on these findings we predict a role for *ANAC040* in the regulation of secondary dormancy. Secondary dormancy occurs in the imbibed stage when the conditions are not optimal for seeds to germinate (Baskin *et al.*, 1998; Buijs, 2020). To test this hypothesis, we made use of public gene expression data describing a dormancy cycling experiment that was performed by following seed germination behaviour and transcriptional changes during one year of seed burial (Buijs *et al.*, 2020). From March to May, buried seeds remained deep dormancy and re-gained germination capacity in June, after which the germination ability gradually increased to 100% in October. During winter seeds became dormant again resulting in a reduced germination capability (20%) in February (Fig. 5A) (Buijs *et al.*, 2020). The expression of *ANAC040*, *ANAC060* and *ANAC089* during dormancy cycling was investigated and compared to that of *DOG1*, a gene that has been reported to show differential expression during dormancy cycling (Footitt *et al.*, 2015; Murphey *et al.*, 2015). The higher relative expression of *ANAC040* in the non-dormant phase of dormancy cycling is in agreement with our earlier findings that *ANAC040* overcomes seed dormancy (Fig. 5B, Fig. 3B). Its peak in expression in October might be explained by the sharp drop in temperature, since we know that low temperatures induce *ANAC040* expression (Fig. 5C, Fig. 4D).

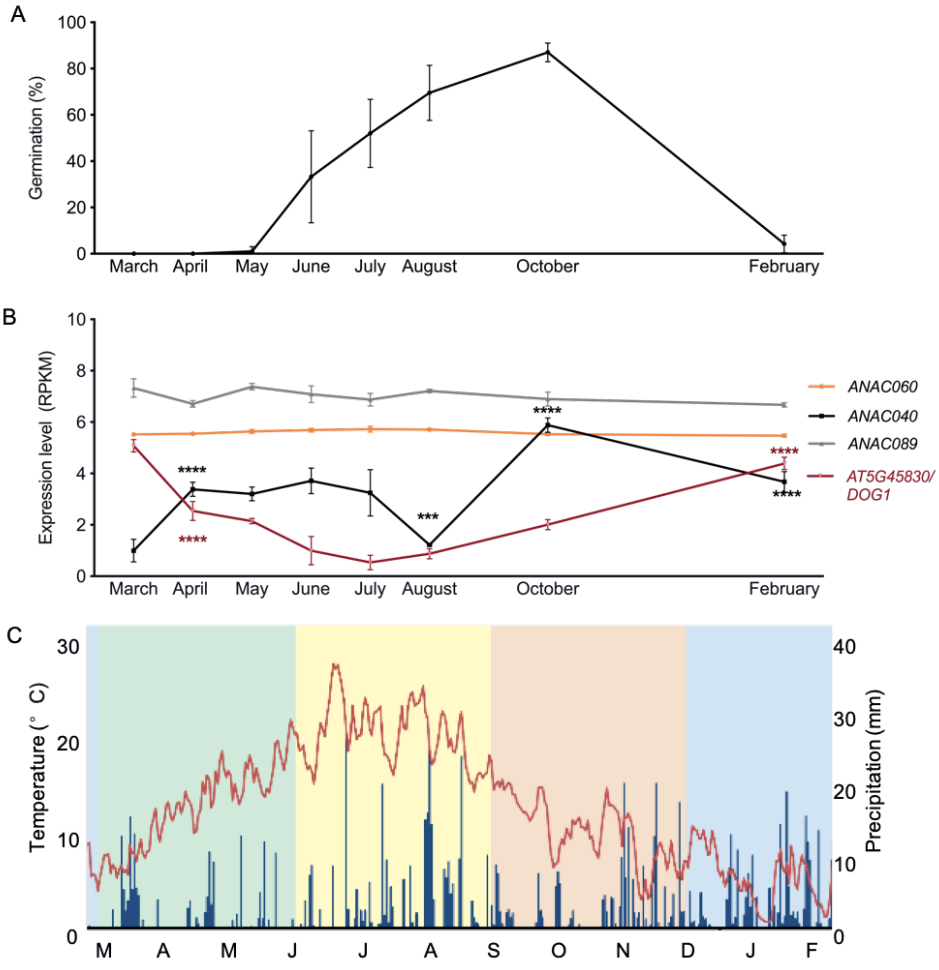


Figure 5. Expression levels of *ANAC060*, *ANAC040*, *ANAC089, DOG1* (*AT5G45830*) in dormancy cycling under soil. A, Germination percentage throughout the seasons. Landsberg *erecta* seeds were collected from field and germinated under lab conditions. B, Relative expression levels of *ANAC040*, *ANAC060* and *ANAC089* and *DELAY OF GERMINATION 1* (*DOG1*) during dormancy cycling. RPKM levels are shown. C, Temperature and moisture content at 5 cm depth in the soil. Red lines indicate field temperature obtained from buried sensors. The data presented in A, B and C is derived from Buijs *et al.* (2020). Two-way ANOVA was performed to analysis the data, every comparison was between adjacent previous month, asterisks indicate significant differences, ***: $P \leq 0.001$, ****: $P \leq 0.0001$. In A and B, the mean value of four replicates and their standard errors are presented.

Motif analyses in promotors of *ANAC060* and *ANAC040*

Motif analyses on the promoters of *ANAC060* (5249 bps) and *ANAC040* (1998 bps) were performed to determine whether these can explain the differential expression of *ANAC060* and *ANAC040* (Lescot *et al.*, 2002). In total 16 and 14 motifs were identified in the *ANAC060* and the *ANAC040* promoter respectively, twelve of these regulatory motifs were present in both promoters. Four motifs are specific for *ANAC060* and two are specifically for *ANAC040* (Table 1). Among these motifs are the ABA-responsive element (ABRE), MYB, G-box, W-box, GT1-motif which have been associated with regulation of Arabidopsis seed development (Belmonte *et al.*, 2013; Yamasaki *et al.*, 2017).

Table 1. Motifs identified in the promotors of *ANAC060* and *ANAC040*. The motifs of *ANAC089* promoter and a description of these motifs can be found in Supplemental table 2.

	Number	Elements	Sequence	Function
Motifs in P*ANAC040* and P*ANAC060*	12	MYB/MYB-like sequence/ Myb	TAACCA/ TAACTG	Responding to hormones during seed dormancy and gemination (Shinozaki *et al.*, 1992)
		activation sequence-1 (as-1)	TGACG	Salicylic acid (SA)- and auxin-responsive element responding to reactive oxygen species (Garretón *et al.*, 2002)
		AT~TATA-box TATA TATA-box	TATATA TATAAAAT TATA	Core promoter element around -30 of transcription start
		ABRE	TACGGTC	*Cis*-acting element involved in the abscisic acid, high salinity, drought responsiveness (Maruyama *et al.*, 2012)
		CAAT-box	CCAAT	Common *cis*-acting element in promoter and enhancer regions
		G-box	CACGTG	*Cis*-acting regulatory element involved in light responsiveness; Regulating seed specific expression (Ouwerkerk *et al.*, 1999)
		GT1-motif	GGTTAA	Light responsive element
		TCT-motif	TCTTAC	Part of a light responsive element
Motifs in P*ANAC060*	4	W box	TTGACC	Bing site for WRKY transcription factors (Rushton *et al.*, 2010)
		Ascorbate Peroxidase Binding Site (AP-1)	TGAGTTAG	Heat shock element (Pati, 2005)
		MYC	CATGTG	Binding site for BHLHZ family controlling cell behaviour (Grandori *et al.*, 2000)
		CAT-box	GCCACT	*Cis*-acting regulatory element related to meristem expression
Motifs in P*ANAC040*	2	STRE	AGGGG	Heat shock elements (Guo *et al.*, 2008)
		GA-motif	ATAGATAA	Part of a light responsive element

There is no functional redundancy between *ANAC060* and *ANAC089*

To test whether also *ANAC089* can rescue the dormancy phenotype of the *anac060-1* seed dormancy levels of *anac060-1*, Col-0 and transgenic plants containing the P*ANAC060*::Cvi_*ANAC089* and P*ANAC089*::L*er*_*ANAC060* constructs in *anac060-1* background were measured. None of these swapped transformants complemented the *anac060-1* dormancy phenotype (Fig. 6). Similarly, we tested whether *ANAC060* could complement the *anac089* fructose sensitivity. The transgenic seeds containing the P*ANAC060*::Cvi_*ANAC089*, P*ANAC089*::L*er*_*ANAC060* constructs in the *anac089* background together with the *anac089* mutants seeds and wild type L*er* were germinated on 6% fructose to assess the sugar sensitivity. The percentage of healthy seedlings was examined based on the presence of green cotyledons. Only the line containing the P*ANAC089*::Cvi_*ANAC089* construct complemented the fructose sensitivity, confirming earlier results of Li *et al.* (2011). None of the swapping transformants revealed resistance to fructose (Fig. 6B). Both experiments indicate that neither the promoter nor the CDS of *ANAC089* and *ANAC060* can replace that of the other gene.

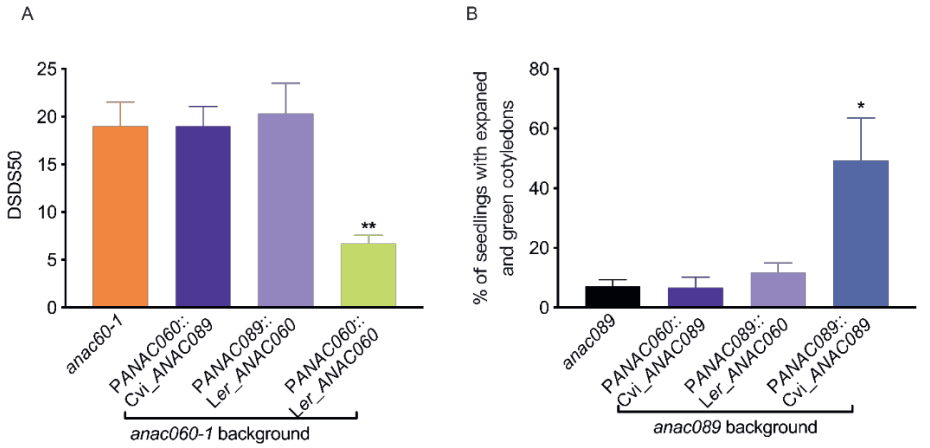


Figure 6. No functional redundancy between *ANAC060* and *ANAC089*. A, Dormancy of *ANAC060* and *ANAC089* swapping lines. Seed dormancy is shown as DSDS50 (days of seed dry storage required to reach 50% germination) levels. B, Fructose sensitivity of transgenic lines. Seeds were sown on ½ MS

plates and ones contain 6% fructose plates. Statistical significances were calculated using one-way ANOVA. Bars indicate the mean value of three replicates and the standard error. Asterisks indicate significant differences compared to the respective single mutants (*: P ≤ 0.05, **: P ≤ 0.01).

Discussion

The NAC family of transcription factors is one of the largest plant specific transcription factors containing over 100 genes that regulate multiple biological processes such as cell division, pathogen resistance, leave senescence (Kim *et al.*, 2006; Riechmann *et al.*, 2000; Seo *et al.*, 2010; Wu *et al.*, 2012; Yang *et al.*, 2011). A comprehensive analysis was performed for the 75 NAC transcription factors in *Oryza sativa* (monocot) and 105 NAC family genes in Arabidopsis (dicot). Based on amino acid similarity, all the NAC family proteins were divided into two main groups and 18 subgroups. *ANAC060*, *ANAC040* and *ANAC089* were classified into same subgroup, also referred to as OsNAC8 in *Oryza sativa* (Ooka *et al.*, 2003). In the Arabidopsis NAC proteins classification, *ANAC060* and *ANAC089* belong to same subfamily IIIa, whereas *ANAC040* is part of the VIa subgroup (Zhu *et al.*, 2012). The phylogenetic tree revealed that *ANAC040* was retained from the At-beta duplication, and *ANAC060* and *ANAC089* were syntenic duplicates from the At-alpha duplication (Fig. 2B). Thus, *ANAC040* is likely more ancestral than *ANAC060* and *ANAC089*. NAC transcription factors are often pleiotropic this is also the case for the here studied genes, *ANAC060* regulates seed dormancy and sugar (glucose and fructose) sensitivity (Li *et al.*, 2014; Yu *et al.*, 2020)(Chapter 2). *ANAC040* inhibits seed germination under high salt concentrations, negatively regulates trichome formation and might play a role during dormancy cycling (Buijs *et al.*, 2020; Kim, Sang-Gyu *et al.*, 2008; Tian *et al.*, 2017). *ANAC089* is negatively regulating floral initiation and promoting seed germination under high concentration of fructose (Li *et al.*, 2010; Li *et al.*, 2011). Based on the high homology between *ANAC060*, *ANAC040* and *ANAC089* and the fact that these genes confer germination related phenotypes we aimed at studying the possible redundancy in more detail.

Promoter activity determines the functional diversity between *ANAC060* and *ANAC040*

We generated transgenic Arabidopsis seeds containing swops of the nuclear localized forms of the ANAC060*,* ANAC089 and ANAC040 proteins driven by their respective promoters P*ANAC060,* P*ANAC089* or P*ANAC040* in *anac060-1*, *anac08*9 and *anac040* backgrounds. We have shown that the CDS of ANAC060 and ANAC040 are interchangeable in the sense that the ANAC060 CDS can complement the *anac040* mutant phenotype and *vice versa* the ANAC040 CDS the *anac060* mutant phenotype (Fig. 3). The lack of complementation of the dormancy phenotype using the *ANAC040* promoter is explained by the very low activity of this promoter in dry seeds (Fig. 4).

Promoter motif analyses were performed to reveal if the differences in promoter activity could be explained by the presence of known motifs. Several of the identified motif had been identified in the promoters of genes that are expressed in seeds before, however most of them are stress related. ABRE is a *cis*-acting element regulating ABA-related gene expression. This motif was initially detected in wheat early-methionine-labelled (Em) gene and recognized by leucine zipper protein during seed maturation (Guiltinan *et al.*, 1990; Morris *et al.*, 1990). In Arabidopsis, it was indicated to be involved in regulating expression of *rd29* responding to drought and salt stresses (Narusaka *et al.*, 2003). MYB motifs have been indicated to play a role in reacting to hormonal signals during seed development and germination. AtMYB2 responds in addition to drought and salt stress also induced by ABA (Shinozaki *et al.*, 1992). The G-box motif determines seed specific expression in tobacco transgenics (Ouwerkerk *et al.*, 1999). The GT1-motif has been suggested to activate transcription of light-dependent genes in tobacco transgenics (Lam *et al.*, 1990). In P*ANAC060* four specific motifs were found, MYC, CAT-box, Activator Protein 1 (AP-1) and W-box. MYC motif with sequence CATGTG acts as a core DNA binding site in Arabidopsis *EARLY RESPONSIVE TO DEHYDRATION STRESS 1* (*ERD1*) which is a drought induced gene (Tran *et al.*, 2004). The CAT-box is described to be related to meristem expression (Lescot *et al.*, 2002). Activator protein 1 (AP-1) was initially

identified to function in mice cell proliferation and survival and has been described to functions in response to oxidative stress in several other organisms (Karin *et al.*, 1997; Shaulian *et al.*, 2001; Lev *et al*., 2005; Scandalios, 2005). The W-box motif is recognized by WRKY transcription factors (Rushton *et al.*, 2010), which are involved in biotic and abiotic stress, including seed dormancy and germination (Rushton *et al.*, 2012). The P*ANAC040* promoter contains two specific motifs. Stress response element (STRE) is able to regulate various stresses-induced genes (Estruch, 2000; Wieser *et al.*, 1991). The GA-motif is a light responsive element (Lescot *et al.*, 2002) (Table. 1).

The *ANAC040* promoter was reported to be regulated by different environmental conditions and might by induced by the binding of the TFs to the ABRE and STRE motif (Table. 1). Both the *ANAC040* transcript and protein levels were highly elevated during cold-imbibition, by 150mM NaCl and the combination of cold-imbibition and 150mM NaCl, suggesting these three distinct conditions induce *ANAC040* parallelly or additively (Kim *et al.*, 2008). Based on this information, we tested the cold induced expression of *ANAC040* in the transgenic lines together with controls. *ANAC040* was clearly induced in the complementation lines, however the expression was significantly lower than in Col-0 (Fig.4D). The promotor sequence that was used for the cloning started 1.722 bp upstream of the start codon, however the final 112 bp before the start codon were missing due to cloning difficulties. Earlier it was reported that that a genomic sequence containing around 2K bp of the native *ANAC040* promoter exhibited full complemented phenotypes after treatment with the gibberellin inhibitor paclobutrazol (Kim *et al.*, 2008). Motif analysis for the 112 missing nucleotides of the *ANAC040* promoter revealed that this region was enriched for the G-box and TATA-box motifs and contained one copy of the ABRE, CAAT-box and as-1 motif in Supplemental figure 2. The lack of these motifs might explain the lower expression of *ANAC040* in the complementation lines and thus the lack of complete complementation. The random insertion of the construct might be another explanation for the lower expression.

Based on the fact that the *ANAC040* CDS could overcome seed dormancy at 25°C (Fig. 3B) we hypothesized that *ANAC040* could be involved in the regulation of secondary dormancy. Secondary dormancy occurs in imbibed seeds and might have occurred already before desiccation tolerance and thus primary seed dormancy existed. This hypothesis is supported by the fact the *ANAC040* is likely the ancestral gene from which *ANAC060*. *DOG1*, which is a main determinant of primary seed dormancy (Alonso-Blanco *et al.*, 2003; Bentsink *et al.*, 2010; Bentsink *et al.*, 2006), is also involved in the regulation of secondary seed dormancy as was revealed from secondary dormancy induced by warm as well as cold stratification (Footitt *et al.*, 2013; Footitt *et al.*, 2015; Murphey *et al.*, 2015). Based on our data, *ANAC060* and *ANAC040* have opposite functions compared to *DOG1*. *DOG1* is important for the induction of seed dormancy, whereas *ANAC060* and *ANAC040* inhibit seed dormancy (Fig .1G, Chapter 2; Fig. 3B). Notably the expression pattern of *ANAC040* is also opposite to that of *DOG1* during dormancy cycling in the field (Fig. 5B).

Respective specific motifs in promoter of *ANAC060* and *ANAC040* might cause the distinct functions

Promoters are known as *cis*-regulatory sequences playing critical roles in regulating genes expression at transcriptional level in all organisms (Wray *et al.*, 2003). There are various motifs in promoter sequences which are involved in regulation of genes (Halees *et al.*, 2003). In Table 1, the twelve overlapping motifs of the *ANAC060* and *ANAC040* promoter are listed. Since these motifs are present in both promoters it is not very likely that they determine the differences in expression of both genes. However, it should be noted that the frequency and location of these motifs in both promoters were not taken into account. In the promoter of *ANAC060*, four specific motifs are present, among them is the W-box (TTGACC) motif that is preferably bound by WRKY transcription factors (Ciolkowski *et al.*, 2008). The WRKY transcription factor family is indicated to be a key regulator in ABA responsive signaling pathways and several members of this family have been related to seed dormancy or germination as well (Rushton *et al.*, 2012). WRKY41 regulates seed dormancy by increasing *ABSCISIC ACID INSENSITIVE 3* expression (Ding *et al.*,

2014), WRKY6 is involved in ABA signaling by down regulating *ETHYLENE RESPONSE DNA BINDING FACTOR 4* (*RAV1*) (Huang *et al.*, 2016) and WRKY2 regulates seed germination and post germination in response to ABA (Jiang *et al.*, 2009). The presence of the W-box in the *ANAC060* promoter might indicate a role for WRKY transcription factors in the regulation of *ANAC060*. The STRE motif (AGGGG), that is found in the *ANAC040* promoter, is reported to be essential for the activation of transcription in response to stress (de Winde *et al.*, 1997)(Table 1).

Divergent functions of *ANAC060* and *ANAC089*

Genome wide expression analysis shows that *ANAC089* is highly expressed in dry seed (Winter *et al.*, 2007), it peaks at the same time as *ANAC060*. Nevertheless, the *ANAC089* promoter is not able to rescue the *anac060* dormancy phenotype when driving the *ANAC060* CDS, neither the *ANAC060* promoter is able to rescue the *anac089* phenotype when driving the *ANAC089* CDS. This might be explained by the different temporal and spatial expression of *ANAC060* and *ANAC089*. Also, the coding region of *ANAC089* could not complement the *anac060* dormancy phenotype when driving by the *ANAC060* promoter as was the case *vice versa* for the *ANAC060* CDS (Fig. 6). The sequence identity of both proteins is high, however the differences in function might be explained by structural differences. This was for example shown for *ANAC019*. Through X-ray crystallography, the ANAC019 NAC domain was identified to contain a twisted β-sheet surrounding with some helical elements instead of the helix-turn-helix motif that is common in several other NAC transcription factors, such a unique structural feature leads to diverse protein functions (Ernst *et al.*, 2004; Olsen *et al.*, 2005).

Conclusions and perspectives

Through a transgenic swapping study, we showed that the NAC transcription factors *ANAC060* and *ANAC040*, but not *ANAC060* and *ANAC089* are functionally redundant. The phenotypic effects that we see are caused by differences in the expression patterns of both genes and these could be further investigated by replacing or knocking out specific motifs in each predictive promoter sequence and

examining the respective phenotypes of these alterations. To provide more evidence for a role for *ANAC040* in dormancy cycling it would be useful to determine the dormancy cycling phenotype of the *anac040* mutant. However, because genetically modified organism (GMO) is prohibited in field experiment in Europe, we will for now depend on the inducting of secondary seed dormancy in lab conditions (Bouchaut *et al.*, 2020). So far, we have not been able to induce secondary dormancy in Col-0 and *anac040* mutant using the thermal treatment that was described by Footitt *et al.* (2017).

Materials and Methods

Plant materials and growth conditions

Seeds of the *Arabidopsis thaliana* accession Columbia (Col-0), *anac060-1* mutant (SALK_127838C), *anac060-2* mutant (SALK-012554C), *anac040* mutant (SM_3.16309; alias *anac040-1* mutant (SM_3.16309) (Kim *et al.*, 2007) and *anac040-2* (SM_3_16309) (Tian *et al.*, 2017)) were obtained from the Nottingham Arabidopsis Stock Centre. The mutant *anac089* (Gt19255) in the Landsberg *erecta* (L*er*) genetic background was obtained from Sheng Teng (Li *et al.*, 2011).

Seeds were sown on water imbibed white filter paper in transparent petri dishes, and then placed in 4°C dark room for cold stratification. After three days, seeds were transferred to a growth chamber at 22°C with continuous light for another day before planting. Germinated seeds with radical protrusion were grown on 4 X 4 cm Rockwool blocks in a climate room at 20°C / 18°C (day/ night) under a 16h artificial light (150 µmol m^{-2} s^{-1}) photoperiod and 70% relative humidity. Plants were watered with a standard nutrient solution and three times per week.

Seed germination experiments

Seed dormancy comparisons: In each plastic tray (15 x 21 cm), 2 layers of blue paper with 48 mL demi water, 50 – 100 seeds were sown using a plastic mask to accurate position the seeds samples. Afterwards the trays were piled and wrapped into transparent plastic bags. The trays were placed into an incubator with 22°C, 30 W m^{-2} continuous light. Pictures to assess the germination were taken for two times per day during a seven-day cycle. Germination was scored using the Germinator package (Joosen *et al.*, 2010). Dormancy level was evaluated by DSDS50, which stands for the days of seed dry storage to reach 50% germination. The method to determine DSDS50 is presented in (Bentsink *et al.*, 2010).

Salt and fructose treatments: Around 50 surface sterilized seeds were plated on sterilised 1.2 % agar ½ MS medium containing either 0 mM and 150 mM salt; 0% or 6,5% fructose. For all conditions the pH was set to at 5.8. Seeds were stratified in 4°C in the dark for 2 days, and transferred to 22°C, in continuous light for 5-7 days.

For salt treatments: germination percentage were calculated on a day base. Seeds with a radicle occurrence were categorized to be germinated (Kim *et al.*, 2008). For Fructose treatment: Fructose sensitivity phenotype was evaluated by counting the percentage of green seedlings with expanded cotyledons (Li *et al.*, 2011).

DNA isolation from seeds

Around 200 seeds and two small bullets were put into 1.5mL tubes, seeds were ground for 1 min, 30HZ, afterwards, 250 µl extraction buffer (2 M NaCl, 200mM Tris-HCl PH 8, 70 mM EDTA, 20mM $Na_2S_2O_5$) was added to each sample and followed by grinding again as indicated for the previous step. Samples were kept at 60°C for 1hour. After that, samples were centrifuged for 10 mins at maximum speed. 75 µl clean supernatant was taken from each tube into a new tube, 75 µl iso-propanol and 30 µl 10 M NH_4AC were added and mixed well in each tube. All samples were kept for 15 minutes at room temperature for precipitation, after that they were centrifuged for 20 mins at maximum speed. Supernatant was discarded and DNA pellet was washed by 70 µl 70% ethanol, after that, the samples were centrifuged for 5 mins at maximum speed. The ethanol was removed using a pipette, the DNA pellet was dried for 10 minutes. Finally, the pellet was dissolved in 30 µl MQ (Cheung *et al.*, 1993).

PCR condition and gel electrophoresis

Polymerase chain reactions (PCR) were operated in a 25 µl total volume including around 100ng DNA, 1.25 µl of forward and reverse primers, 12.5 µl VYO HIFI mix (SOPACHEM PB10.43). The cycling programme was as follows: First denaturation at 94°C for 3 minutes in one cycle, following by second denaturation 94°C for 10 seconds. Annealing temperature was from 59-64°C which depended on the primers for 30 seconds and a 45 seconds extension at 72°C, this cycle was replayed for 35 times. The final extension was at 72°C for 5 minutes. The PCR products were checked by 1% agarose gel electrophoresis.

Synteny network construction and phylogenetic reconstruction

The synteny network of the NAC gene family was constructed by extracting the syntenic relations of all the NAC genes from the synteny network database of 107 plant genomes (Zhao *et al*., 2019). To do this, the NAC gene IDs identified from the 107 plant genomes using HMMER3 (Mistry *et al.*, 2013) was used to query the network database. The edgelist of the NAC subnetwork was clustered and visualized in Gephi (Bastian *et al*., 2009). We located and highlighted the synteny cluster containing *ANAC060*, *ANAC040* and *ANAC089*. A closely inter-connected synteny cluster indicates a shared genomic context. The phylogenetic tree was then reconstructed for the nodes/genes in this cluster. Multiple sequence alignments were performed using MAFFT (version 7.187). Alignment trimming were conducted by trimAl (Capella-Gutiérrez *et al.*, 2009). Maximum-likelihood analyses were conducted using IQ-TREE (Nguyen *et al.*, 2015). We used the 'JTT+R' model for the protein alignment, with 1000 bootstrap replicates (-bb 1000)).

Phylogenetic analysis

ANAC060 was analysed by plant membrane protein database (Schwacke *et al.*, 2003), and 15 high homology genes were collected. The protein sequences were built to be phylogenetic tree by (Kumar *et al.*, 2016).

Swapping experiment

Recombined constructs by swapping the promoters and genomic CDS among *ANAC060*, *ANAC040*, *ANAC089* was shown in Table 2.

Table 2. Swapping transgenic lines for *ANAC060* and its homologous genes. Transgenic lines were made by switching promoters and genomic CDS of *ANAC060* and *ANAC040* (A), *ANAC060* and *ANAC089* (B), the recombined constructs were transformed into respective backgrounds as indicated in the tables.

A

Background	Recombined constructs		
Anac060	P*ANAC060*::L*er_ANAC060*	P*ANAC060*::Col-0_*ANAC040ΔC*	P*ANAC040*::L*er_ANAC060*
Anac040	P*ANAC040*::Col-0_*ANAC040ΔC*	P*ANAC060*::Col-0_*ANAC040ΔC*	P*ANAC040*::L*er_ANAC060*

B

Background	Recombined constructs		
Anac060	P*ANAC060*::L*er_ANAC060*	P*ANAC060*::Cvi_*ANAC089*	P*ANAC089*::L*er_ANAC060*
Anac089	P*ANAC089*::Cvi_*ANAC089*	P*ANAC089*::L*er_ANAC060*	P*ANAC060*::Cvi_*ANAC089*

Expression vectors construction by Gateway cloning

Gateway cloning was performed according to the manual Multisite Gateway® Pro: Using Gateway® Technology. This method allows to simultaneously clone multiple DNA fragments. All primers used for cloning are listed in Supplemental table 3.

PCol-0_*ANAC040*::Col-0_*ANAC040ΔC* construct: The native *ANAC040* promoter was cloned from accession Col-0 (pCol-0_*ANAC040*) and the genomic CDS of *ANAC040* with an added stop codon was also cloned from wildtype Col-0 (Col-0_*ANAC040ΔC*) (Kim *et al.*, 2007). Both of the fragments were amplified using the VYO HIFI mix and cloned into gateway entry vectors pDONR 221 + P1P5 and pDONR 221 + P5P2 respectively. Afterwards, the two entry vectors were cloned into the destination vector 428pKGW red seed + R1R2 which allows transformant selection based on fluorescence.

PCol-0_*ANAC089*::Cvi_*ANANC089* construct: The native *ANAC089* promoter was cloned from ecotype Col-0 (PCol-0_*ANAC089*) and the genomic CDS of *ANAC089* was cloned from accession Cvi (Cvi_ *ANAC089*) (Li *et al.*, 2011). The way to make

the construct is comparable with explained in PCol-0_*ANAC040*::Col-0_*ANAC040ΔC* construct.

PCol-0_ *ANAC060*::L*er*_*ANAC060* construct: The native *ANAC060* promoter was cloned from PMD18_Col-0 (Li *et al.*, 2014). The genomic CDS of *ANAC060* was cloned from L*er* (Chapter 2). The method to make the construct is identical with described in PCol-0_*ANAC040*::Col-0_*ANAC040ΔC* construct.

The further procedures to build the constructs of swapped promoters and genomic CDS between *ANAC060* and *ANAC040*/*ANAC089* were same with what were described above.

Transformation of *Arabidopsis thaliana*

Cell cultures are centrifuged for 15 mins at 4000 rpm under room temperature in 250ml tubes, the supernatant was poured off, after that the pellet was resuspended in infiltration medium (½ MS medium + vitamins, 5% sucrose) to an OD600 of 1. Prior to be used, 0.03% Silwet L-77 was added. Afterwards, the flowering plants were dipped into the culture for 15 seconds. The plants were placed in plastic bags in horizontal position for overnight. After the plastic bags were opened, the plants were grown in upright position until seed to be harvested.

RNA isolation and characterization

The detailed and optimized protocol for RNA isolation and quality evaluation from seeds is present in Supplemental protocol 1.

cDNA synthesis and qPCR analysis

iScript cDNA synthesis kit (Bio Rad) was used for making cDNAs, in this project, 1 µg RNA was used from each sample. cDNA samples were one-tenth diluted. qPCR was done and analyzed according to qPCR guide Eurogentec. The master mix was 10 µl for each: 2.5 µl cDNAs, 5 µl SYBR green (sopachem VYBA01-31), 0.5 µl primers mixtures, 2 µl MQ water. The CFX Bio Rad was used to generate threshold cycle value for each reaction. Excel and Prism 8 software were used to analysis qPCR result.

qPCR primers: *ANAC040*, forward primer: AGGATGCATTAGTGGTGTGC, reverse primer: TTGTCCTCCTTCTCCAAACC. *ANAC060*, forward primer: AGCCTTGGGATTTACCTGA, reverse primer: TTGGTTGCTCTTCTGTTCTGT. Two reference genes, *At4G12590* and A*t4G23270,* are designed based on study by (Dekkers *et al.*, 2012).

Motif analyses

Predictive promoter sequences of three genes were analysed in PlantCARE database (plant cis-acting regulatory elements) (Lescot *et al*., 2002), qualitative motifs of three single genes were presented and described in Supplemental table 3.

Supporting Information

Supporting information can be downloaded by scanning the QR code or from http://www.wageningenseedlab.nl/thesis/ssong/SI/

Supplemental Figure 1. Fructose sensitivity in the single *anac060-1* and *anac089* mutants and their respective wildtypes.

Supplemental Figure 2. 112 missing base pairs in the cloned *ANAC040* promoter.

Supplemental Table 1. List of 107 species used for synteny network analysis.

Supplemental Table 2. Motif lists and the function description of promoter of *ANAC060*, *ANAC040*, *ANAC089*.

Supplemental Table 3. List of primers used in Gateway cloning.

Supplemental Protocol 1. Optimized RNA extraction protocol from dry Arabidopsis seeds.

Chapter 4

Degradation patterns of seed stored mRNAs and proteins during seed ageing

S. Song, M. Coenen, P. Shakya, A. Liefferink, L. A.J. Willems, L. Bentsink, H. Nijveen

Abstract

Seeds are usually stored in well controlled storehouses with restrained relative low temperature and humidity to conserve seed quality for a long period. Nevertheless, seeds will gradually lose vigor and eventually die. Oxygen has been identified as a main factor influencing seed viability during long term storage, however it plays a dual role during the seed lifespan. On the one hand, its product Reactive Oxygen Species (ROS) are beneficial for seed dormancy alleviation during after ripening. On the other hand, ROS progressively cause damage in the seed by altering the fluidity of cell membrane, influencing the transport of ions, enzyme deactivation, protein carbonylation and DNA and RNA damage, which all together ultimately result in seed death. Among these, the integrity of genetic materials largely impacts seed viability. RNA is vulnerable to ROS as it has single strand structure and an RNA repair mechanism is lacking at dry seed state. Oxidation of messenger (m) RNA results in the inhibition of translation or the production of abnormal proteins. Protein oxidation causes protein carbonylation, which irreversibly alters protein function and leads to detrimental protein cross-links. To uncover the oxidation pattern of seed stored RNAs and proteins along with seed ageing in a limited time span, an artificial ageing system elevated partial pressure of oxygen (EPPO), which mimics and accelerates seed ageing, has been applied. The integrity of total stored RNAs, the dynamic expression patterns of seed stored mRNAs and the oxidation profiles of seed storage proteins were explored. We have identified severe mRNA degradation during seed ageing and show that this degradation correlates with the germination capacity of seeds. Moreover, we have a strong indication that the degradation of monosome associated transcripts is delayed.

Introduction

Storing seeds in an appropriate way is critical not only for breeders to ensure high crop yield and keep good quality seeds for planting the next generation, but also for gene banks to conserve plant genetic biodiversity over decades or even centuries (Bewley *et al.*, 2012). Hence, seed research and understanding the factors which deliver detrimental effects during seed storage would offer great help for worldwide sustainable agriculture.

In the end of seed maturation, seeds start to lose water and become dry, this process is named 'desiccation' (Kermode *et al.*, 2002). During desiccation, seeds enter a glass state and maintain only extreme low levels of metabolic activity (Buitink *et al.*, 2008; Burke, 1986; Rajjou *et al.*, 2008; Walters *et al.*, 2005). Along with this process, seeds acquire seed dormancy which temperately retains seeds from germination (Bewley, 1997; Hilhorst, 1995). Mature dry dormant seeds are harvested and stored in storehouses or gene banks for future use. During dry storage, seeds gradually release dormancy until they reach full germination capacity, this phenomenon is called after ripening. At the same time the ageing process starts, seeds start to deteriorate and will not be able to germinate eventually. The period during which seeds maintain their high germination ability has been defined as seed longevity (Koornneef *et al.*, 2002; Rajjou *et al.*, 2008; Walters *et al.*, 2005). The level of seed longevity correlates to the ability of seeds to become desiccation tolerant (Hong *et al.*, 1996). Seeds which are extremely sensitive to drying have generally short lifespans, e.g., *Acer saccharinum* seeds die after a few days of desiccation (Connor *et al.*, 2001). Seeds which are moderately sensitive to desiccation tend to have intermediate long lifespans, like coffee seeds which can survive for 12 months with a low water content (Ellis *et al.*, 1990). Extreme long lifespan seeds are generally tolerant to desiccation, like the Judean date palm for which germination has been reported after nearly 2000 years of storage (Erlanger, 2005). Seeds degrade at different rates during storage. However, for all of them counts that they become increasingly vulnerable to environmental stresses, for instance, humidity of the air, oxygen and temperature during storage (Belletti *et al.*, 1991; Hendry, 1993).

Eventually seeds lose vigour and cannot germinate anymore, this phenomenon is referred to as seed ageing (Walters, 1998). Of the factors influencing seed ageing, oxygen has been indicated to have the most distinct injurious consequences (Hendry, 1993; Kurek *et al.*, 2019; Ohlrogge *et al.*, 1982). The insufficient or unfinished reduction of oxygen, or energy transfer can cause the production of by-products. Examples of these are, superoxide $O_2(-)$, hydroxyl radical ($OH\cdot$), singlet oxygen ($^1O_2$), hydrogen peroxide (H_2O_2). All these derivatives are referred to as reactive oxygen species (ROS) (Apel *et al.*, 2004). Oxygen is known to accelerate seed deterioration in all kinds of seeds. For example, dehydration in *Castanea sativa* results in a rapid boost of superoxide which causes seed viability loss and hydrogen peroxide is involved in sunflower seed deterioration. Generally higher oxygen concentrations in the storage environment result in a shorter seed longevity (Justice *et al.*, 1978; Kibinza *et al.*, 2006; Ohlrogge *et al.*, 1982; Roach *et al.*, 2010). Moreover, the beneficial effects of specific variety of ROS in seed physiology were also studied and discussed. For instance: Hydrogen peroxide releases seed dormancy in sunflower, barley, soybean and Arabidopsis. Superoxide accumulation promotes growth of embryonic axes in soybean, *Vigna radiata* (L.) Wilczek (Bahin *et al.*, 2011; Leymarie *et al.*, 2012; Oracz *et al.*, 2009; Puntarulo *et al.*, 1988; Singh *et al.*, 2014). Next to the afore mentioned effects, there are numerous underlying ways by which ROS interacts with components in seeds, this leads to the oxidation of lipids, proteins, DNA and RNA. In this chapter, interaction between ROS and proteins, mRNAs is studied (Kranner *et al.*, 2011; Oracz *et al.*, 2007; Terskikh *et al.*, 2008).

Embryos in mature dry Arabidopsis seeds accumulate proteins and mRNAs that are essential for initial transcription upon imbibition (Bewley *et al.*, 1994; Kimura *et al.*, 2010; Nakabayashi *et al.*, 2005). Generally, mRNAs were found to be degraded in minutes to hours in Arabidopsis cell cultures (Narsai *et al.*, 2007). However, seed stored mRNAs in mature dry rice, cotton and Arabidopsis seeds remain intact for years. Therefore, these mRNAs are also called long-lived messenger RNAs (Dure *et al.*, 1965; Galland *et al.*, 2014; Sano *et al.*, 2012). After imbibition, non-dormant Arabidopsis seeds complete germination in the presence of the transcriptional

inhibitor α-amanitin, while germination is suppressed by the translational inhibitor cycloheximide (Rajjou *et al.*, 2004). This underlines the importance of translation of stored mRNAs for germination and is confirmed by the finding that 17% of the seed stored mRNAs that are bound to monosomes (one single ribosome) in dry seeds are translationally upregulated during early seed germination (Bai *et al.*, 2020). Gene ontology analyses revealed that these selectively up-regulated transcripts belong to the categories “response to water deprivation”, “redox balance”, the proteins in monosome fractions are related to protection from oxidation as well. Taken together, these studies suggest certain groups of seed stored transcripts and proteins may play roles in antioxidation during long term seed storage (Bai *et al.*, 2020; Dure *et al.*, 1965; Narsai *et al.*, 2007; Rajjou *et al.*, 2004; Sano *et al.*, 2015).

In order to address how ROS impacts seed stored mRNAs and proteins, we need to divide seeds into two physiological stages, dry seeds and imbibed seeds. In dry mature seeds, due to the extreme low water content, ROS mainly derive from non-enzymatic reactions, for instance, Amadori and Maillard reactions, lipid peroxidation and protein carbonylation which is an irreversible and one of the most detected and destructive oxidation modifications causing protein dysfunction (Arc *et al.*, 2011; Dalle-Donne *et al.*, 2006; Priestley, 1986; Sun *et al.*, 1995). In parallel, the non-enzymatic ROS scavenging antioxidants, such as, flavonoids and tocopherol, are functioning as oxidation buffers to protect other essential compounds (Apel *et al.*, 2004).

At the dry seed state, ROS oxidize mRNA, resulting in structural changes by generating 8-hydroxyguanosine (8-OHG) and strand breaks, causing translational errors and protein synthesis deficiency upon imbibition, which eventually lead to cell death (Bazin *et al.*, 2011; Fleming *et al.*, 2018). Evidence for mRNA oxidation has been reported several times for different species. In sunflower seeds, 24 seed stored mRNAs are reported to be selectively oxidized based on the increasing content of the RNA oxidation marker 8-OHG during dormancy release. Among the oxidized genes are several genes that are functioning in stress responses which also have been related to the regulation of seed dormancy (Bazin *et al.*, 2011; Bentsink *et al.*,

2010; Carrera *et al.*, 2008). In ageing soybean seeds, longer seed stored transcripts become fragmented by ROS attacks at random bases and although shorter transcripts keep their high integrity, seeds still die. Thus, it is hypothesized that the longer transcripts are critical for preserving seed vigour (Fleming *et al.*, 2018).

In sunflower, accumulation of ROS during dry storage leads to lipid peroxidation and selective embryo proteins carbonylation causing release dormancy release (Oracz *et al.*, 2007). The amount of carbonylated proteins is also positively correlated with seed viability loss. Especially the main seed storage protein 12-cruciferins are dominantly oxidized in dry seed (Job *et al.*, 2005; Terskikh *et al.*, 2008). In addition, the cruciferin (CRU) *crua crub cruc* triple mutant shows much shorter seed longevity compared with wildtype Col-0 after an artificial ageing treatment. These results imply that protein carbonylation affects seed after ripening and longevity during dry storage and suggests that the abundant seed storage proteins service as a protective shield to defend essential germination related proteins from being carbonylated (Job *et al.*, 2013; Nguyen *et al.*, 2015; Oracz *et al.*, 2007; Rajjou *et al.*, 2008).

When dry seeds are exposed to water they will imbibe. During imbibition the glassy state gradually changes into a fluid state. There are three phases of seed imbibition. In phase I seeds physically absorb water, cells are rehydrated until seeds are totally imbibed. In phase II there is no additional water uptake, however, metabolism is reactivated, seed stored mRNAs are translated or degraded and new mRNAs are transcribed and translated (Dekkers *et al.*, 2013; Marcus *et al.*, 1964). Moreover, mitochondrial repair occurs. This phase ends when the radicle protrudes through the embryo surrounding tissues to complete seed germination. Phase III is marked by further water uptake and seedling growth (Nonogaki *et al.*, 2010). During rehydration diverse ROS are produced by all types of cells and various cellular enzymes (Apel *et al.*, 2004; Finkel, 1998; Vranová *et al.*, 2002). Simultaneously, enzymatic activities like antioxidant machinery are reactivated to scavenge ROS and control cellular redox balance (Bailly, 2004; Tommasi *et al.*, 2001; Wojtyla *et al.*, 2006). In hydrated eukaryotic cells, oxidized mRNAs might be repaired by certain metabolic activities and continue to be translated (Aas *et al.*, 2003; Li, *et al.*, 2006; Yan *et al.*, 2019). 5’

to 3' mRNA decay during seed imbibition is thought to regulate seed dormancy release. Mutants in genes that affect 5' to 3' mRNA decay show dormancy phenotypes, *exoribonuclease4* (*xrn4*) shows higher dormancy and *varicose* (*vcs*) exhibits lower dormancy compared with wildtype Col-0 (Basbouss-Serhal *et al.*, 2017).

Mild carbonylated proteins could be eliminated by enzymatic activities, whereas strongly carbonylated proteins might be resistant to degradation therefore increase the level of damaged or aberrant proteins in hydrated cells (Dalle-Donne *et al.*, 2006). During rice germination, 66 carbonylated proteins are analysed for their dynamic carbonylation intensity patterns after 0, 24, 48 and 72 hours imbibition. Carbonylation abundance of seven stress responsive proteins is significantly decreased and some storage proteins are gradually reduced while other proteins either significantly or gradually increased during germination. For example, carbonylation pattern of stress responsive small heat shock proteins and Heat Shock Protein 70 is decreased and enhanced, respectively. This implicates that different proteins display specific functions in response to ROS stress during germination (Zhang *et al.*, 2016). Arabidopsis 12S cruciferins, as a main target of oxidation, display severe carbonylation in dry aged seeds. In contrast, their oxidative products steadily start to fade away whereas other proteins are carbonylated after imbibition, implying that some specific proteins with certain functions are selectively oxidized for completing seed germination (Galland *et al.*, 2014; Job *et al.*, 2005; Rajjou *et al.*, 2006; Rajjou *et al.*, 2008).

To study seed longevity under dry storage in a limited time, elevated partial pressure of oxygen (EPPO) is applied as it is an innovative way to mimic and stimulate seed dormancy release and ageing under controlled relative humidity and temperatures (Groot *et al.*, 2012). In brief, the principal is that higher pressure increases the absolute oxygen concentration but maintains relative oxygen concentration like that is present in ambient air. In Arabidopsis, EPPO exposure shortens the whole seed lifespan curve to around half a year, whereas seeds are able to survive almost a decade under lab bench conditions (Buijs *et al.*, 2018; Buijs *et al.*, 2020; Groot *et al.*,

2012). In this chapter we aim at investigating how and if degradation patterns of seed stored transcripts and proteins under oxidative stress correlate to transitions of physiological seeds states during the Arabidopsis seed life span. We hypothesize that the translation of seed stored mRNAs is essential for seed germination and that degradation of these critical transcripts are the cause of the inability of aged seeds to germinate. We used degradome sequencing and RT-qPCR to determine the degradation patterns of the mRNAs of three main seed physiological stages: freshly harvested dormant seeds, fully after-ripened (non-dormant) seeds and non-germinating (aged) seeds obtained by EPPO treatment. We revealed mRNA degradation patterns during seed ageing and show that this degradation correlates with the germination capacity of seeds. Moreover, we have a strong indication that the degradation of monosome associated transcripts is delayed. In addition, the carbonylation of storage proteins increased during seed ageing eventually also leading to the carbonylation of other high molecular weight proteins.

Results

Stable RNA integrity during seed life span under EPPO treatment

To study the effect of oxidation on the seed life span, *delay of germination 6* (*dog6-2*)/*anac060-2* seeds were aged using the EPPO treatment (Fig. 1A). The mutant was chosen for this study since it has a prolonged dormancy compared to the standard laboratory accessions Landsberg *erecta* or Columbia. Seed germination was analysed at different time-points from 0 to 115 days of EPPO treatment. During the first period, 0 day to 13 days, seed dormancy was fully released. From 13 days onwards seeds started to gradually lose vigour and could not germinate anymore. To confirm the loss of viability in the 115 days EPPO treated seeds the seed coats were taken off. Removal of the seed coat releases the physical constrain of the embryos surrounding layers and allows elongation of the radicle of seeds that have a reduced viability, however not of dead seeds. Viable seeds germinate within two days after removal of the seed coat, seeds after 115 days of EPPO did still not germinate five days later after removing the seed coat and were considered to be dead.

Total RNA was extracted at every time-point of seed life span curve (Fig. 1A). RNA integrity was detected using the Qubit RNA IQ assay and the Agilent bioanalyzer 2100 (Mueller *et al.*, 2004). The Qubit RNA IQ Assay is a fast and easy method to determine the quality and integrity of two different sizes of RNAs. Large/intact RNAs and small/degraded RNAs are bound by different dyes. RNA integrity numbers (RIN) represent the RNA fragments size distributions which are measured through laser-activated fluorescence, higher RIN values indicate a better RNA integrity (Schroeder *et al.*, 2006). Both methods indicate that the RNA quality was stable over the whole seed life span (Fig. 1B and 1C).

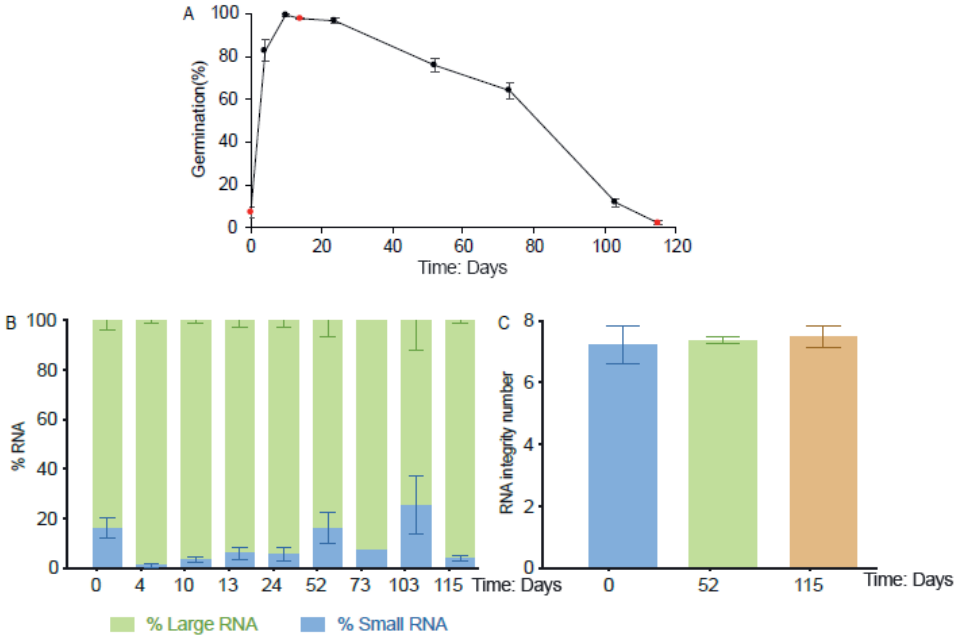


Figure 1. RNA integrity remains stable in seed life span curve generated using EPPO treatment. A, Germination percentages of *dog6-2* mutant seeds after EPPO treatment. From 0 until 13 days the pressure was 8MPa from 13 days onwards the pressure was increased to 20MPa. The conditions during the EPPO treatment were kept constant at 55% RH and ambient temperature. Three representative physiological seed states are indicated by the red dot, from left to right these represent freshly harvested, dormant seeds, fully after-ripened non-dormant seeds and aged, dead seeds. Means and the standard error of three replicates are shown. B, RNA integrities of samples at different time-points (0, 4, 10, 13, 24, 52, 73, 103 and 115 days) during EPPO measured by the Qubit RNA IQ Assay. C, RINs measured using the bioanalyzer, after 0 (freshly harvested dormant seeds), 52 (fully after-ripened, non-dormant seeds) and 115 days (aged, dead seed) of EPPO treatment. RIN: RNA integrity number, higher value stands for better RNA integrity level. The RIN values are means of triplicates ± SE. Bars are the standard error of the mean value.

Seven representative mRNA degradation patterns during seed ageing

To discover whether mRNA degradation patterns during seed ageing correlate with seed physiological states, degradome analyses were performed (German *et al.*, 2009). These analyses are based on the identification of transcripts that lack the 5'-cap (and instead contain a 5' monophosphate) but have an intact 3' poly(A) tail. The

analyses have been performed on dry seeds that represent three different seed ages; freshly harvested dormant seeds (0 day of EPPO), fully after-ripened non-dormant seeds (13 days of EPPO) and aged, dead seeds (115 days of EPPO) (Fig. 1A). Using high-throughput sequencing specifically the 5' ends of partially degraded mRNA molecules are sequenced. Quantification of the resulting sequencing reads was performed and further normalized based on the transcript length, demonstrated as transcript-per-million values (TPM) ultimately (Langmead, Ben *et al.*, 2012; Li *et al.*, 2011). To reveal expression patterns that correlate with the germination patterns the transcripts were clustered based on their expression values. In total, seven representative clusters have been identified. The transcript numbers and length distributions of each cluster are shown in Fig. 2. Clusters 1 and 2 represent transcripts for which the degradation increases with ageing time. In Cluster 1 the transcripts gradually become more degraded from 0 to 13 days of EPPO and show a steep increase between 13 and 115 days. In cluster 2 the number of degraded transcripts accumulates following a nearly linear pattern. Cluster 3 to 6 represent transcripts for which the degradation decreases, either from the beginning or at later phases during the EPPO treatment. In cluster 3 transcripts start with a higher level of degradation which slowly decreases in the after-ripened seeds and thereafter dramatically drop in the aged seeds. Cluster 4 differs from cluster 3 in the sense that there is an increase of degraded transcripts in the fully after-ripened dormant seeds, thereafter the levels of degraded transcripts drop again. Cluster 5 shows a striking decline in number of degraded transcripts from 0 to 13 days of EPPO whereas the level of degradation remains stable afterwards. In cluster 6 there is an initial low level of transcript degradation which is followed by a strong increase in the fully after-ripened seeds and a reduction again in aged seeds. In Cluster 7 the initial level of transcript degradation is much higher than that in the fully after-ripened seeds and the level of degraded transcripts increase again in the aged seed to levels similar as at the starting point (Fig. 2A). The clusters where investigated for their average transcripts length and this revealed that the transcripts in cluster 3 and 4 were significantly longer than those in other clusters (Fig. 2B).

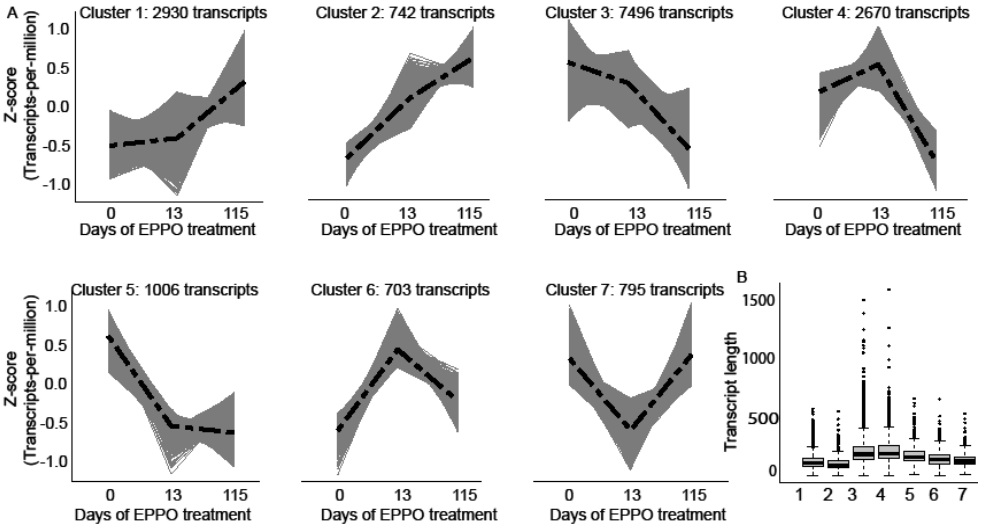


Figure. 2 Seven distinctive degradation patterns of seed stored mRNAs during seed ageing. A, Cluster analyses performed on all identified mRNAs. Clusters representing the degradation patterns identified in seeds at three physiological stages, 0 (freshly harvested dormant seeds), 13 (fully after-ripened, non-dormant seeds) and 115 days (aged, dead seed) of EPPO treatment. For each transcript the Z-score (number of standard deviations above or below the mean) is presented (gray lines) for each physiological stage. A representative gene model is represented by the black line in every cluster. The number of transcripts in each cluster is indicated. B, Transcript length distribution of each cluster.

The critical roles of seed stored mRNAs have been highlighted by different studies (Bai *et al.*, 2020; Bazin *et al.*, 2011; Kimura *et al.*, 2010; Nakabayashi *et al.*, 2005; Rajjou *et al.*, 2004; Zhao *et al.*, 2020). Bai *et al.* (2020) identified 112 transcripts that are monosome bound and translationally up-regulated during early seed germination (Supplemental table 1). These mRNAs are predicted to be protected by ribosome complexes in the dry seed and their protection is hypothesized to be important for seed germination. To investigate the degradation patterns of these 112 transcripts, we looked for overlap with the genes identified in the different clusters. Although the monosome associated mRNAs appear in all clusters, by far the most of them show an increased degradation pattern, represented by cluster 1 and 2 (Fig. 3).

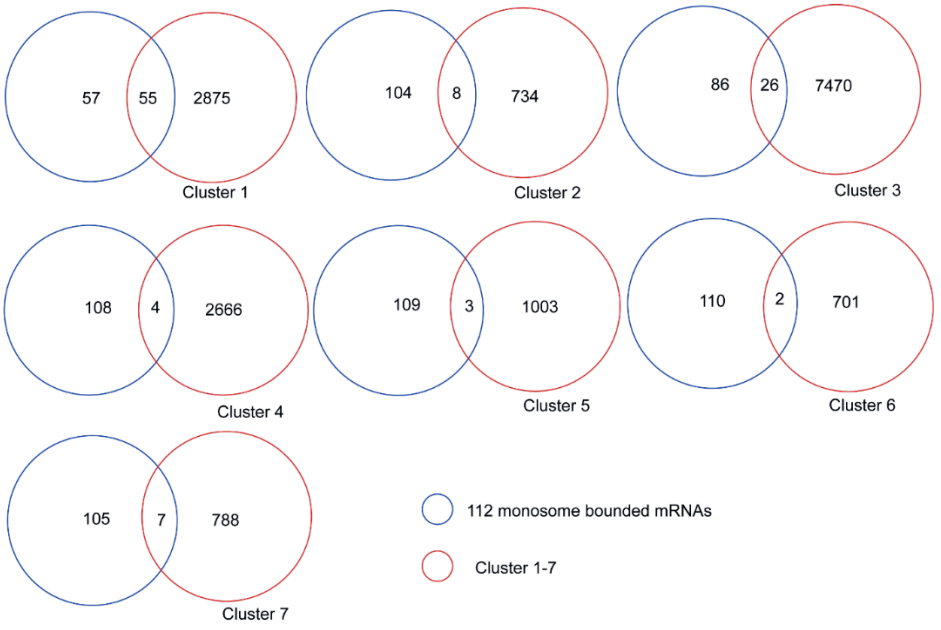


Figure 3. Distribution of the 112 monosome bound mRNAs in the seven degradation clusters. Venn diagram showing the distribution of the 112 monosome bound mRNAs that are translationally up-regulated during earlier seed germination. The number of mRNAs in each comparison are indicated. List of overlapped mRNAs between 112 monosome bound mRNAs and every cluster is presented in Supplemental table 2.

To identify the differentially degraded mRNAs in dry seed following seed ageing, every physiological seed state was compared to the other stages, i.e 0 vs. 13 days of EPPO, 13 vs. 115 days of EPPO and 0 vs. 115 days of EPPO. From 0 to 13 days of EPPO, 140 differentially degraded mRNAs were idenfitied. Seven of them show increased degradation and rest of them decreased degradation (Fig. 4A, Supplemental table 3). Notably, only one of these transcripts (*AT2G41905*) belongs to the 112 monosome associated transcripts which is down-regulated during the seed life span. This gene is described to be a transmembrane protein matches arabinogalactan protein 23 and is related to higher cell wall extensibility (Mabuchi *et al.*, 2016). From 13 to 115 days of EPPO the ratio of transcripts that show an increased degradation (546) and those that show a decreased degradation (425) is

close to one. The monosome associated mRNAs are mostly found among the transcripts that show an increased degradation (Fig. 4B, Supplemental table 3). The highest number of differentially degraded transcripts are found when comparing 0 to 115 days of EPPO, 1008 and 940 mRNAs are significantly elevated and declined in degradation respectively. 18 monosome bound transcripts are remarkably more degraded in the dead seeds (Fig. 4C, Supplemental table 3).

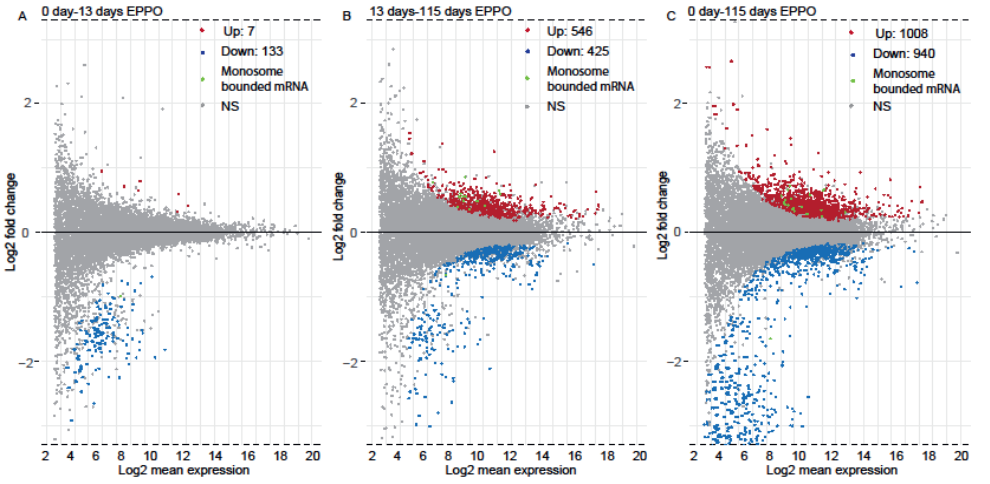


Figure 4. Differentially degraded transcripts during seed ageing. MA-plots, representing the log ratio of differential expression as a function of the mean intensity, for the different comparisons, A, 0 to 13 days of EPPO, B, 13 to 115 days of EPPO and C, 0 to 115 days of EPPO. The mRNAs that are significantly increased in degradation are highlighted in red, mRNAs are significantly reduced in degradation are highlighted in blue, the ones that are overlapped between 112 specific mRNAs and significantly up- and down- expressed in degradome are displayed in light green, non-significant mRNAs are shown in grey.

Fate of seed stored mRNAs in imbibed seeds

The degradome analyses identified mRNAs which were degraded from the 5' site of the transcript but did not provide insight into the degradation information close to 3' site of the mRNA neither allowed for quantification of the (full length) transcript levels. To provide more insight into these aspects we performed RT-qPCR analysis on a subset of the 112 monosome associated transcripts that are translationally upregulated during early seed germination. Note that RT-qPCR does not determine

whether mRNAs are oxidized. If present, the mRNA it will be synthesized into cDNA and amplified in the consecutive PCR. In total 23 gene transcripts have been investigated by RT-qPCR (Fig. 5 and Supplemental figure 1), the ones that were also found to be differentially degraded are shown in Table. 1 with a description of their function. There is no mRNA among these 23 selected ones being more degraded during dormancy release process (0 day EPPO-13 days EPPO), six mRNAs are more degraded during seed ageing (13 days-115 days EPPO), all together nine mRNAs significantly accumulate degradation from fresh seeds (0 day EPPO) to aged seeds (115 days EPPO). Generally, four different expression patterns were recognized, these are described in more detail using an example gene for each category (Fig. 5), profiles of all tested genes can be found in Supplemental figure 1.

Table 1. Nine out of 23 monosome bound mRNAs significantly show increased degradation during the seed life span. The respective gene ID, gene names and description are indicated, the bold ones are mRNAs showing significant changes in these two comparisons: 13 to 115 days EPPO and 0 day to 115 days EPPO.

Gene ID	Other names	Description
AT5G37670.1	HSP15.7	HSP20-like chaperones superfamily protein
AT5G04750.1		F1F0-ATPase inhibitor protein
AT4G35570.1	HMGB5	High mobility group B5
AT4G22220.1	ATISU1	SufE/NifU family protein
AT3G60640.1	ATG8G	Ubiquitin-like superfamily protein
AT2G25720.1		Hypothetical protein
AT4G29350.1	ATPRF2	Profilin 2
AT1G45145.1	ATH5	Thioredoxin H-type 5
AT2G16600.1	ATCYP19-1	Rotamase CYP 3

The patterns are described using four example genes, 1. *AT2G16600*, 2. *AT4G22220*, 3. *AT4G35770* and 4. *AT5G04750*. *AT2G16600*, *AT4G35770* and *AT5G04750* belong to cluster 1 and *AT4G2220* is in cluster 2 (Fig. 3). Both of the clusters display an increasing degradation level during seed ageing which suggests a correlation between the degradation of these mRNAs with a reduction in germination capacity (Fig. 2). For *AT2G16600* and *AT4G2220*, the accumulation of

degraded transcripts is gradual, but significant when comparing the initial degradation level at 0 days of EPPO to that at 115 days of EPPO. *AT4G35770* did not significantly change during the ageing period. *AT5G04750* transcripts degraded significantly during seed ageing (0-115 days of EPPO) with a steep increase from 13 to 115 days of EPPO (Fig. 5A). For the representative transcripts' expression pattern 1, *AT2G16600*, the transcriptome and translatome expression levels were both significantly increased after 6 hours imbibition (Bai *et al.*, 2020) (Fig. 5B). *AT2G16600* expression was stable during the ageing process in dry seeds, however it significantly decreased after 115 days of oxidative ageing in 3 hours imbibed seeds, detected at both 5' and 3' site of the transcripts (Fig. 5C, 5D). This gene is annotated as *ATCYP19-1, CYCLOPHILIN 19, CYP19, ROC3, ROTAMASE CYP 3* and is predicted to be involved in protein peptidyl-prolyl isomerization, protein refolding, signal transduction and is located in the mitochondrion, chloroplast and plasma membrane (Swarbreck *et al.*, 2007). Pattern 2, the expression of *AT4G22220* severely declines at the transcriptome level whereas it elevates at the translatome level after 6 hours imbibition (Bai *et al.*, 2020) (Fig. 5B). Different from previously mentioned genes, fold change of 5'-site amplified expression level of this gene was significantly lower than 3'-site amplification when comparing the absolute transcripts levels between these two sites at both dry and 3 hours imbibed seed states in seed lifespan. At the 3' site, from 13 days to 115 days ageing process, the expression was sharply reduced in imbibed seeds (Fig. 5C, 5D). *AT4G22220* is annotated as *ATISU1*, *ISU1* and encodes a mitochondrial protein and is located in the mitochondrion (Swarbreck *et al.*, 2007). Pattern 3, the expression of *AT4G35770* is dramatically decreased at the transcriptome level after 6 hours imbibition, whereas expression at translatome is stable during imbibition (Bai *et al.*, 2020) (Fig. 5B). Expression of both the 5' and 3' site was stable throughout the whole seed life span at the dry seed state. After 3 hours of imbibition, the transcript levels increased from dormant to non-dormant seed state, for both sites of the mRNAs. For the 3'-site, expression levels also increased when the seeds further deteriorated. There were no additional changes visible for the 5'-site of the mRNAs (Fig. 5C, 5D). This gene

is named *ARABIDOPSIS THALIANA SENESCENCE 1, ATSEN1, DARK INDUCIBLE 1, DIN1, SEN1, SENESCENCE 1, SENESCENCE ASSOCIATED GENE 1.* It is a senescence and ageing associated gene, responds to jasmonic acid and oxidative stress (Swarbreck *et al.*, 2007). For pattern 4 (*AT5G04750)* the transcriptome levels are constant whereas the translatome is reduced after 6 hours of imbibition (Bai *et al.*, 2020) (Fig. 5B). In the dry seeds, the expression levels dropped already after 13 days of EPPO treatment. From 13 days onwards until the aged state, the expression was slightly increased but was not significant compared with first two time- points. In imbibed seeds, the expression decreased along with the germination capacity during the seed life span curve, the lowest expression was reached after 115 days of EPPO, as identified for both sites of the mRNAs (Fig. 5C, 5D). *AT5G04750* encodes a T1F0-ATPase inhibitor protein which is involved in negatively regulating ATPase activity. It is located in various places including mitochondrion and the plasma membrane (Swarbreck *et al.*, 2007).

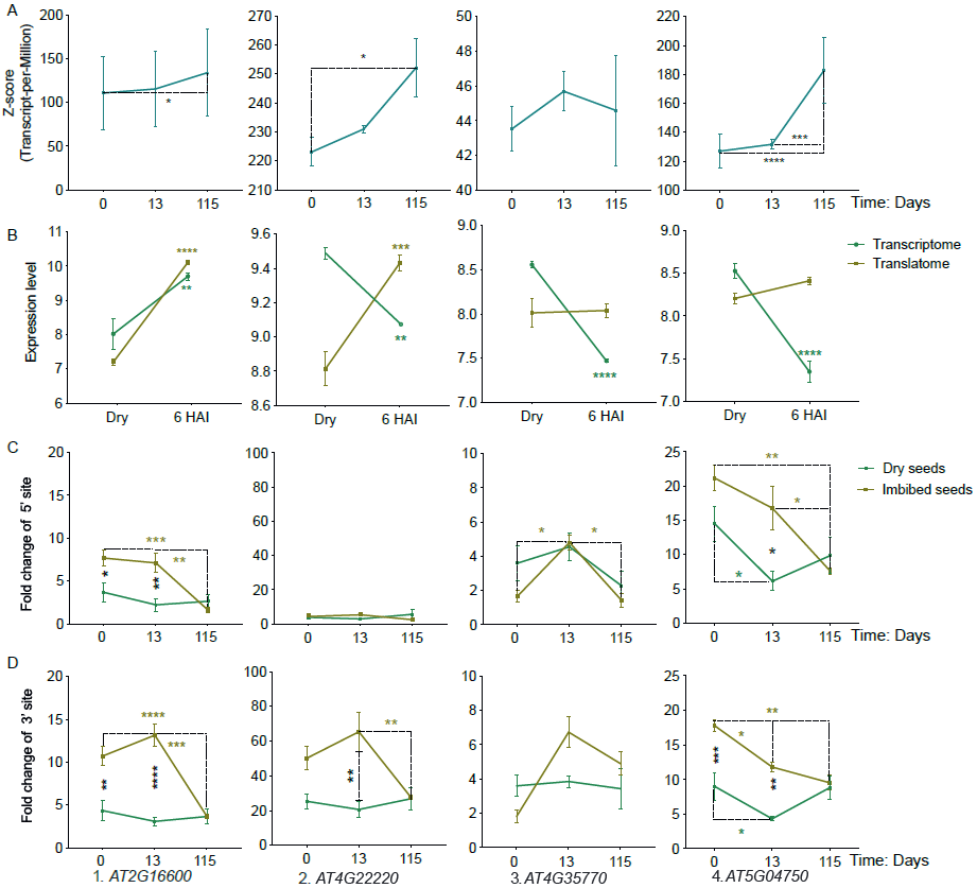


Figure 5. Transcript degradation as investigated by RT-qPCR. 1. *AT2G16600*; 2. *AT4G22220*; 3. *AT4G35770*; 4. *AT5G04750.* A, Degradation patterns of four example genes during seed ageing. Asterisks stand for the significant changes as indicated by dashed lines between the different EPPO treatments. B, Transcriptional and translational patterns of the selected gene in dry and 6 hours imbibed (HAI) seeds as published by Bai *et al.* (2020). Green asterisks represent significant changes for the transcriptome and dark yellow asterisks for the translatome. C and D respectively present the fold change of 5' and 3' sites of mRNAs in dry and 3 hours imbibed seeds at three time-points during seed ageing (0, 13 and 115 days of EPPO treatment). Green asterisks indicate the significant differences for the dry seed and dark yellow asterisks for imbibed seeds between the different physiological seed states. The horizontal lines refer to comparisons between the time-points, vertical lines and black asterisks refer to the significance between dry and imbibed seeds at the different time-points. The average represents relative mRNA levels ± standard errors of three replicates are presented. Statistical analyses were performed by two-way ANOVA. *: P ≤ 0.05, **: P ≤ 0.01, ***: P ≤ 0.001, ****: P ≤ 0.0001.

Seed storage proteins carbonylation accumulates during seed ageing

The level of protein carbonylation accumulates both during artificial seed ageing as concluded from experiments that used a controlled deterioration test (CDT) and during natural ageing in dry seeds, whereas it disappears after seed imbibition (Job *et al.*, 2005; Rajjou *et al.*, 2008). To investigate carbonylation patterns of storage proteins after EPPO treatment, total proteins were extracted from dry and 3 hours imbibed seeds after 0, 13 and 115 days of EPPO. Protein carbonylation patterns were detected based on the switch from 2,4-dinitrophenylhydrazine to form the 2,4-dinitrophenylhydrazone in carbonylated proteins (Levine *et al.*, 1990; Levine *et al.*, 1994). In dry seeds, there is an increase in carbonylated proteins during seed ageing (Figure. 6B). The α-cruciferin band which is visible already in the mature dry seeds became more intense over time, while the β-cruciferins and napins that have smaller molecular weights become carbonylated from 13 days of EPPO onwards. Notably, some high molecular weight non-storage proteins also become carbonylated after 115 days of seed ageing. A similar protein carbonylation pattern was also observed in 3 hours imbibed seeds (Fig. 6D).

Seed storage proteins were hypothesized to act as buffers against oxidative stress (Job *et al.*, 2005; Nguyen *et al.*, 2015). To analyse the role of seed storage proteins over natural seed ageing, the carbonylation patterns of the cruciferin *crua crub cruc* triple mutant (*cruabc*) and its wild type Col-0 were studied after seven years of natural ageing. The cruciferin triple mutant showed a significantly lower germination percentage than the wild type Col-0 seeds, however both still germinated to higher than 85 percent (Fig. 7A). The cruciferin triple mutant exhibited a stronger carbonylation in the β-CRU and NAPIN area of the blot, possibly caused by other proteins that are prone to carbonylation and have similar molecular weights. Also, here high molecular weight non-storage proteins are stronger carbonylated in the triple mutant seeds than in Col-0 (Fig. 7C).

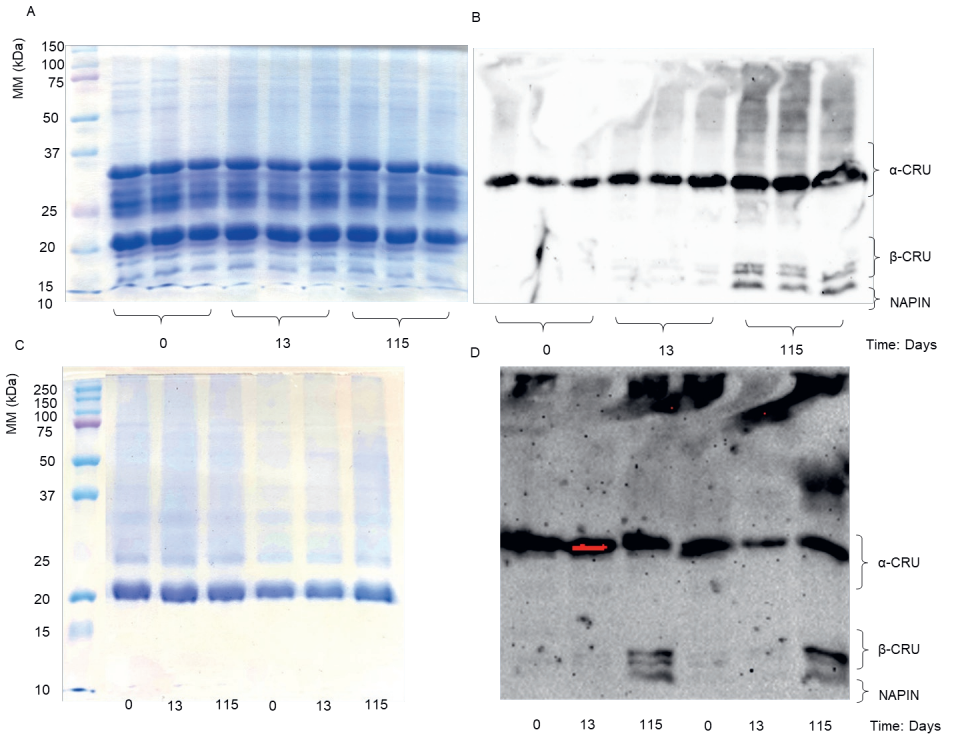


Figure 6. Protein carbonylation patterns of dry and imbibed Arabidopsis seeds using gel electrophoresis. A, Total protein loading control of dry seeds stained with Coomassie Brilliant Blue of dry seeds. B, Carbonylated proteins in dry seeds. Detected by immunoblotting using the anti-DNPH antibody, after derivatization of the carbonyl groups with 2,4-dinitrophenylhydrazine (DNPH). C, total protein of imbibed seeds stained with Coomassie Brilliant Blue of imbibed seeds. D, Anti-DNPH immunoblot of imbibed seeds. Seeds were treated with EPPO for 0, 13 and 115 days. Three and two biological replicates for each time-points are shown respectively in B and D. The protein bands were annotated based on Nguyen *et al.* 2015.

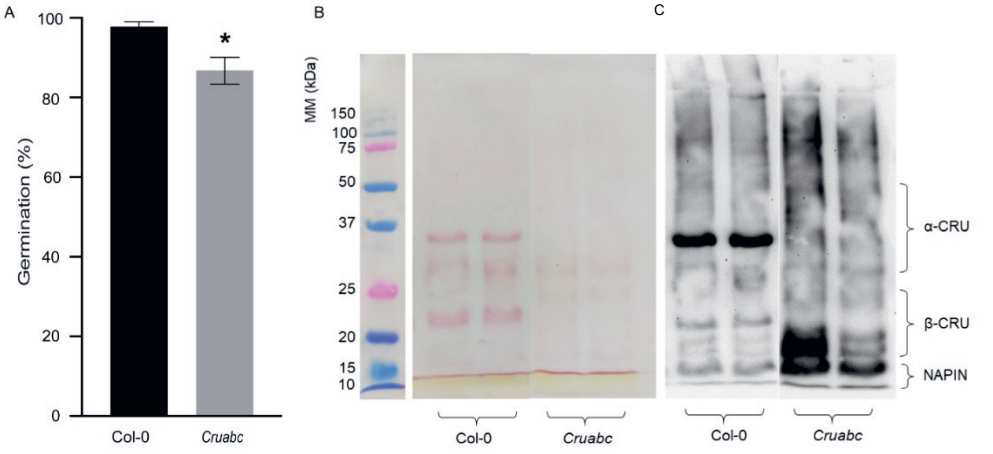


Figure 7. Protein carbonylation of seed storage protein in natural ageing seeds. A, Germination test for Col-0 and *cruabc*. B, Gel electrophoresis, stained with Ponceau staining, of total protein extract of the seeds. C, Carbonylated proteins of the same samples. It was detected by immunoblotting with anti-DNPH antibody, after derivatization is the carbonyl groups with 2,4-dinitrophenylhydrazine. Col-0 and triple cruciferin mutant abc (*cruabc*) seeds were both harvested in 2013 and naturally aged. Standard errors of four replicates are presented. Statistical analyses were performed by student t-test. *: $P \leq 0.05$.

Discussion

Seeds contain large amounts of stored molecules, including mRNAs and proteins. We have studied the integrity of these molecules to reveal whether this correlates with seed longevity and thus can explain why aged seeds cannot germinate. The importance of the protection of seed stored mRNAs for seed survival was hypothesized based on the finding that a large proportion of the mRNAs are associated with monosomes and translationally upregulated during early seed germination (Bai *et al.*, 2020).

Seed ageing using EPPO treatment

To study the effect of ageing on the integrity of seed stored molecules seeds have to be aged. Long storage times can be overcome by the use of artificial ageing techniques such as controlled deterioration test (CDT), accelerated ageing (AA) and EPPO (Buijs *et al.*, 2018; Buijs *et al.*, 2020; Demir *et al.*, 2008; Groot *et al.*, 2012; Gutierrez *et al.*, 1993; Powell *et al.*, 1981; Rodo *et al.*, 2003). CDT and AA make use of high temperature and high relative humidity. EPPO which is based on the increase of the absolute amount of oxygen can be performed at ambient relative humidity and temperature (Groot *et al.*, 2012; Rajjou, Loïc *et al.*, 2008; Tesnier *et al.*, 2002). EPPO treatments have been performed on seeds of several species, including barley, soybean, cabbage, lettuce and Arabidopsis. Recently, using a genetic approach it was shown that EPPO very well mimics dry storage in Arabidopsis. Quantitative trait locus (QTL) analysis after EPPO treatment led to the identification of both dormancy (*DELAY OF GERMINATION*; *DOG* loci) and ageing related QTLs (*GERMINATION ABILITY AFTER STORAGE*; *GAAS*) that were earlier identified in naturally stored dry seeds (Buijs *et al.*, 2018; Buijs *et al.*, 2020; Groot *et al.*, 2012; Nagel *et al.*, 2016). The overlap between loci identified by natural ageing and EPPO was much higher than that for seeds aged by CDT (40-60% relative humidity, 20-22°C) (Buijs *et al.*, 2018; Buijs *et al.*, 2020; Nguyen *et al.*, 2012). Moreover, the higher absolute amount of oxygen in the EPPO treatment contributes to studying the specific roles of oxygen during the seed life span. To control the ageing and to ensure that also seed dormancy release could be measured we have made used of EPPO at a controlled

relative humidity (55%) and temperature (20°C) (Buijs *et al.*, 2018; Groot *et al.*, 2012). The freshly harvested seeds were dormant, meaning that they did not germinate under conditions that favour seed germination. EPPO treatment for 13 days at 8MPa resulted in full after-ripening and increased the germination capacity of the seeds to 100%. A continuation of the EPPO treatment at 20MPa led to deterioration and a complete loss of the germination ability. The germination capacity seeds at different time-points during the EPPO treatment resulted in the typical seed life span curve (Fig. 1A).

Arabidopsis dry seed RNA integrity parameters are no read-out for seed germination ability

our study did not identify differences in integrity of total RNA extracted from dry seeds at different physiological stages, indicating that the RNA integrity of dry Arabidopsis seeds cannot be used to predict seed viability during seed ageing (Fig. 1B, 1C). This is in contrast to an earlier finding where for dry soybean seeds, which were collected and kept at 5°C and around 35% relative humidity from 1989 to 2015, a negative correlation between the integrity of the RNA and seed viability was shown (Fleming *et al.*, 2017). The differences between the study by Fleming *et al* (2017) and our analyses might be caused by the different environmental conditions, such as temperature and relative humidity. Another study by Fleming, M. B. *et al.* (2018) on the same batch of soybean seeds showed that the fragmentation level of mRNA which were extracted from embryonic axes correlated to the length of the transcripts, where longer transcripts show more fragmentation. This suggests that longer RNAs are more vulnerable for degradation whereas shorter mRNAs are more stable.

Seed degradome analyses reveal severe mRNA degradation during seed ageing

To more specifically investigate the integrity of mRNAs we have performed degradome analyses. Degradome analysis are widely applied to detect cleavage sites of degraded mRNAs by microRNAs (Karlova *et al.*, 2013; Pantaleo *et al.*, 2010; Xu *et al.*, 2012). Based on the selection of monophosphorylated 5' parts of 3'poly(A)

containing mRNAs, parallel analysis of RNA ends (PARE) sequencing for the identification of the fragments was performed (German *et al.*, 2009; German *et al.*, 2008).

In total, we found 16,342 5' monophosphorylated transcripts which represented seven different degradation patterns (Fig. 2). We also identified differentially expressed genes in each physiological state comparison during seed life span, for example, from fresh and dormant seed state to aged and dead seed state, 1008 up- and 940 down- regulated genes were respectively present (Fig. 4). Earlier it has been suggested that mRNA breakage and fragmentation occur randomly in aged Arabidopsis and soybean seeds (Fleming *et al.*, 2018; Zhao *et al.*, 2020). Our current analyses cannot confirm that since we have not yet analysed the mapping of the sequenced fragments on the genome. Moreover, we specifically isolated 5' monophosphorylated with intact 3' Poly (A) mRNAs which excludes the identification of mRNAs that lack the 3'part.

mRNA degradation during seed dormancy release

Based on the degradome expression patterns seven different clusters were identified (Fig. 2A). The dormancy release phase (from 0 to 13 days EPPO) is represented by cluster 1, 2, 4, 6 showing an elevated level of degradation at 13 day of EPPO compared to 0 days of EPPO (Fig. 2A). This implies that these transcripts become more degraded during after-ripening. To explore the degradation patterns of specific genes differential expression analyses were performed (Fig. 4). During dormancy release, comparing freshly harvested seeds to fully after-ripened non-dormant seeds, mostly transcripts show a reduction in degradation (meaning that the abundance was higher in the dormant stage compared to the non-dormant stage). We hypothesize that these mRNAs are remnants from seed maturation and that the pool of these mRNAs is depleted when seeds are after ripened. It could even be that these mRNAs have to be degraded before the seed is able to germinate, meaning that they inhibit seed germination. We have only investigated dry seeds, which are thought to be in a glassy and quiescent state lacking enzymatic activities (Buitink *et al.*, 2008; Sun *et al.*, 1994). Thus, the degradome that we identify is most likely a

consequence of mRNA turnover by non-enzymatic activities or environmental stimuli such as oxidation (Borbolis *et al.*, 2015). However, the degradome that we identify in dry dormant seeds could also be the result of degradation during seed maturation, a stage in which enzymatic reactions are still possible. Earlier the 5' to 3' decay pathway has been confirmed to be involved in Arabidopsis seed dormancy release. *Exoribonuclease4* (*xrn4*) and *varicose* (*vcs*) mutants showed definite dormancy phenotypes, transcriptome analysis performed on these two mutants discovered the degradation of certain specific mRNAs regulating seed dormancy (Basbouss-Serhal *et al.*, 2017). Hence, it would be interesting to select genes from our analysis and perform mutant analysis to confirm the roles in seed dormancy release. Moreover, mRNA degradation during dormancy release is consistent with results presented by Bazin *et al.* (2011) who evaluated the oxidation of stored mRNAs by analyzing the level of oxidative nucleic acid product 8-hydroxyguanosine (8-OHG) leading to protein translation alteration in sunflower seeds. For 24 mRNAs a positive correlation between increased mRNA oxidation and a reduced dormancy level was found (Bazin *et al.*, 2011).

mRNA degradation during seed deterioration

Half of the seed stored mRNAs are associated with monosome complexes and 17% of these (112 genes) are translationally upregulated during early seed germination, suggesting that a subset of the stored mRNAs is being protected by ribosomes to ensure translation during early germination (Bai *et al.*, 2020; Kimura *et al.*, 2010; Nakabayashi *et al.*, 2005; Rajjou *et al.*, 2004; Sano *et al.*, 2012). To explore this assumption, we analysed the appearance of the 112 monosome associated translationally upregulated mRNAs in the different clusters. Interestingly, 50% of these mRNAs occurred in clusters (1, 2 and 6) that represent mRNAs that accumulate in the degradome during seed ageing (Fig. 3). The monosome associated transcripts are also more pronounced in the degradation upregulated differentially expressed mRNAs compared to the down regulated genes. We consider transcripts that decrease in the degradome as being depleted and expect that these transcripts were more prone to degradation than the transcripts that

increase in the degradome and thus the latter could be considered to be more stable. In line with this theory our data suggests that the monosome associated transcripts are more stable, which would confirm the hypothesis that association to ribosomes could protect mRNAs (Bai *et al.*, 2020).

From dry to rehydrated seeds: can de novo transcription of seed stored mRNAs rescue seed germination

Upon seed imbibition de novo transcripts are synthesized (Bewley *et al.*, 2012). Whether the damage that has occurred to seed stored mRNAs is detrimental also depends on whether these transcripts can be produced during seed imbibition. To explore this, we have performed RT-qPCR for a set of 23 genes in dry and imbibed seeds at the three physiological seeds that were also investigated at the degradome level (Fig. 5). The degradation of RNAs mainly starts from three positions by three major enzymes, 5' exonuclease trims RNA from 5' site, 3' deadenylase shortens RNA from 3' site, and endonuclease cleavages RNA internally (Houseley *et al.*, 2009; Tourrière *et al.*, 2002). Thus, to be able to detect the whole mRNA we designed primers at both the 5' and 3' site of the mRNA. In dry seed states, transcripts expression levels were stable overall in seed life span (Fig. 5C, 5D). This seems in contrast to the degradation patterns that we earlier described but can be explained by the fact that broken mRNAs and de-capped mRNAs were reversely transcribed to cDNAs and thus amplified. Although most of the genes showed an increased expression in the imbibed seeds when compared to dry seeds this expressed ceased in the aged seeds (Fig. 5C, 5D). A clear example is *AT2G16600*, at 3 hours imbibition its transcript levels are significantly higher than in dry seed in freshly harvested (0 day EPPO) and fully after-ripened non dormant seed (13 days EPPO), however not in aged, dead seeds (115 days EPPO) (Fig. 5C, 5D). This suggests that *At2G16600* might be degraded by oxidation and mRNA surveillance or *de novo* transcription could not ensure the presence of the transcript and thus the contribution of this gene to the seed germination. Likely it is not the lack of a single gene that causes the lack of germination and it cannot be excluded that other factors e.g., cell wall integrity, other cellular damage or a combination of factors are the actual reason that the seed

is not able to germinate anymore. The importance of certain mRNAs can be investigated using mutants or overexpression lines for their seed longevity phenotypes.

Seed storage proteins act as buffers against oxidative stress during seed ageing in natural and artificial ageing systems

During seed maturation seeds become desiccation tolerant and accumulate storage compounds (Goldberg *et al.*, 1994). In Arabidopsis seeds, the major seed storage proteins (SSPs) are 2S and 12S proteins that together represent around forty percent of the dry seed weight (Baud *et al.*, 2002; Mansfield *et al.*, 1992). Because of their abundance SSPs serve as a main target of oxidation (Davies, 2005) and especially are carbonylated during Arabidopsis and pea seed germination (Barba-espin *et al.*, 2011; Job *et al.*, 2005). The two main SSPs, cruciferins (12S globulins) and napins (2S albumins), have been reported to play a role in seed longevity. The cruciferin triple mutant (*cruabc)* and the *NAPIN* RNAi line exhibited shorter longevity compared to wildtype Col-0 (Nguyen *et al.*, 2015). Furthermore, it was shown that SSPs, mainly cruciferins, are carbonylated by oxidation and this has been proposed to be important to protect other essential proteins from oxidative damage (Job *et al.*, 2005; Nguyen *et al.*, 2015; Wan *et al.*, 2007). We show that the protein carbonylation pattern in dry seed state after both EPPO treatment and seven years of natural ageing was consistent with the study by (Nguyen *et al.*, 2015). The level of oxidized cruciferins and napins increased during ageing, confirming the protective role of these proteins (Fig. 6B, 7C). Although most seed storage proteins have disappeared during seed imbibition, still severe carbonylation of α-CRU was detected (Fig. 6D). This is in contrast to earlier studies that described the disappearance of SSPs carbonylation and other protein carbonylation during imbibition (Job *et al.*, 2005). The differences between our and earlier studies might be based on the imbibition time and conditions. Moreover, the identification of the carbonylated proteins in the SSP mutants could reveal which protein might be essential for germination and thus need to be protected during seed storage.

Conclusions and perspectives

We found that RNA integrity is not associated with seeds viability when seeds are kept at dry state. We identified severe mRNA degradation during seed storage and our data suggests that the degradation of mRNAs that are known to be associated to ribosomes in dry seeds correlated with lack of germination in aged seeds. Phenotypic analyses in mutants that lack these mRNAs should provide more information on whether these mRNAs are indeed essential for seed germination. In addition, it would also be interesting to see where in the cell these mRNAs are stored and whether in addition to ribosomes other proteins contribute to their protection. Besides the functional analyses of specific genes more information can also be obtained from the degradome data analyses. The information that is present in the data set is ample and so far, only a fraction of this is analysed. Additional analyses could include motif, guanine-cytosine (GC) content and Gene Ontology analyses on both the clusters and the differentially expressed mRNAs. Motif analysis could lead to the identification of specific sequence motifs in the mRNAs that are involved in recruiting certain specific RNA-binding proteins that act as a protective structure during storage. Gene Ontology analyses can help to understand the roles of transcripts during seed germination or other stages during the seed life span. As known, the mRNA degradation pathway needs enzymatic activity, but degradation pattern analysis in our study is performed in dry seed state which lacks active enzymes. Therefore, it might be interesting to explore the specific interaction between mRNA and oxidation which results in seed death eventually.

Materials and Methods

Plant materials and growth conditions

The mutant *dog6-2* (*SALK_012554*, T-DNA insertion in the second intron region) originates from NASC (the Nottingham Arabidopsis Stock Centre).

After-ripened seeds were grown on 4 X 4 cm Rockwool blocks in a climate room at 20 ℃/ 18℃ (day/ night) under a 16hs artificial light (150 µmol m^{-2} s^{-1}) photoperiod and 70% relative humidity. Plants were watered with a standard nutrient solution and three times per week.

Seed germination experiments

Germination test were derived from (Joosen *et al.*, 2010) for a period of 7 days. Basically, in each plastic tray (15 x 21 cm), 2 layers of blue paper with 48 mL demi water, 50-100 seeds were sown using a plastic sheet with 6 round holes for the accurate position. After sowing the trays were piled and packed into transparent plastic bags and placed into an incubator with 22℃, 30 W m^{-2} continuous light.

Germination parameters

Each day, pictures were taken at 2 times at 9 AM and 5PM for 7 days. Germinator package was used to analysis germination (Joosen *et al.*, 2010). GMAX: maximum germination percentage was applied for describing the germination ability. The germinator curve-fitting script makes it possible to statistically compare the significance among the samples by Student's t-test.

Seed longevity experiments under EPPO treatment

Seeds were harvested and the initial germination percentage was investigated before EPPO treatment, it was confirmed that at 0-day seed samples were dormant. In total, eight 1.5L steel tanks and 32 tubes were prepared for releasing seed dormancy and reducing seed longevity during several periods. Seeds were aliquoted to 50mg and stored into tubes with a lid that contains holes, one piece of paper was used between seeds samples and lid to avoid seeds dispersal. The samples were placed into tanks with a controlled relative humidity (55%) by adding dry silica gel.

The pressure in the tanks, initially 8MPa was used for releasing seed dormancy, afterwards, 20MPa was applied for ageing seeds. After different storage periods, samples were taken out and compared the germination test until seeds did not germinate anymore (Groot *et al.*, 2012).

Candidate genes selection and primers design for gene expression analysis by RT-qPCR

The selection was based on (Bai *et al.*, 2020), after 6 hours imbibition, 112 stored monosomes associated mRNAs translationally up regulated, among these, 23 genes expression level were high enough to be detected, including Heat Shock Proteins (HSPs), Late Embryogenesis Abundant (LEA) proteins, Calcium Binding Proteins, Proline and Glycine rich proteins and some unknown transcripts. The Arabidopsis Information Resource (TAIR) website and Arabidopsis eFP browser were also used to confirm the high expression level for dry and 3 hours imbibed seeds. The list of genes could be found in Supplemental table 4.

The primers were designed by software Primer3Plus (Untergasser *et al.*, 2007). Two pairs of primers were designed for each candidate gene, one pair of primer amplification position was close to the ATG site, the other one was near the TAA site. Both of the primers' efficiencies were tested and compared significance between each other, the list of primers could also be found in Supplemental table 5.

cDNA synthesis and qPCR analysis

iScript cDNA synthesis kit (Bio Rad) was used for making cDNAs, in this project, 1 µg RNA was used from each sample. cDNA samples were one – tenth diluted. qPCR was done and analysed according to qPCR guide eurogentec. The master mix was 10 µl for each: 2.5 µl cDNAs, 5 µl SYBR green, 1 µl primers mixtures, 1.5 µl MQ water. The CFX Bio Rad was used to generate threshold cycle value for each reaction. Excel and Prism 8 software were used to analysis qPCR result.

Preparation of library and sequencing

Dry Arabidopsis seeds were frozen and ground in liquid nitrogen, and RNA extraction from dry seeds was executed using a modified protocol which is detailly described

in Supplemental protocol 1. Purification of 5' monophosphorylated with 3' poly(A) RNA was prepared for the PARE library by Vertis Biotechologie AG (Freising, Germany) following a modified protocol of (German *et al.*, 2009; German *et al.*, 2008) for high throughput sequencing. The library pool was fractionated in a size range of 200-600 bp. The primers used for PCR amplification were designed for TruSeq sequencing according to the instructions of Illumina. The cDNA pool was sequenced on an Illumina NextSeq 500 system using 75 bp read length.

Degradome sequencing analysis and gene expression analysis

The quality of the raw reads was assessed for each sample using FastQC v0.11.5 (Andrews, 2010). rRNA depletion was performed with SortMeRNA v4.1.0 to filter rRNA fragments from the degradome reads datasets (Kopylova *et al.*, 2012). After filtering, the degradome data was quantified using RNA-Seq by Expectation Maximization v1.3.1 (RSEM) (Li *et al.*, 2011). The Araport11 reference cDNA gene model was prepared for mapping with RSEM-prepare-reference, the reads were subsequently mapped to the reference genome using a bowtie2 plugin of RSEM-calculate-expression for strand specific mapping which is able to deal with ambiguously mapped reads (Cheng *et al.*, 2017; Langmead *et al.*). Expected counts were estimated as the sum of posterior probabilities of the maximum likelihood generated model using an expectation-maximization algorithm (RSEM). Mapped reads were then (de) compressed and checked using SAMtools and Integrative Genome Viewer (IGV) (Li *et al.*, 2009; Robinson *et al.*, 2011). Hierarchical clustering with complete linkage was performed to visualize distinct degradation kinetics over the course of EPPO treatments. To filter out genes that were very low in abundance, filtering of the gene set was performed to only keep genes that have a minimum of 3 reads per sample. In hierarchical clustering, transcript-per-million (TPM) abundance estimates by RSEM were used. Differential expression analysis was done in DESeq2 v1.28.1 (Love *et al.*, 2014). DESeq2 worked by first estimating size factors to correct for the library size (Anders *et al.*, 2010). Dispersion estimates were then calculated for fitting the model and counts were shrunk towards the fitted curve (McCarthy *et al.*, 2012). Differential expression was tested in a negative binomial

generalized linear model (negative binomial Wald test) (Love *et al.*, 2014). Results were extracted for *p*-value significance adjusted for false discovery rate cut-off of 0.05 with the Benjamini-Hochberg approach. No log-fold-change threshold was set, and a Cook's distance cut-off was used (Benjamini *et al.*, 1995; Cook, 1977).

Protein carbonylation setup

The full protocol is described in supplemental protocol 2.

Supporting Information

Supporting information can be downloaded by scanning the QR code or from http://www.wageningenseedlab.nl/thesis/ssong/SI/

Supplemental Figure 1. Expression of selected mRNAs both in dry and 3 hours imbibed seeds.

Supplemental Table 1. List of 112 monosome bounded mRNAs.

Supplemental Table 2. Overlap of monosome associated mRNAs and the mRNAs of the different clusters.

Supplemental Table 3. List of differentially expressed mRNAs in each physiological seed state comparison during seed life span.

Supplemental Table 4. List of 23 selected monosome bound mRNAs and the function descriptions.

Supplemental Table 5. List of primers used in RT-qPCR analysis.

Supplemental Protocol 1. Optimized RNA extraction protocol from dry Arabidopsis seeds.

Supplemental Protocol 2. Protein carbonylation setup.

Chapter 5

General Discussion

The world population is predicted to be increased to around 12.3 billion by the end of the 21st century. This rapid population growth strongly correlates with a higher demand of the food supply system. The success of food production is largely depended on how seeds are produced and stored, this is especially important for the daily food-cereals. Therefore, improving global cereal production has been one of the main ways to face the global expanding demand for food (Dyson, 1996; Gerland *et al.*, 2014; Gilland, 2002; Godfray *et al.*, 2010). Seeds are not only used as commodities but are also the starting material for producing cereals and growing vegetable crops. Seed dormancy and seed longevity are known to be two of the most vital features for seed quality which determine crop yield eventually (Colville *et al.*, 2019; Martínez-Andújar *et al.*, 2012). Seed dormancy is a quiescent phase that avoids germination of a mature seed. It determines the uniformity of seed germination which is beneficial for better management either for smallholder famers or large breeding companies. The preferred seed dormancy level depends on the specific species and on the geographical area in which the species is grown. Seed longevity is primarily referring to the time that seeds remain viable after storage. Depending on the purpose seeds should remain viable for a certain time. For regeneration by farmers seeds should at least survive until the next growing season. For gene banks which aim to safeguard plant genetic variation, viability should be extended much longer. This will largely reduce the frequency of seed vigor monitoring and thus reduce the costs spend on testing the seeds and the human labor that is required for these analyses. Longer viability also reduces the potentially loss of genetic information that might occur when the seed batches are renewed (Copeland *et al.*, 1999; Hay *et al.*, 2017; Rajjou *et al.*, 2008; Walters *et al.*, 2005). In order to understand more about seed dormancy and longevity, this thesis focusses on understanding regulatory mechanism of seed dormancy and longevity in the model plant Arabidopsis. The final goal is to translate the achieved knowledge to crops, to allow crop improvement in the future.

The presence or absence of the transmembrane domain in NAC genes

In my thesis I studied the functional analyses of *ANAC060* and its two nearest homologs *ANAC040* and *ANAC089* (**Chapter 2 and 3**). These three NAC transcription factors (TFs) are highly homologous but have distinct functions. All of them contain a conserved NAC domain and diverse C-terminal that determines the absence or presence of the transmembrane domain (TMD) (Ooka *et al.*, 2003). Sequencing of the *ANAC060* allele in different accessions including Landsberg *erecta* (L*er*), Columbia-0 (Col-0), Antwerp-1 (An-1) and Tacoma (Tac) revealed polymorphisms that result in a lack of the TMD, whereas the *ANAC060* alleles of Santa Maria da Feira-1 (Fei-1), Calver (Cal), Shakdara (Sha), Cape Verde Islands (Cvi), Kashmir and C24 encode the TMD. The causes of the loss of the TMD are not uniform, for instance, the L*er* genomic sequence has a deletion of 482 bp that encodes the TMD. Col-0 carries a point mutation in third intron affecting *ANAC060* mRNA splicing patterns which results in a short form of cDNA without the TMD (**Chapter 2**)(Li *et al.*, 2014). The presence and absence of the TMD determine whether the protein is localized in nuclear membrane or in the nucleus respectively. The localization of the protein determines if it is able to activate transcription and thus may lead to opposite phenotypes as has been indicated for several processes (**Chapter 2**)(Kim *et al.*, 2008; Li *et al.*, 2014; Yu *et al.*, 2020). Genetic variation that determines the presence or absence of the TMD in genomic sequences is not unique for *ANAC060*. The *ANAC060* homolog *ANAC089* has been reported to show a similar membrane-mediated regulation in the accessions L*er* and Cvi. The *ANAC089* L*er* allele encodes an identical polypeptide to the Col-0 allele containing the full length TMD. Seedlings containing the ANAC089 L*er* protein, which is localized in cytoplasm, show a fructose sensitive phenotype. In contrast, the *ANAC089* Cvi allele containing a one base pair deletion, that leads to a premature stop codon in third exon, is present in the nucleus. This allele leads to an insensitivity to high concentrations of fructose. Moreover, the one bp deletion in Cvi is identified by investigating the sequence collection of Arabidopsis 1001 genome project, showing the potential of genome sequencing in identifying sequence polymorphisms between

accessions (Li *et al.*, 2011; Weigel *et al.*, 2009). For the other *ANAC060*, homolog *ANAC040*, no allelic differences have been reported so far. Sequence information as present in the 1001 genome database (Weigel *et al.*, 2009) could be investigated to reveal whether also the localization of *ANAC040* is determined by naturally occurring polymorphisms. Identifying genetic variation in the genomic sequence of NAC genes, which are known to regulate developmental and stress responses, is interesting since this provides a mechanism by which plants have adapted themselves to their environment during the evolution.

Cleavage of the TMD has shown to be induced by environmental factors and thus provide another level at which plants are able to respond to the environment, increasing their chance to survive. It is interesting to discuss the role of the TMD a bit more. Why do TFs contain a TMD if this inhibits their function as a transcription factor? The fact that these genes exist suggests that the protein eventually ends up in the nucleus. How does this occur? The activation of membrane bound TFs have been described and hypothesized to be regulated by proteolytic cleavage including ubiquitin/proteosome-dependent processing and intramembrane proteolysis, both of them are predicted to be strategies to quickly respond to environmental stresses (Hoppe *et al.*, 2001; Seo, Pil *et al.*, 2008; Vik *et al.*, 2000). Based on this theory, I performed several experiments to determine whether the ANAC060 protein could translocate from the membrane into the nucleus. For this the GFP tagged *ANAC060* allele of Kashmir, that would retain the protein from the nucleus, was used. The alleles lacking the TMD were none dormant and thus we hypothesized that the ANAC060 protein was required in the nucleus to inhibit seed dormancy. Treatments with gibberellic acid, KNO_3, H_2O_2 and cold stratification that were known to release seed dormancy (Koyuncu, 2005; Liu *et al.*, 2010; Rosbakh *et al.*, 2019; Tang *et al.*, 2008) and an additional treatment with dithiothreitol, that had shown to cause the translocation of ANAC089 from the membrane to nucleus, did not allow us to detect ANAC060 in the nucleus (data not shown) (Klein *et al.*, 2012). We could however not exclude that the protein was translocated to the nucleus. The nuclear form of the protein was dominant over the membrane bound form as was confirmed by analyses

using L*er* and the *DOG6* Kashmir near isogenic line (**Chapter 2**) and transgenic plants containing Cvi ANAC089 protein in L*er* background (Li *et al.*, 2011). It could be that low, not detectable amounts of the protein are sufficient to activate downstream targets. Another reason for not detecting the protein in the nucleus could be that the protein becomes instable because of the loss of the TMD (Flecha, 2017). The evidence of dynamic protein trans localization of ANAC089 after one hour of dithiothreitol treatment also confirms that environmental stimuli could cause cleavage of TMD from the transcription factor. Once in the nucleus this protein acts as a negative regulator of stromal ascorbate peroxidase (Klein *et al.*, 2012). For ANAC040 there is no evidence about its dynamic translocation. It seems that cues for releasing NAC TFs from the membrane are specific for different transcription factors, indicating that specific environmental stresses regulate these TFs.

The potential of the elevated partial pressure of oxygen (EPPO) treatment to study seed dormancy and seed longevity

The elevated partial pressure of oxygen (EPPO) method is a system that shortens seed life span and is invented by Groot *et al.* (2012). Briefly, EPPO increases the oxygen concentration surrounding the seeds while it keeps the seeds in the dry state, thus avoiding enzymatic activity (Fernández-Marín *et al.*, 2013; Groot *et al.*, 2012). As a result, the release of seed dormancy and loss of seed viability are stimulated. Compared to other artificial ageing techniques it is the best artificial ageing system to mimic natural ageing identified so far (Buijs *et al.*, 2018; Buijs *et al.*, 2020). I have used the EPPO method obtain seeds that represent different stages of the bell-shaped seed life span curve (Fig. 1). This has allowed us to address the question if mRNA degradation correlates with a loss of seed viability during ageing (**Chapter 4**). However, the use of EPPO could advance our understanding of seed storage much more. For example, the process of after-ripening, that is known to release seed dormancy is relatively unknown. Using EPPO solely the effect of oxidation on after-ripening could be investigated. A previous study in our lab revealed that transcriptome analyses comparing dormant to after-ripened seeds hardly revealed differences, however already after three hours of rehydration differences in the

transcriptome occurred (Dekkers *et al.*, 2016). These differences are the result of phenomena that occurred during seed dry storage. It has been suggested that these are caused by reactive oxygen species (ROS) that are formed during seed dry storage (Bewley *et al.*, 2013; Job *et al.*, 2013). The release of seed dormancy has been associated with the accumulation of 8-oxo-7,8-dihydroguanine (8-OHG) in specific transcripts in sunflower (Bazin *et al.*, 2011). This positive relationship between the accumulation of 8-OHG has been not found for total RNA during dry storage of soybean seeds (Fleming *et al.*, 2018). We analyzed the total RNA integrity and found that the RIN did not change during seed storage (**Chapter 4**). This might indicate that seed dormancy is more affected at the mRNA level then at the total RNA level, another explanation is that the effect is species depended. It will be interesting to check the level of 8-OHG of mRNAs in dormant and after-ripened Arabidopsis seeds. This will also allow to investigate whether mRNA degradation is linked to 8-OHG accumulation. Noticeably, we have used *anac060-2* mutant seeds for our degradome analyses. Performing similar analyses in wild type Col-0 EPPO after-ripened seeds would allow to investigate the downstream targets of *ANAC060*.

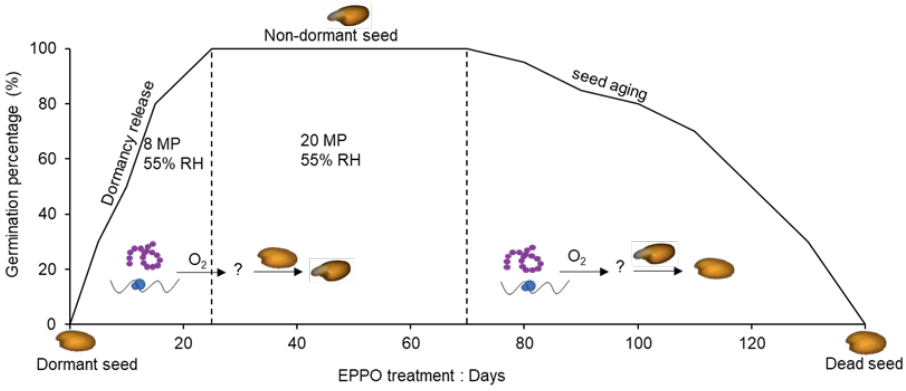


Figure 1. Schematic presentation of the seed life span curve. Germination percentage of seed dormancy release and ageing under EPPO conditions. During dormancy release, the dormant seeds become to be non-dormant seeds under 8 MPa, 55% relative humidity (RH), in parallel, stored proteins and mRNAs are oxidized with a non-specific pattern. From non-dormant seeds to dead seeds, the

pressure is elevated to 20 MPa but with same RH, at the meantime, stored proteins and mRNAs are surrounded with a much higher concentration of oxygen, being oxidized at another non-specific pattern.

In addition to studying after-ripening, EPPO treatment has a great potential to address seed longevity related questions. So far, we have only focused on the degradation of mRNA (**Chapter 4**). These analyses have allowed us to identify mRNAs that are critical for seed longevity, however further exploration of this data is necessary, for example, it will be encouraging to find a seed ageing marker for seed research. Such a marker could be used to better predict the seed viability before seeds exhibit ageing symptoms such as decreased seed germination percentage and thus make it possible to observe different levels of seed vigor at any time during seed life span without doing germination test. Moreover, it would be interesting to explore whether the aged seeds are truly dead or whether there is still synthesis of new proteins from stored mRNAs or even from newly synthesized mRNAs. One could perform polysome profiling on imbibed aged seeds to determine whether translation occurs (Bai *et al.*, 2017; Bai *et al.*, 2020). Except analyzing the mRNA degradation pattern we have also detected patterns of protein carbonylation, which is a form of protein oxidation by ROS (Suzuki *et al.*, 2010). Due to the high abundance of seed storage protein (SSP) and hardness of protein carbonylation detection technique, we could only confirm that SSP functioned as an oxidative buffer during seed germination and seed ageing (Job *et al.*, 2005; Nguyen *et al.*, 2015). It will be promising to find a way to identify and investigate specific essential germination related protein carbonylation patterns during whole seed life span. However, oxidative damage in mRNA and proteins might not be the only reasons for the reduced seed viability. In dry seeds, also DNA is being oxidized by ROS, causing single and double strand breakages. It also has been indicated that ROS modifies guanine to 8-oxogunine (8-oxoG) in DNA during ageing (Boesch *et al.*, 2011; Chen *et al.*, 2013; Johnston *et al.*, 2010; Walters *et al.*, 2006). In Arabidopsis, the overexpression of AtOGG1, which encodes a DNA glycosylase/apurinic/apyrimidinic (AP) lyase, could erase 8-oxoG from DNA and result in longer seed longevity after a controlled deterioration treatment (CDT) and better germination ability under

various stress conditions after imbibition (Chen *et al.*, 2012). Besides, the decrease of 5-methylcytosine (m^5C) is positively correlated with reduced seed germination of the desiccation sensitive *Acer. Pseudoplatanus* L. seeds (Plitta-Michalak *et al.*, 2018). It is not only the seed's genetic material that is affected during ageing. Also, the increase of lipid oxidation leads to seed vigor loss as has been shown in wheat, barley and soybean and to programmed cell death in Arabidopsis (Stewart *et al.*, 1980; Wiebach *et al.*, 2020; Zoeller *et al.*, 2012). Cell membrane permeability and integrity are quantified by checking the level of electrical conductivity, amino acids, water soluble sugars, α-tocopherol (antioxidants), phospholipids, lipid peroxidation respectively in *Azadirachta indica* seeds which are stored for various times from different aged mother trees. These analyses show that the indicators for cell membrane permeability are higher in freshly harvested seeds than the ones are stored for longer time. Cell integrity indicators such as phospholipids and α-tocopherol decline with increased seed storage times (Kumar *et al.*, 2014). Overall, these compounds that are related to the seed ageing could be other potential research targets to investigate in EPPO aged seeds.

Transferring knowledge from the model plant Arabidopsis to crops

Arabidopsis as a model plant has a great value for research (Koornneef *et al.*, 2010; Somerville *et al.*, 2002), however, it does not have direct benefit for agriculture and economy. With the improvement of genome sequencing, rice, wheat, maize are becoming to be the new model plants. Generally, transferring the knowledge from orthologous genes that share function, and thus have a common ancestor, from Arabidopsis to crops is preferred (Spannagl *et al.*, 2011). However, this is not always applicable even within species. We have analyzed the three homologous genes *ANAC060*, *ANAC040* and *ANAC089* evolutionarily and through swapping experiments (**Chapter 3**). We observed that functional redundancy was not absolutely correlated to the high identity of genes. *ANAC089,* for example, has a 64% identity to *ANAC060* and both were derived form At-Alpha duplication event, yet there was no functional redundancy between these two genes. Whereas, functional

redundancy was identified between *ANAC060* and *ANAC040* even though they had a lower sequence identity (41%).

Nethertheless, the knowledge that could be delivered from model plants still has a great potential. Quantative trait loci (QTL) analysis to identify dormancy releated loci are not only performed on Arabidopsis (Alonso-Blanco *et al.*, 2003; Bentsink *et al.*, 2010; Kearsey *et al.*, 1998; van der Schaar *et al.*, 1997), but also for crops such as, rice, wheat and barley (Guo *et al.*, 2004; Osa *et al.*, 2003; Sato *et al.*, 2009). Identifying the underlying genes in these crops might even be more laborious than in Arabidopsis, however can be advanced by comparative genome mapping and synteny analyses (McCouch, 2001). For example, genomic regions containing dormancy QTL in different species could be compared to see whether they encode similar genes. Moreover, dormancy regulators like the NAC membrane bound transcription factor *ANAC060* (**Chapter 2)** could be regulating seed dormancy also in crops. NACs are known for their conserved gene structure, for example, there are around 140 NAC TFs in rice and tomato (Fang *et al.*, 2008; Kim *et al.*, 2010; Kim *et al.*, 2007; Su *et al.*, 2015). How these NAC TFs relate to *ANAC060* and whether they confer a similar function could be studied by combing phylogenetic analysis with synteny network analysis, like that has been performed for MADS box genes (Zhao *et al.*, 2017; Zhao *et al.*, 2019).

In our laboratory we use the EPPO treatment to study seed dormancy and seed longevity, with the aim to identify the mechanisms by which these processes are regulated in Arabidopsis (**Chapter 4**) (Buijs *et al.*, 2018; Buijs *et al.*, 2020; Nguyen *et al.*, 2012). EPPO was initially applied on barley, cabbage, lettuce and soybean seeds (Groot *et al.*, 2012), additional analyses in these species, for example a degradome analyses like we performed in **Chapter 4** might help to identify ageing markers that can be used to improve seed quality.

Moreover, based on phenotypic analyses, plant growth regulators tested in Arabidopsis are shown to have a high translatability to different crop species, such as lettuce, carrot, tomato (Rodriguez-Furlán *et al.*, 2016). Drought stress and survival strategies are also transferable between Arabidopsis and tomato (Krukowski *et al.*,

2020). In conclusion, even though it is often a long way to transfer knowledge from the model plant Arabidopsis to crops, it might bring huge benefits to global agriculture once it works.

References

Aas, P. A., Otterlei, M., Falnes, P. Ø., Vågbø, C. B., Skorpen, F., Akbari, M., . . . Seeberg, E. (2003). Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature, 421*(6925), 859-863.

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H., & Tasaka, M. (1997). Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. *The Plant Cell, 9*(6), 841-857.

Alonso-Blanco, C., Aarts, M. G., Bentsink, L., Keurentjes, J. J., Reymond, M., Vreugdenhil, D., & Koornneef, M. (2009). What has natural variation taught us about plant development, physiology, and adaptation? *Plant Cell, 21*(7), 1877-1896. doi:10.1105/tpc.109.068114

Alonso-Blanco, C., Bentsink, L., Hanhart, C. J., Blankestijn-de Vries, H., & Koornneef, M. (2003). Analysis of natural allelic variation at seed dormancy loci of *Arabidopsis thaliana*. *Genetics, 164*(2), 711-729.

Alonso-Blanco, C., & Koornneef, M. (2000). Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. *Trends in plant science, 5*(1), 22-29.

Alvarado, V., & Bradford, K. (2002). A hydrothermal time model explains the cardinal temperatures for seed germination. *Plant, cell & environment, 25*(8), 1061-1069.

Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Nature Precedings*, 1-1.

Andrews. (2010). FastQC: a quality control tool for high throughput sequence data. In: Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Apel, K., & Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol., 55*, 373-399.

Arc, E., Galland, M., Cueff, G., Godin, B., Lounifi, I., Job, D., & Rajjou, L. (2011). Reboot the system thanks to protein post-translational modifications and proteome diversity: how quiescent seeds restart their metabolism to prepare seedling establishment. *Proteomics, 11*(9), 1606-1618.

Arc, E., Sechet, J., Corbineau, F., Rajjou, L., & Marion-Poll, A. (2013). ABA crosstalk with ethylene and nitric oxide in seed dormancy and germination. *Frontiers in Plant Science, 4*, 63.

Bahin, E., Bailly, C., Sotta, B., Kranner, I., Corbineau, F., & Leymarie, J. (2011). Crosstalk between reactive oxygen species and hormonal signalling pathways regulates grain dormancy in barley. *Plant, cell & environment, 34*(6), 980-993.

Bai, B., Peviani, A., van der Horst, S., Gamm, M., Snel, B., Bentsink, L., & Hanson, J. (2017). Extensive translational regulation during seed germination revealed by polysomal profiling. *New Phytologist, 214*(1), 233-244.

Bai, B., van der Horst, S., Cordewener, J., America, T., Hanson, J., & Bentsink, L. (2020). Seed stored mRNAs that are specifically associated to monosome are translationally regulated during germination. *Plant physiology*, pp. 00644.02019.

Bailly, C. (2004). Active oxygen species and antioxidants in seed biology. *Seed Science Research, 14*(2), 93-107.

Bailly, C. (2019). The signalling role of ROS in the regulation of seed germination and dormancy. *Biochem J, 476*(20), 3019-3032. doi:10.1042/BCJ20190159.

Bailly, C., El-Maarouf-Bouteau, H., & Corbineau, F. (2008). From intracellular signaling networks to cell death: the dual role of reactive oxygen species in seed physiology. *Comptes rendus biologies, 331*(10), 806-814.

Barba-espin, G., DIAZ-VIVANCOS, P., Job, D., Belghazi, M., Job, C., & HERNÁNDEZ, J. A. (2011). Understanding the role of H2O2 during pea seed germination: a combined proteomic and hormone profiling approach. *Plant, cell & environment, 34*(11), 1907-1919.

Basbouss-Serhal, I., Leymarie, J., & Bailly, C. (2016). Fluctuation of Arabidopsis seed dormancy with relative humidity and temperature during dry storage. *Journal of experimental botany, 67*(1), 119-130.

Basbouss-Serhal, I., Pateyron, S., Cochet, F., Leymarie, J., & Bailly, C. (2017). 5′ to 3′ mRNA decay contributes to the regulation of Arabidopsis seed germination by dormancy. *Plant physiology, 173*(3), 1709-1723.

Baskin, C. C., & Baskin, J. M. (1998). *Seeds: ecology, biogeography, and evolution of dormancy and germination*: Elsevier.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 3, No. 1).

Baud, S., Boutin, J.-P., Miquel, M., Lepiniec, L., & Rochat, C. (2002). An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. *Plant Physiology and Biochemistry, 40*(2), 151-160.

Baud, S., Dubreucq, B., Miquel, M., Rochat, C., & Lepiniec, L. (2008). Storage reserve accumulation in Arabidopsis: metabolic and developmental control of seed filling. *The Arabidopsis book/American Society of Plant Biologists, 6*.

Bäumlein, H., Miséra, S., Luerßen, H., Kölle, K., Horstmann, C., Wobus, U., & Müller, A. J. (1994). The FUS3 gene of *Arabidopsis thaliana* is a regulator of gene expression during late embryogenesis. *The Plant Journal, 6*(3), 379-387.

Baxter, I., Brazelton, J. N., Yu, D., Huang, Y. S., Lahner, B., Yakubova, E., . . . Salt, D. E. (2010). A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS Genet, 6*(11), e1001193.

Bazin, J., Langlade, N., Vincourt, P., Arribat, S., Balzergue, S., El-Maarouf-Bouteau, H., & Bailly, C. (2011). Targeted mRNA oxidation regulates sunflower seed dormancy alleviation during dry after-ripening. *The Plant Cell, 23*(6), 2196-2208.

Belletti, P., Lanteri, S., & Lotito, S. (1991). The influence of temperature and moisture on seed ageing in Iceland poppy (*Papaver nudicaule* L.). *Scientia horticulturae, 48*(1-2), 153-158.

Belmonte, M. F., Kirkbride, R. C., Stone, S. L., Pelletier, J. M., Bui, A. Q., Yeung, E. C., . . . Munoz, M. D. (2013). Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. *Proceedings of the National Academy of Sciences, 110*(5), E435-E444.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological), 57*(1), 289-300.

Bentsink, L., Alonso-Blanco, C., Vreugdenhil, D., Tesnier, K., Groot, S. P., & Koornneef, M. (2000). Genetic analysis of seed-soluble oligosaccharides in relation to seed storability of Arabidopsis. *Plant physiology, 124*(4), 1595-1604.

Bentsink, L., Hanson, J., Hanhart, C. J., Blankestijn-de Vries, H., Coltrane, C., Keizer, P., . . . Reymond, M. (2010). Natural variation for seed dormancy in Arabidopsis is regulated by additive genetic and molecular pathways. *Proceedings of the National Academy of Sciences, 107*(9), 4264-4269.

Bentsink, L., Jowett, J., Hanhart, C. J., & Koornneef, M. (2006). Cloning of *DOG1*, a quantitative trait locus controlling seed dormancy in Arabidopsis. *Proceedings of the National Academy of Sciences, 103*(45), 17042-17047.

Bevan, M., & Walsh, S. (2005). The Arabidopsis genome: a foundation for plant research. *Genome Research, 15*(12), 1632-1642.

Bewley, J. D. (1997). Seed germination and dormancy. *The plant cell, 9*(7), 1055.

Bewley, J. D., & Black, M. (1994). Seeds. In *Seeds* (pp. 1-33): Springer.

Bewley, J. D., & Black, M. (2013). *Seeds: physiology of development and germination*: Springer Science & Business Media.

Bewley, J. D., Bradford, K., & Hilhorst, H. (2012). *Seeds: physiology of development, germination and dormancy*: Springer Science & Business Media.

Black, M. (1989). Seed research—past, present and future. In *Recent advances in the development and germination of seeds* (pp. 1-6): Springer.

Boesch, P., Weber-Lotfi, F., Ibrahim, N., Tarasenko, V., Cosset, A., Paulus, F., . . . Dietrich, A. (2011). DNA repair in organelles: pathways, organization, regulation, relevance in disease and aging. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1813*(1), 186-200.

Borbolis, F., & Syntichaki, P. (2015). Cytoplasmic mRNA turnover and ageing. *Mechanisms of ageing and development, 152*, 32-42.

Boswell, V. R. (1961). What seeds are and do: an introduction. *Seed. The Yearbook of Agricultural USDA*, 1-2.

Bouchaut, B., & Asveld, L. (2020). Safe-by-Design: Stakeholders' Perceptions and Expectations of How to Deal with Uncertain Risks of Emerging Biotechnologies in the Netherlands. *Risk Analysis*.

Bouteau, H. E.-M., Job, C., Job, D., Corbineau, F., & Bailly, C. (2007). ROS signaling in seed dormancy alleviation. *Plant Signaling & Behavior, 2*(5), 362-364.

Brady, S. M., Sarkar, S. F., Bonetta, D., & McCourt, P. (2003). The *ABSCISIC ACID INSENSITIVE 3* (*ABI3*) gene is modulated by farnesylation and is involved in auxin signaling and lateral root development in Arabidopsis. *The Plant Journal, 34*(1), 67-75.

Brocard-Gifford, I., Lynch, T. J., Garcia, M. E., Malhotra, B., & Finkelstein, R. R. (2004). The *Arabidopsis thaliana ABSCISIC ACID-INSENSITIVE8* locus encodes a novel protein mediating abscisic acid and sugar responses essential for growth. *The Plant Cell, 16*(2), 406-421.

Bryant, F. M., Hughes, D., Hassani-Pak, K., & Eastmond, P. J. (2019). Basic *LEUCINE ZIPPER TRANSCRIPTION FACTOR 67* transactivates *DELAY OF GERMINATION 1* to establish primary seed dormancy in Arabidopsis. *Plant Cell*. doi:10.1105/tpc.18.00892.

Buijs, G. (2020). A Perspective on Secondary Seed Dormancy in *Arabidopsis thaliana*. *Plants, 9*(6), 749.

Buijs, G., Kodde, J., Groot, S. P., & Bentsink, L. (2018). Seed dormancy release accelerated by elevated partial pressure of oxygen is associated with *DOG* loci. *Journal of experimental botany, 69*(15), 3601-3608.

Buijs, G., Vogelzang, A., Nijveen, H., & Bentsink, L. (2020). Dormancy cycling: translation-related transcripts are the main difference between dormant and non-dormant seeds in the field. *The Plant Journal, 102*(2), 327-339.

Buijs, G., Willems, L. A., Kodde, J., Groot, S. P., & Bentsink, L. (2020). Evaluating the EPPO method for seed longevity analyses in Arabidopsis. *Plant Science, 301*, 110644.

Buitink, J., & Leprince, O. (2008). Intracellular glasses and seed survival in the dry state. *Comptes rendus biologies, 331*(10), 788-795.

Burke, M. J. (1986). The glass state and survival of anhydrous biological systems. *Membranes, metabolism, and dry organisms*, 358-363.

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics, 25*(15), 1972-1973.

Carrera, E., Holman, T., Medhurst, A., Dietrich, D., Footitt, S., Theodoulou, F. L., & Holdsworth, M. J. (2008). Seed after-ripening is a discrete developmental pathway associated with specific gene networks in Arabidopsis. *The Plant Journal, 53*(2), 214-224.

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., Wang, X., . . . Samans, B. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *science, 345*(6199), 950-953.

Chang, C., Bowman, J. L., & Meyerowitz, E. M. (2016). Field guide to plant model systems. *Cell, 167*(2), 325-339.

Chen, H., Chu, P., Zhou, Y., Li, Y., Liu, J., Ding, Y., . . . Huang, S. (2012). Overexpression of AtOGG1, a DNA glycosylase/AP lyase, enhances seed longevity and abiotic stress tolerance in Arabidopsis. *Journal of experimental botany, 63*(11), 4107-4121.

Chen, H., Osuna, D., Colville, L., Lorenzo, O., Graeber, K., Kuester, H., . . . Kranner, I. (2013). Transcriptome-wide mapping of pea seed ageing reveals a pivotal role for genes related to oxidative stress and programmed cell death. *PLoS One, 8*(10), e78471.

Cheng, C. Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., & Town, C. D. (2017). Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *The Plant Journal, 89*(4), 789-804.

Cheung, W., Hubert, N., & Landry, B. (1993). A simple and rapid DNA microextraction method for plant, animal, and insect suitable for RAPD and other PCR analyses. *Genome Research, 3*(1), 69-70.

Chiang, G. C., Bartsch, M., Barua, D., Nakabayashi, K., Debieu, M., Kronholm, I., . . . de Meaux, J. (2011). *DOG1* expression is predicted by the seed-maturation environment and contributes to geographical variation in germination in *Arabidopsis thaliana*. *Molecular ecology, 20*(16), 3336-3349.

Ciolkowski, I., Wanke, D., Birkenbihl, R. P., & Somssich, I. E. (2008). Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant molecular biology, 68*(1-2), 81-92.

Clerkx, E. J., Blankestijn-De Vries, H., Ruys, G. J., Groot, S. P., & Koornneef, M. (2004). Genetic differences in seed longevity of various Arabidopsis mutants. *Physiologia Plantarum, 121*(3), 448-461.

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J, 16*(6), 735-743. Retrieved from PM:10069079.

Colville, L., & Pritchard, H. W. (2019). Seed life span and food security. *New Phytologist, 224*(2), 557-562.

Connor, K. F., & Bonner, F. T. (2001). The effects of desiccation on seeds of *Acer saccharinum* and *Aesculus pavia*: recalcitrance in temperate tree seeds. *Trees, 15*(3), 131-136.

Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics, 19*(1), 15-18.

Copeland, L. O., & McDonald, M. B. (1999). Seed longevity and deterioration. In *Principles of seed science and technology* (pp. 181-220): Springer.

Corbineau, F., Rudnicki, R., & Côme, D. (1988). Induction of secondary dormancy in sunflower seeds by high temperature. Possible involvement of ethylene biosynthesis. *Physiologia Plantarum, 73*(3), 368-373.

Dalle-Donne, I., Aldini, G., Carini, M., Colombo, R., Rossi, R., & Milzani, A. (2006). Protein carbonylation, cellular dysfunction, and disease progression. *Journal of cellular and molecular medicine, 10*(2), 389-406.

Davies, M. J. (2005). The oxidative environment and protein damage. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 1703*(2), 93-109.

de Winde, J., Thevelein, J., & Winderickx, J. (1997). Yeast stress responses.

Debeaujon, I., Léon-Kloosterziel, K. M., & Koornneef, M. (2000). Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. *Plant physiology, 122*(2), 403-414.

Debieu, M., Tang, C., Stich, B., Sikosek, T., Effgen, S., Josephs, E., . . . de Meaux, J. (2013). Co-variation between seed dormancy, growth rate and flowering time changes with latitude in *Arabidopsis thaliana. PLoS One, 8*(5), e61075. doi:10.1371/journal.pone.0061075

Dekkers, B. J., Pearce, S., van Bolderen-Veldkamp, R., Marshall, A., Widera, P., Gilbert, J., . . . King, J. R. (2013). Transcriptional dynamics of two seed compartments with opposing roles in Arabidopsis seed germination. *Plant physiology, 163*(1), 205-215.

Dekkers, B. J., Pearce, S. P., van Bolderen-Veldkamp, R., Holdsworth, M. J., & Bentsink, L. (2016). Dormant and after-ripened *Arabidopsis thaliana* seeds are distinguished by early transcriptional differences in the imbibed state. *Frontiers in Plant Science, 7*, 1323.

Dekkers, B. J., Willems, L., Bassel, G. W., van Bolderen-Veldkamp, R., Ligterink, W., Hilhorst, H. W., & Bentsink, L. (2012). Identification of reference genes for RT–qPCR expression analysis in Arabidopsis and tomato seeds. *Plant and cell physiology, 53*(1), 28-37.

Demir, I., & Mavi, K. (2008). Controlled deterioration and accelerated aging tests to estimate the relative storage potential of cucurbit seed lots. *HortScience, 43*(5), 1544-1548.

Dickie, J., Ellis, R., Kraak, H., Ryder, K., & Tompsett, P. (1990). Temperature and seed storage longevity. *Annals of botany, 65*(2), 197-204.

Ding, Z. J., Yan, J. Y., Li, G. X., Wu, Z. C., Zhang, S. Q., & Zheng, S. J. (2014). *WRKY 41* controls Arabidopsis seed dormancy via direct regulation of *ABI3* transcript levels not downstream of ABA. *The Plant Journal, 79*(5), 810-823.

Dure, L., & Waters, L. (1965). Long-lived messenger RNA: evidence from cotton seed germination. *science, 147*(3656), 410-412.

Dyson, T. (1996). *Population and food: global trends and future prospects*: Psychology Press.

el-Lithy, M. E., Bentsink, L., Hanhart, C. J., Ruys, G. J., Rovito, D., Broekhof, J. L., . . . Koornneef, M. (2006). New Arabidopsis recombinant inbred line populations genotyped using SNPWave and their use for mapping flowering-time quantitative trait loci. *Genetics, 172*(3), 1867-1876. doi:10.1534/genetics.105.050617.

Ellis, R., & Hong, T. (2006). Temperature sensitivity of the low-moisture-content limit to negative seed longevity–moisture content relationships in hermetic storage. *Annals of botany, 97*(5), 785-791.

Ellis, R., Hong, T., & Roberts, E. (1989). A comparison of the low-moisture-content limit to the logarithmic relation between seed moisture and longevity in twelve species. *Annals of botany, 63*(6), 601-611.

Ellis, R., Hong, T., & Roberts, E. (1990). An intermediate category of seed storage behaviour? I. Coffee. *Journal of experimental botany, 41*(9), 1167-1174.

Ellis, R., Hong, T., Roberts, E., & Tao, K.-L. (1990). Low moisture content limits to relations between seed longevity and moisture. *Annals of botany, 65*(5), 493-504.

Erlanger, S. (2005). After 2,000 Years, a Seed From Ancient Judea Sprouts. *The New York Times*.

Ernst, H. A., Nina Olsen, A., Skriver, K., Larsen, S., & Lo Leggio, L. (2004). Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. *EMBO reports, 5*(3), 297-303.

Estruch, F. (2000). Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast. *FEMS microbiology reviews, 24*(4), 469-486.

Fait, A., Angelovici, R., Less, H., Ohad, I., Urbanczyk-Wochniak, E., Fernie, A. R., & Galili, G. (2006). Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. *Plant physiology, 142*(3), 839-854.

Fang, Y., You, J., Xie, K., Xie, W., & Xiong, L. (2008). Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice. *Molecular Genetics and Genomics, 280*(6), 547-563.

Fennimore, S., Nyquist, W., Shaner, G., Doerge, R., & Foley, M. (1999). A genetic model and molecular markers for wild oat (*Avena fatua* L.) seed dormancy. *Theoretical and Applied Genetics, 99*(3-4), 711-718.

Fernández-Marín, B., Kranner, I., Sebastián, M. S., Artetxe, U., Laza, J. M., Vilas, J. L., . . . Becerril, J. M. (2013). Evidence for the absence of enzymatic reactions in the glassy state. A case study of xanthophyll cycle pigments in the desiccation-tolerant moss *Syntrichia ruralis*. *Journal of experimental botany, 64*(10), 3033-3043.

Finkel, T. (1998). Oxygen radicals and signaling. *Current opinion in cell biology, 10*(2), 248-253.

Finkelstein, R., Reeves, W., Ariizumi, T., & Steber, C. (2008). Molecular aspects of seed dormancy. *Annu. Rev. Plant Biol., 59*, 387-415.

Flecha, F. L. G. (2017). Kinetic stability of membrane proteins. *Biophysical Reviews, 9*(5), 563-572.

Fleming, M. B., Patterson, E. L., Reeves, P. A., Richards, C. M., Gaines, T. A., & Walters, C. (2018). Exploring the fate of mRNA in aging seeds: protection, destruction, or slow decay? *Journal of experimental botany, 69*(18), 4309-4321.

Fleming, M. B., Richards, C. M., & Walters, C. (2017). Decline in RNA integrity of dry-stored soybean seeds correlates with loss of germination potential. *Journal of experimental botany, 68*(9), 2219-2230.

Footitt, S., Huang, Z., Clay, H. A., Mead, A., & Finch-Savage, W. E. (2013). Temperature, light and nitrate sensing coordinate Arabidopsis seed dormancy cycling, resulting in winter and summer annual phenotypes. *The Plant Journal, 74*(6), 1003-1015.

Footitt, S., Müller, K., Kermode, A. R., & Finch-Savage, W. E. (2015). Seed dormancy cycling in Arabidopsis: chromatin remodelling and regulation of *DOG1* in response to seasonal environmental signals. *The Plant Journal, 81*(3), 413-425.

Footitt, S., Ölçer-Footitt, H., Hambidge, A. J., & Finch-Savage, W. E. (2017). A laboratory simulation of Arabidopsis seed dormancy cycling provides new insight into its regulation by clock genes and the dormancy-related genes *DOG1*, *MFT*, *CIPK23* and *PHYA*. *Plant, cell & environment, 40*(8), 1474-1486.

Fujita, M., Fujita, Y., Maruyama, K., Seki, M., Hiratsu, K., Ohme-Takagi, M., . . . Shinozaki, K. (2004). A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. *The Plant Journal, 39*(6), 863-876.

Galland, M., Huguet, R., Arc, E., Cueff, G., Job, D., & Rajjou, L. (2014). Dynamic proteomics emphasizes the importance of selective mRNA translation and protein turnover during Arabidopsis seed germination. *Molecular & Cellular Proteomics, 13*(1), 252-268.

Gandhi, S. D., Heesacker, A. F., Freeman, C. A., Argyris, J., Bradford, K., & Knapp, S. J. (2005). The self-incompatibility locus (S) and quantitative trait loci for self-pollination and seed dormancy in sunflower. *Theoretical and Applied Genetics, 111*(4), 619-629.

Garretón, V., Carpinelli, J., Jordana, X., & Holuigue, L. (2002). The *as-1* promoter element is an oxidative stress-responsive element and salicylic acid activates it via oxidative species. *Plant physiology, 130*(3), 1516-1526.

Gerland, P., Raftery, A. E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., . . . Lalic, N. (2014). World population stabilization unlikely this century. *science, 346*(6206), 234-237.

German, M. A., Luo, S., Schroth, G., Meyers, B. C., & Green, P. J. (2009). Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nature protocols, 4*(3), 356-362.

German, M. A., Pillay, M., Jeong, D.-H., Hetawal, A., Luo, S., Janardhanan, P., . . . German, R. (2008). Global identification of microRNA–target RNA pairs by parallel analysis of RNA ends. *Nature biotechnology, 26*(8), 941-946.

Gilland, B. (2002). World population and food supply: can food production keep pace with population growth in the next half-century? *Food policy, 27*(1), 47-63.

Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., . . . Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *science, 327*(5967), 812-818.

Goldberg, R. B., Barker, S. J., & Perez-Grau, L. (1989). Regulation of gene expression during plant embryogenesis. *Cell, 56*(2), 149-160.

Goldberg, R. B., De Paiva, G., & Yadegari, R. (1994). Plant embryogenesis: zygote to seed. *science, 266*(5185), 605-614.

Grandori, C., Cowley, S. M., James, L. P., & Eisenman, R. N. (2000). The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annual review of cell and developmental biology, 16*(1), 653-699.

Groot, S., & Karssen, C. (1987). Gibberellins regulate seed germination in tomato by endosperm weakening: a study with gibberellin-deficient mutants. *Planta, 171*(4), 525-531.

Groot, S., Surki, A., De Vos, R., & Kodde, J. (2012). Seed storage at elevated partial pressure of oxygen, a fast method for analysing seed ageing under dry conditions. *Annals of botany, 110*(6), 1149-1159.

Gualberti, G., Papi, M., Bellucci, L., Ricci, I., Bouchez, D., Camilleri, C., . . . Vittorioso, P. (2002). Mutations in the Dof zinc finger genes *DAG2* and *DAG1* influence with opposite effects the germination of Arabidopsis seeds. *The Plant Cell, 14*(6), 1253-1263.

Guiltinan, M. J., Marcotte, W. R., & Quatrano, R. S. (1990). A plant leucine zipper protein that recognizes an abscisic acid response element. *science, 250*(4978), 267-271.

Guo, L., Zhu, L., Xu, Y., Zeng, D., Wu, P., & Qian, Q. (2004). QTL analysis of seed dormancy in rice (*Oryza sativa* L.). *Euphytica, 140*(3), 155-162.

Guo, L., Chen, S., Liu, K., Liu, Y., Ni, L., Zhang, K., & Zhang, L. (2008). Isolation of heat shock factor HsfA1a-binding sites in vivo revealed variations of heat shock elements in *Arabidopsis thaliana*. *Plant and Cell Physiology, 49*(9), 1306-1315.

Gutierrez, G., Cruz, F., Moreno, J., Gonzalez-Hernandez, V. A., & Vazquez-Ramos, J. M. (1993). Natural and artificial seed ageing in maize: germination and DNA synthesis. *Seed Science Research, 3*, 279-279.

Halees, A. S., Leyfer, D., & Weng, Z. (2003). PromoSer: a large-scale mammalian promoter and transcription start site identification service. *Nucleic acids research, 31*(13), 3554-3559.

Hay, F. R., & Whitehouse, K. J. (2017). Rethinking the approach to viability monitoring in seed genebanks. *Conservation physiology, 5*(1).

He, H., de Souza Vidigal, D., Snoek, L. B., Schnabel, S., Nijveen, H., Hilhorst, H., & Bentsink, L. (2014). Interaction between parental environment and genotype affects plant and seed performance in Arabidopsis. *J Exp Bot, 65*(22), 6603-6615. doi:10.1093/jxb/eru378.

Hegedus, D., Yu, M., Baldwin, D., Gruber, M., Sharpe, A., Parkin, I., . . . Lydiate, D. (2003). Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress. *Plant molecular biology, 53*(3), 383-397.

Hendry, G. A. (1993). Oxygen, free radical processes and seed longevity. *Seed Science Research, 3*(3), 141-153.

Hilhorst, H. W. (1995). A critical update on seed dormancy. I. Primary dormancy. *Seed Science Research, 5*(2), 61-73.

Holdsworth, M. J., Bentsink, L., & Soppe, W. J. (2008). Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytologist, 179*(1), 33-54.

Hong, J., Lee, H., Lee, J., Kim, H., & Ryu, H. (2019). ABSCISIC ACID-INSENSITIVE 3 is involved in brassinosteroid-mediated regulation of flowering in plants. *Plant Physiology and Biochemistry, 139*, 207-214.

Hong, T., & Ellis, R. H. (1996). *A protocol to determine seed storage behaviour*: Bioversity International.

Hoppe, T., Rape, M., & Jentsch, S. (2001). Membrane-bound transcription factors: regulated release by RIP or RUP. *Current opinion in cell biology, 13*(3), 344-348.

Houseley, J., & Tollervey, D. (2009). The many pathways of RNA degradation. *Cell, 136*(4), 763-776.

Hu, H., Dai, M., Yao, J., Xiao, B., Li, X., Zhang, Q., & Xiong, L. (2006). Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proceedings of the National Academy of Sciences, 103*(35), 12987-12992.

Huang, Y., Feng, C.-Z., Ye, Q., Wu, W.-H., & Chen, Y.-F. (2016). *Arabidopsis WRKY6* transcription factor acts as a positive regulator of abscisic acid signaling during seed germination and early seedling development. *PLoS Genet, 12*(2), e1005833.

Jiang, W., & Yu, D. (2009). *Arabidopsis WRKY2* transcription factor mediates seed germination and postgermination arrest of development by abscisic acid. *BMC plant biology, 9*(1), 96.

Job, C., & Bailly, C. (2013). Role of protein and mRNA oxidation in seed dormancy and germination. *Frontiers in Plant Science, 4*, 77.

Job, C., Rajjou, L., Lovigny, Y., Belghazi, M., & Job, D. (2005). Patterns of protein oxidation in Arabidopsis seeds and during germination. *Plant physiology, 138*(2), 790-802.

Johnston, J. W., Pimbley, I., Harding, K., & Benson, E. E. (2010). Detection of 8-hydroxy-2'-deoxyguanosine as a marker of oxidative damage in DNA and germplasm exposed to cryogenic treatments. *CryoLetters, 31*(1), 1-13.

Joosen, R. V., Kodde, J., Willems, L. A., Ligterink, W., van der Plas, L. H., & Hilhorst, H. W. (2010). GERMINATOR: a software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. *The Plant Journal, 62*(1), 148-159.

Justice, O. L., & Bass, L. N. (1978). *Principles and practices of seed storage*: Department of Agriculture, Science and Education Administration.

Karin, M., Liu, Z.-g., & Zandi, E. (1997). AP-1 function and regulation. *Current opinion in cell biology, 9*(2), 240-246.

Karlova, R., van Haarst, J. C., Maliepaard, C., van de Geest, H., Bovy, A. G., Lammers, M., . . . de Maagd, R. A. (2013). Identification of microRNA targets in tomato fruit development using high-throughput sequencing and degradome analysis. *Journal of experimental botany, 64*(7), 1863-1878.

Kato, K., Nakamura, W., Tabiki, T., Miura, H., & Sawada, S. (2001). Detection of loci controlling seed dormancy on group 4 chromosomes of wheat and comparative mapping with rice and barley genomes. *Theoretical and Applied Genetics, 102*(6-7), 980-985.

Kearsey, M., & Farquhar, A. (1998). QTL analysis in plants; where are we now? *Heredity, 80*(2), 137-142.

Kendall, S. L., Hellwege, A., Marriot, P., Whalley, C., Graham, I. A., & Penfield, S. (2011). Induction of Dormancy in Arabidopsis Summer Annuals Requires Parallel Regulation of *DOG1* and Hormone Metabolism by Low Temperature and CBF Transcription Factors. *The Plant Cell*. doi:10.1105/tpc.111.087643.

Kermode, A. R., & Finch-Savage, B. E. (2002). Desiccation sensitivity in orthodox and recalcitrant seeds in relation to development. *Desiccation and survival in plants: Drying without dying*, 149-184.

Khan, A. A. (1980). Hormonal regulation of primary and secondary seed dormancy. *Israel Journal of Plant Sciences, 29*(1-4), 207-224.

Khandelwal, A., Cho, S. H., Marella, H., Sakata, Y., Perroud, P.-F., Pan, A., & Quatrano, R. (2010). Role of ABA and *ABI3* in desiccation tolerance. *science, 327*(5965), 546-546.

Kibinza, S., Vinel, D., Côme, D., Bailly, C., & Corbineau, F. (2006). Sunflower seed deterioration as related to moisture content during ageing, energy metabolism and active oxygen species scavenging. *Physiologia Plantarum, 128*(3), 496-506.

Kim, S.-G., Kim, S.-Y., & Park, C.-M. (2007). A membrane-associated NAC transcription factor regulates salt-responsive flowering via *FLOWERING LOCUS T* in Arabidopsis. *Planta, 226*(3), 647-654.

Kim, S.-G., Lee, S., Ryu, J., & Park, C.-M. (2010). Probing protein structural requirements for activation of membrane-bound NAC transcription factors in Arabidopsis and rice. *Plant Science, 178*(3), 239-244.

Kim, S.-G., Lee, S., Seo, P. J., Kim, S.-K., Kim, J.-K., & Park, C.-M. (2010). Genome-scale screening and molecular characterization of membrane-bound transcription factors in Arabidopsis and rice. *Genomics, 95*(1), 56-65.

Kim, S.-Y., Kim, S.-G., Kim, Y.-S., Seo, P. J., Bae, M., Yoon, H.-K., & Park, C.-M. (2007). Exploring membrane-associated NAC transcription factors in Arabidopsis: implications for membrane biology in genome regulation. *Nucleic acids research, 35*(1), 203-213.

Kim, S. G., Lee, A. K., Yoon, H. K., & Park, C. M. (2008). A membrane-bound NAC transcription factor *NTL8* regulates gibberellic acid-mediated salt signaling in Arabidopsis seed germination. *The Plant Journal, 55*(1), 77-88.

Kim, Y.-S., Kim, S.-G., Park, J.-E., Park, H.-Y., Lim, M.-H., Chua, N.-H., & Park, C.-M. (2006). A membrane-bound NAC transcription factor regulates cell division in Arabidopsis. *The Plant Cell, 18*(11), 3132-3144.

Kimura, M., & Nambara, E. (2010). Stored and neosynthesized mRNA in Arabidopsis seeds: effects of cycloheximide and controlled deterioration treatment on the resumption of transcription during imbibition. *Plant molecular biology, 73*(1-2), 119-129.

Klein, P., Seidel, T., Stöcker, B., & Dietz, K.-J. (2012). The membrane-tethered transcription factor *ANAC089* serves as redox-dependent suppressor of stromal ascorbate peroxidase gene expression. *Frontiers in Plant Science, 3*, 247.

Knott, G. J., & Doudna, J. A. (2018). CRISPR-Cas guides the future of genetic engineering. *science, 361*(6405), 866-869.

Koornneef, M., Bentsink, L., & Hilhorst, H. (2002). Seed dormancy and germination. *Current opinion in plant biology, 5*(1), 33-36.

Koornneef, M., & Meinke, D. (2010). The development of Arabidopsis as a model plant. *The Plant Journal, 61*(6), 909-921.

Kopylova, E., Noé, L., & Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics, 28*(24), 3211-3217.

Koyuncu, F. (2005). Breaking seed dormancy in black mulberry (*Morus nigra* L.) by cold stratification and exogenous application of gibberellic acid. *Acta Biologica Cracoviensia Series Botanica, 47*(2), 23-26.

Kranner, I., Chen, H., Pritchard, H. W., Pearce, S. R., & Birtić, S. (2011). Inter-nucleosomal DNA fragmentation and loss of RNA integrity during seed ageing. *Plant Growth Regulation, 63*(1), 63-72.

Kroj, T., Savino, G., Valon, C., Giraudat, J., & Parcy, F. (2003). Regulation of storage protein gene expression in Arabidopsis. *Development, 130*(24), 6065-6073. doi:Doi 10.1242/Dev.00814.

Krukowski, P. K., Ellenberger, J., Röhlen-Schmittgen, S., Schubert, A., & Cardinale, F. (2020). Phenotyping in *Arabidopsis* and Crops—Are We Addressing the Same Traits? A Case Study in Tomato. *Genes, 11*(9), 1011.

Kucera, B., Cohn, M. A., & Leubner-Metzger, G. (2005). Plant hormone interactions during seed dormancy release and germination. *Seed Science Research, 15*(4), 281-307.

Kuhn, J. M., Boisson-Dernier, A., Dizon, M. B., Maktabi, M. H., & Schroeder, J. I. (2006). The protein phosphatase *AtPP2CA* negatively regulates abscisic acid signal transduction in Arabidopsis, and effects of *abh1* on *AtPP2CA* mRNA. *Plant Physiol, 140*(1), 127-139. doi:10.1104/pp.105.070318.

Kumar, D., & Mishra, D. K. (2014). Variability in permeability and integrity of cell membrane and depletion of food reserves in neem (*Azadirachta indica*) seeds from trees of different age classes. *Journal of forestry research, 25*(1), 147-153.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular biology and evolution, 33*(7), 1870-1874.

Kurek, K., Plitta-Michalak, B., & Ratajczak, E. (2019). Reactive oxygen species as potential drivers of the seed aging process. *Plants, 8*(6), 174.

Laibach, F. (1907). *Zur Frage nach der Individualitat der Chromosomen im Pflanzenreich.* (Doctoral dissertation, Bonn.).

Lam, E., & Chua, N.-H. (1990). GT-1 binding site confers light responsive expression in transgenic tobacco. *science, 248*(4954), 471-474.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods, 9*(4), 357.

Lawrence, M. (1976). Variations in natural populations of *Arabidopsis thaliana* (L.) Heynh. *the Biology and Chemistry of the Cruciferae. JG Vaughan, AJ Macleod & BMG Jones, eds*.

Lawrenson, T., Shorinola, O., Stacey, N., Li, C., Østergaard, L., Patron, N., . . . Harwood, W. (2015). Induction of targeted, heritable mutations in barley and *Brassica oleracea* using RNA-guided Cas9 nuclease. *Genome biology, 16*(1), 258.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., . . . Rombauts, S. (2002). PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic acids research, 30*(1), 325-327.

Leubner-Metzger, G. (2001). Brassinosteroids and gibberellins promote tobacco seed germination by distinct pathways. *Planta, 213*(5), 758-763.

Levine, R. L., Garland, D., Oliver, C. N., Amici, A., Climent, I., Lenz, A.-G., . . . Stadtman, E. R. (1990). [49] Determination of carbonyl content in oxidatively modified proteins. *Methods in enzymology, 186*, 464-478.

Levine, R. L., Williams, J. A., Stadtman, E. P., & Shacter, E. (1994). [37] Carbonyl assays for determination of oxidatively modified proteins. *Methods in enzymology, 233*, 346-357.

Leymarie, J., Vitkauskaité, G., Hoang, H. H., Gendreau, E., Chazoule, V., Meimoun, P., . . . Bailly, C. (2012). Role of reactive oxygen species in the regulation of Arabidopsis seed dormancy. *Plant and cell physiology, 53*(1), 96-106.

Li, B., & Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics, 12*(1), 323.

Li, C., Tarr, A., Lance, R., Harasymow, S., Uhlmann, J., Westcot, S., . . . Broughton, S. (2003). A major QTL controlling seed dormancy and pre-harvest sprouting/grain α-amylase in two-rowed barley (*Hordeum vulgare* L.). *Australian Journal of Agricultural Research, 54*(12), 1303-1313.

Li, D.-Z., & Pritchard, H. W. (2009). The science and economics of *ex situ* plant conservation. *Trends in plant science, 14*(11), 614-621.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics, 25*(16), 2078-2079.

Li, J., Zhang, J., Wang, X., & Chen, J. (2010). A membrane-tethered transcription factor *ANAC089* negatively regulates floral initiation in *Arabidopsis thaliana*. *Science China Life Sciences, 53*(11), 1299-1306.

Li, P., Wind, J. J., Shi, X., Zhang, H., Hanson, J., Smeekens, S. C., & Teng, S. (2011). Fructose sensitivity is suppressed in Arabidopsis by the transcription factor *ANAC089* lacking the membrane-bound domain. *Proceedings of the National Academy of Sciences, 108*(8), 3436-3441.

Li, P., Zhou, H., Shi, X., Yu, B., Zhou, Y., Chen, S., . . . Smeekens, S. C. (2014). The *ABI4*-induced Arabidopsis *ANAC060* transcription factor attenuates ABA signaling and renders seedlings sugar insensitive when present in the nucleus. *PLoS Genet, 10*(3), e1004213.

Li, X., Chen, T., Li, Y., Wang, Z., Cao, H., Chen, F., . . . Liu, Y. (2019). ETR1/RDO3 Regulates Seed Dormancy by Relieving the Inhibitory Effect of the ERF12-TPL Complex on *DELAY OF GERMINATION1* Expression. *Plant Cell, 31*(4), 832-847. doi:10.1105/tpc.18.00449.

Li, Z., Wu, J., & DeLeo, C. J. (2006). RNA damage and surveillance under oxidative stress. *IUBMB life, 58*(10), 581-588.

Linkies, A., Müller, K., Morris, K., Turečková, V., Wenk, M., Cadman, C. S., . . . Finch-Savage, W. E. (2009). Ethylene interacts with abscisic acid to regulate endosperm rupture during germination: a comparative approach using *Lepidium sativum* and *Arabidopsis thaliana*. *The Plant Cell, 21*(12), 3803-3822.

Liu, X., Zhang, H., Zhao, Y., Feng, Z., Li, Q., Yang, H.-Q., . . . He, Z.-H. (2013). Auxin controls seed dormancy through stimulation of abscisic acid signaling by inducing ARF-mediated *ABI3* activation in Arabidopsis. *Proceedings of the National Academy of Sciences, 110*(38), 15485-15490.

Liu, Y., Ye, N., Liu, R., Chen, M., & Zhang, J. (2010). H2O2 mediates the regulation of ABA catabolism and GA biosynthesis in Arabidopsis seed dormancy and germination. *Journal of experimental botany, 61*(11), 2979-2990.

Long, R. L., Gorecki, M. J., Renton, M., Scott, J. K., Colville, L., Goggin, D. E., . . . Finch-Savage, W. E. (2015). The ecophysiology of seed persistence: a mechanistic view of the journey to germination or demise. *Biological Reviews, 90*(1), 31-59.

Lopez-Molina, L., Mongrand, S., Kinoshita, N., & Chua, N. H. (2003). AFP is a novel negative regulator of ABA signaling that promotes ABI5 protein degradation. *Genes Dev, 17*(3), 410-418.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology, 15*(12), 550.

Mabuchi, A., Soga, K., Wakabayashi, K., & Hoson, T. (2016). Phenotypic screening of Arabidopsis T-DNA insertion lines for cell wall mechanical properties revealed *ANTHOCYANINLESS2*, a cell wall-related gene. *Journal of plant physiology, 191*, 29-35.

Maia, J., Dekkers, B. J., Provart, N. J., Ligterink, W., & Hilhorst, H. W. (2011). The re-establishment of desiccation tolerance in germinated *Arabidopsis thaliana* seeds and its associated transcriptome. *PLoS One, 6*(12), e29123. doi:10.1371/journal.pone.0029123.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., . . . Yakunin, A. F. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology, 9*(6), 467-477.

Mansfield, S., & Briarty, L. (1991). Early embryogenesis in *Arabidopsis thaliana*. II. The developing embryo. *Canadian journal of botany, 69*(3), 461-476.

Mansfield, S., & Briarty, L. (1992). Cotyledon cell development in *Arabidopsis thaliana* during reserve deposition. *Canadian journal of botany, 70*(1), 151-164.

Mansfield, S., & Briarty, L. (1994). Arabidopsis: An atlas of morphology and development. In: New York: Springer.

Marcus, A., & Feeley, J. (1964). Activation of protein synthesis in the imbibition phase of seed germination. *Proceedings of the National Academy of Sciences of the United States of America, 51*(6), 1075.

Martínez-Andújar, C., Martin, R. C., & Nonogaki, H. (2012). Seed traits and genes important for translational biology—highlights from recent discoveries. *Plant and cell physiology, 53*(1), 5-15.

Maruyama, K., Todaka, D., Mizoi, J., Yoshida, T., Kidokoro, S., Matsukura, S., . . . Yoshiwara, K. (2012). Identification of *cis*-acting promoter elements in cold-and dehydration-induced transcriptional pathways in Arabidopsis, rice, and soybean. *DNA research, 19*(1), 37-49.

Mayer, U., & Jürgens, G. (1998). *Pattern formation in plant embryogenesis: a reassessment.* Paper presented at the Seminars in cell & developmental biology.

Mayer, U., Ruiz, R. A. T., Berleth, T., Miséra, S., & Jürgens, G. (1991). Mutations affecting body organization in the Arabidopsis embryo. *Nature, 353*(6343), 402-407.

McCarthy, D. J., Chen, Y., & Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic acids research, 40*(10), 4288-4297.

McCouch, S. R. (2001). Genomics and synteny. *Plant physiology, 125*(1), 152-155.

Meinke, D. W. (1992). A homoeotic mutant of *Arabidopsis thaliana* with leafy cotyledons. *science, 258*(5088), 1647-1650.

Meinke, D. W., Franzmann, L. H., Nickle, T. C., & Yeung, E. C. (1994). Leafy cotyledon mutants of Arabidopsis. *The Plant Cell, 6*(8), 1049-1064.

Mira, S., Estrelles, E., & González-Benito, M. (2015). Effect of water content and temperature on seed longevity of seven Brassicaceae species after 5 years of storage. *Plant Biology, 17*(1), 153-162.

Miransari, M., & Smith, D. (2014). Plant hormones and seed germination. *Environmental and Experimental Botany, 99*, 110-121.

Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A., & Punta, M. (2013). Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic acids research, 41*(12), e121-e121.

MORRIS, P. C., KUMAR, A., BOWLES, D. J., & CUMING, A. C. (1990). Osmotic stress and abscisic acid induce expression of the wheat Em genes. *European Journal of Biochemistry, 190*(3), 625-630.

Mu, J., Tan, H., Zheng, Q., Fu, F., Liang, Y., Zhang, J., . . . Wang, X.-J. (2008). *LEAFY COTYLEDON1* is a key regulator of fatty acid biosynthesis in Arabidopsis. *Plant physiology, 148*(2), 1042-1054.

Mueller, O., Lightfoot, S., & Schroeder, A. (2004). RNA integrity number (RIN)–standardization of RNA quality control. *Agilent application note, publication, 1*, 1-8.

Murphey, M., Kovach, K., Elnacash, T., He, H., Bentsink, L., & Donohue, K. (2015). *DOG1*-imposed dormancy mediates germination responses to temperature cues. *Environmental and Experimental Botany, 112*, 33-43.

Nagel, M., Kodde, J., Pistrick, S., Mascher, M., Börner, A., & Groot, S. P. (2016). Barley seed aging: genetics behind the dry elevated pressure of oxygen aging and moist controlled deterioration. *Frontiers in Plant Science, 7*, 388.

Nakabayashi, K., Bartsch, M., Xiang, Y., Miatton, E., Pellengahr, S., Yano, R., . . . Soppe, W. J. (2012). The time required for dormancy release in Arabidopsis is determined by DELAY OF GERMINATION1 protein levels in freshly harvested seeds. *The Plant Cell, 24*(7), 2826-2838.

Nakabayashi, K., Okamoto, M., Koshiba, T., Kamiya, Y., & Nambara, E. (2005). Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: epigenetic and genetic regulation of transcription in seed. *The Plant Journal, 41*(5), 697-709.

Nambara, E., Kawaide, H., Kamiya, Y., & Naito, S. (1998). Characterization of an *Arabidopsis thaliana* mutant that has a defect in ABA accumulation: ABA-dependent and ABA-independent accumulation of free amino acids during dehydration. *Plant and cell physiology, 39*(8), 853-858.

Nambara, E., McCourt, P., & Naito, S. (1995). A regulatory role for the *ABI3* gene in the establishment of embryo maturation in *Arabidopsis thaliana*. *Development, 121*(3), 629-636.

Nambara, E., Okamoto, M., Tatematsu, K., Yano, R., Seo, M., & Kamiya, Y. (2010). Abscisic acid and the control of seed dormancy and germination. *Seed Science Research, 20*(2), 55-67.

Nannas, N. J., & Dawe, R. K. (2015). Genetic and genomic toolbox of *Zea mays*. *Genetics, 199*(3), 655-669.

Narsai, R., Howell, K. A., Millar, A. H., O'Toole, N., Small, I., & Whelan, J. (2007). Genome-wide analysis of mRNA decay rates and their determinants in *Arabidopsis thaliana*. *The Plant Cell, 19*(11), 3418-3436.

Narusaka, Y., Nakashima, K., Shinwari, Z. K., Sakuma, Y., Furihata, T., Abe, H., . . . Yamaguchi-Shinozaki, K. (2003). Interaction between two *cis*-acting elements, ABRE and DRE, in ABA-

dependent expression of *Arabidopsis rd29A* gene in response to dehydration and high-salinity stresses. *The Plant Journal, 34*(2), 137-148.

Nee, G., Kramer, K., Nakabayashi, K., Yuan, B., Xiang, Y., Miatton, E., . . . Soppe, W. J. J. (2017). DELAY OF GERMINATION1 requires PP2C phosphatases of the ABA signalling pathway to control seed dormancy. *Nat Commun, 8*(1), 72. doi:10.1038/s41467-017-00113-6.

Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution, 32*(1), 268-274.

Nguyen, T.-P., Cueff, G., Hegedus, D. D., Rajjou, L., & Bentsink, L. (2015). A role for seed storage proteins in Arabidopsis seed longevity. *Journal of experimental botany, 66*(20), 6399-6413.

Nguyen, T.-P., Keizer, P., van Eeuwijk, F., Smeekens, S., & Bentsink, L. (2012). Natural variation for seed longevity and seed dormancy are negatively correlated in Arabidopsis. *Plant physiology, 160*(4), 2083-2092.

Nikolić, R., Mitić, N., Miletić, R., & Nešković, M. (2006). Effects of cytokinins on *in vitro* seed germination and early seedling morphogenesis in *Lotus corniculatus* L. *Journal of plant growth Regulation, 25*(3), 187.

Nonogaki, H., Bassel, G. W., & Bewley, J. D. (2010). Germination—still a mystery. *Plant Science, 179*(6), 574-581.

North, H., Baud, S., Debeaujon, I., Dubos, C., Dubreucq, B., Grappin, P., . . . Miquel, M. (2010). Arabidopsis seed secrets unravelled after a decade of genetic and omics-driven research. *The Plant Journal, 61*(6), 971-981.

Nyachiro, J., Clarke, F., DePauw, R., Knox, R., & Armstrong, K. (2002). Temperature effects on seed germination and expression of seed dormancy in wheat. *Euphytica, 126*(1), 123-127.

Ohlrogge, J. B., & Kernan, T. P. (1982). Oxygen-dependent aging of seeds. *Plant physiology, 70*(3), 791-794.

Olsen, A. N., Ernst, H. A., Leggio, L. L., & Skriver, K. (2005). NAC transcription factors: structurally distinct, functionally diverse. *Trends in plant science, 10*(2), 79-87.

Ooka, H., Satoh, K., Doi, K., Nagata, T., Otomo, Y., Murakami, K., . . . Carninci, P. (2003). Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. *DNA Research, 10*(6), 239-247.

Ooms, J. J., Leon-Kloosterziel, K. M., Bartels, D., Koornneef, M., & Karssen, C. M. (1993). Acquisition of desiccation tolerance and longevity in seeds of *Arabidopsis thaliana* (a comparative study using abscisic acid-insensitive *abi3* mutants). *Plant physiology, 102*(4), 1185-1191.

Oracz, K., Bouteau, H. E. M., Farrant, J. M., Cooper, K., Belghazi, M., Job, C., . . . Bailly, C. (2007). ROS production and protein oxidation as a novel mechanism for seed dormancy alleviation. *The Plant Journal, 50*(3), 452-465.

Oracz, K., El-Maarouf-Bouteau, H., Kranner, I., Bogatek, R., Corbineau, F., & Bailly, C. (2009). The mechanisms involved in seed dormancy alleviation by hydrogen cyanide unravel the role of reactive oxygen species as key factors of cellular signaling during germination. *Plant physiology, 150*(1), 494-505.

Osa, M., Kato, K., Mori, M., Shindo, C., Torada, A., & Miura, H. (2003). Mapping QTLs for seed dormancy and the *Vp1* homologue on chromosome 3A in wheat. *Theoretical and Applied Genetics, 106*(8), 1491-1496.

Ouwerkerk, P., & Memelink, J. (1999). A G-box element from the *Catharanthus roseus* strictosidine synthase (*Str*) gene promoter confers seed-specific expression in transgenic tobacco plants. *Molecular and General Genetics MGG, 261*(4-5), 635-643.

Pantaleo, V., Szittya, G., Moxon, S., Miozzi, L., Moulton, V., Dalmay, T., & Burgyan, J. (2010). Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. *The Plant Journal, 62*(6), 960-976.

Parcy, F., Valon, C., Kohara, A., Miséra, S., & Giraudat, J. (1997). The *ABSCISIC ACID-INSENSITIVE3*, *FUSCA3*, and *LEAFY COTYLEDON1* loci act in concert to control multiple aspects of Arabidopsis seed development. *The Plant Cell, 9*(8), 1265-1277.

Pati, A. (2005). *Modeling and Analysis of Regulatory Elements in Arabidopsis thaliana from Annotated Genomes and Gene Expression Data.* (Doctoral dissertation, Virginia Tech).

Pearcon, W., & Lipman, D. (1988). Improved tools for biological sequence analysis. *Proc Natl Acad Sci USA, 85*(2444), 8.

Pearson, W. R. (2000). Flexible sequence similarity searching with the FASTA3 program package. In *Bioinformatics methods and protocols* (pp. 185-219): Springer.

Pelletier, J. M., Kwong, R. W., Park, S., Le, B. H., Baden, R., Cagliari, A., . . . Goldberg, R. B. (2017). LEC1 sequentially regulates the transcription of genes involved in diverse developmental processes during seed development. *Proceedings of the National Academy of Sciences, 114*(32), E6710-E6719.

Piquerez, S. J., Harvey, S. E., Beynon, J. L., & Ntoukakis, V. (2014). Improving crop disease resistance: lessons from research on Arabidopsis and tomato. *Frontiers in Plant Science, 5*, 671.

Plitta-Michalak, B. P., Naskręt-Barciszewska, M. Z., Kotlarski, S., Tomaszewski, D., Tylkowski, T., Barciszewski, J., . . . Michalak, M. (2018). Changes in genomic 5-methylcytosine level mirror the response of orthodox (*Acer platanoides* L.) and recalcitrant (*Acer pseudoplatanus* L.) seeds to severe desiccation. *Tree physiology, 38*(4), 617-629.

Polyak, K., & Meyerson, M. (2003). Overview: gene structure. In *Holland-Frei Cancer Medicine. 6th edition*: BC Decker.

Powell, A. A., & Matthews, S. (1981). Evaluation of controlled deterioration, a new vigour test for small seeded vegetables. *Seed Science and Technology, 9*(2), 633-640.

Priestley, D. A. (1986). *Seed aging: implications for seed storage and persistence in the soil*: Comstock Associates.

Prusinski, J., & Khan, A. A. (1990). Relationship of ethylene production to stress alleviation in seeds of lettuce cultivars. *Journal of the American Society for Horticultural Science, 115*(2), 294-298.

Puntarulo, S., Sánchez, R. A., & Boveris, A. (1988). Hydrogen peroxide metabolism in soybean embryonic axes at the onset of germination. *Plant physiology, 86*(2), 626-630.

Rajjou, L., Belghazi, M., Huguet, R., Robin, C., Moreau, A., Job, C., & Job, D. (2006). Proteomic investigation of the effect of salicylic acid on Arabidopsis seed germination and establishment of early defense mechanisms. *Plant physiology, 141*(3), 910-923.

Rajjou, L., & Debeaujon, I. (2008). Seed longevity: survival and maintenance of high germination ability of dry seeds. *Comptes rendus biologies, 331*(10), 796-805.

Rajjou, L., Gallardo, K., Debeaujon, I., Vandekerckhove, J., Job, C., & Job, D. (2004). The effect of α-amanitin on the Arabidopsis seed proteome highlights the distinct roles of stored and neosynthesized mRNAs during germination. *Plant physiology, 134*(4), 1598-1613.

Rajjou, L., Lovigny, Y., Groot, S. P., Belghazi, M., Job, C., & Job, D. (2008). Proteome-wide characterization of seed aging in Arabidopsis: a comparison between artificial and natural aging protocols. *Plant physiology, 148*(1), 620-641.

Ramaih, S., Guedira, M., & Paulsen, G. M. (2003). Relationship of indoleacetic acid and tryptophan to dormancy and preharvest sprouting of wheat. *Functional Plant Biology, 30*(9), 939-945.

Ratcliffe, D. (1976). Germination characteristics and their inter-and intra-population variability in Arabidopsis. *Arabidopsis Information Service, 13*, 34-45.

Ren, T., Qu, F., & Morris, T. J. (2000). *HRT* gene function requires interaction between a NAC protein and viral capsid protein to confer resistance to turnip crinkle virus. *The Plant Cell, 12*(10), 1917-1925.

Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C.-Z., Keddie, J., . . . Samaha, R. (2000). Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *science, 290*(5499), 2105-2110.

Riefler, M., Novak, O., Strnad, M., & Schmülling, T. (2006). Arabidopsis cytokinin receptor mutants reveal functions in shoot growth, leaf senescence, seed size, germination, root development, and cytokinin metabolism. *The Plant Cell, 18*(1), 40-54.

Roach, T., Beckett, R. P., Minibayeva, F. V., Colville, L., Whitaker, C., Chen, H., . . . Kranner, I. (2010). Extracellular superoxide production, viability and redox poise in response to desiccation in recalcitrant *Castanea sativa* seeds. *Plant, cell & environment, 33*(1), 59-75.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature biotechnology, 29*(1), 24-26.

Rodo, A. B., & Marcos Filho, J. (2003). Accelerated aging and controlled deterioration for the determination of the physiological potential of onion seeds. *Scientia Agricola, 60*(3), 465-469.

Rodriguez-Furlán, C., Miranda, G., Reggiardo, M., Hicks, G. R., & Norambuena, L. (2016). High throughput selection of novel plant growth regulators: assessing the translatability of small bioactive molecules from Arabidopsis to crops. *Plant Science, 245*, 50-60.

Rosbakh, S., Hülsmann, L., Weinberger, I., Bleicher, M., & Poschlod, P. (2019). Bleaching and cold stratification can break dormancy and improve seed germination in Cyperaceae. *Aquatic Botany, 158*, 103128.

Rushton, D. L., Tripathi, P., Rabara, R. C., Lin, J., Ringler, P., Boken, A. K., . . . Emme, N. J. (2012). WRKY transcription factors: key components in abscisic acid signalling. *Plant biotechnology journal, 10*(1), 2-11.

Rushton, P. J., Somssich, I. E., Ringler, P., & Shen, Q. J. (2010). WRKY transcription factors. *Trends in plant science, 15*(5), 247-258.

Saatkamp, A., Cochrane, A., Commander, L., Guja, L. K., Jimenez-Alfaro, B., Larson, J., . . . Cross, A. T. (2019). A research agenda for seed-trait functional ecology. *New Phytologist, 221*(4), 1764-1775.

Sablowski, R. W., & Meyerowitz, E. M. (1998). A homolog of *NO APICAL MERISTEM* is an immediate target of the floral homeotic genes *APETALA3/PISTILLATA*. *Cell, 92*(1), 93-103.

Sajeev, N., Bai, B., & Bentsink, L. (2019). Seeds: A unique system to study translational regulation. *Trends in plant science, 24*(6), 487-495.

Sall, K., Dekkers, B. J. W., Nonogaki, M., Katsuragawa, Y., Koyari, R., Hendrix, D., . . . Nonogaki, H. (2019). *DELAY OF GERMINATION1-LIKE4* acts as an inducer of seed reserve accumulation. *Plant J*. doi:10.1111/tpj.14485.

Sallon, S., Solowey, E., Cohen, Y., Korchinsky, R., Egli, M., Woodhatch, I., . . . Kislev, M. (2008). Germination, genetics, and growth of an ancient date seed. *science, 320*(5882), 1464-1464.

Sano, N., Ono, H., Murata, K., Yamada, T., Hirasawa, T., & Kanekatsu, M. (2015). Accumulation of long-lived mRNAs associated with germination in embryos during seed development of rice. *Journal of experimental botany, 66*(13), 4035-4046.

Sano, N., Permana, H., Kumada, R., Shinozaki, Y., Tanabata, T., Yamada, T., . . . Kanekatsu, M. (2012). Proteomic analysis of embryonic proteins synthesized from long-lived mRNAs during germination of rice seeds. *Plant and cell physiology, 53*(4), 687-698.

Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M., & Lepiniec, L. (2008). Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. *The Plant Journal, 54*(4), 608-620.

Sato, K., Matsumoto, T., Ooe, N., & Takeda, K. (2009). Genetic analysis of seed dormancy QTL in barley. *Breeding science, 59*(5), 645-650.

Sattler, S. E., Gilliland, L. U., Magallanes-Lundback, M., Pollard, M., & DellaPenna, D. (2004). Vitamin E is essential for seed longevity and for preventing lipid peroxidation during germination. *The Plant Cell, 16*(6), 1419-1432.

Savadi, S. (2018). Molecular regulation of seed development and strategies for engineering seed size in crop plants. *Plant Growth Regulation, 84*(3), 401-422.

Schatzki, J., Schoo, B., Ecke, W., Herrfurth, C., Feussner, I., Becker, H. C., & Möllers, C. (2013). Mapping of QTL for seed dormancy in a winter oilseed rape doubled haploid population. *Theoretical and Applied Genetics, 126*(9), 2405-2415.

Schnittger, A., Jurgens, G., & Hulskamp, M. (1998). Tissue layer and organ specificity of trichome formation are regulated by *GLABRA1* and *TRIPTYCHON* in *Arabidopsis*. *Development, 125*(12), 2283-2289.

Scholl, R. L., May, S. T., & Ware, D. H. (2000). Seed and molecular resources for Arabidopsis. *Plant physiology, 124*(4), 1477-1480.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., . . . Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC molecular biology, 7*(1), 1-14.

Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., . . . Kunze, R. (2003). ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. *Plant physiology, 131*(1), 16-26.

Seo, P. J., Kim, M. J., Park, J. Y., Kim, S. Y., Jeon, J., Lee, Y. H., . . . Park, C. M. (2010). Cold activation of a plasma membrane-tethered NAC transcription factor induces a pathogen resistance response in Arabidopsis. *The Plant Journal, 61*(4), 661-671.

Seo, P. J., Kim, M. J., Song, J. S., Kim, Y. S., Kim, H. J., & Park, C. M. (2010). Proteolytic processing of an Arabidopsis membrane-bound NAC transcription factor is triggered by cold-induced changes in membrane fluidity. *Biochem J, 427*(3), 359-367. doi:10.1042/BJ20091762.

Seo, P. J., Kim, S.-G., & Park, C.-M. (2008). Membrane-bound transcription factors in plants. *Trends in plant science, 13*(10), 550-556.

Shaulian, E., & Karin, M. (2001). AP-1 in cell proliferation and survival. *Oncogene, 20*(19), 2390-2400.

Shen-Miller, J., Mudgett, M. B., Schopf, J. W., Clarke, S., & Berger, R. (1995). Exceptional seed longevity and robust growth: ancient sacred lotus from China. *American Journal of Botany, 82*(11), 1367-1380.

Shinomura, T., Nagatani, A., Hanzawa, H., Kubota, M., Watanabe, M., & Furuya, M. (1996). Action spectra for phytochrome A-and B-specific photoinduction of seed germination in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, 93*(15), 8129-8133.

Shinozaki, K., Yamaguchi-Shinozaki, K., Urao, T., & Koizumi, M. (1992). Nucleotide sequence of a gene from *Arabidopsis thaliana* encoding a *myb* homologue. *Plant molecular biology, 19*(3), 493-499.

Singh, K. L., Chaudhuri, A., & Kar, R. K. (2014). Superoxide and its metabolism during germination and axis growth of *Vigna radiata* (L.) Wilczek seeds. *Plant Signaling & Behavior, 9*(8), e29278.

Sneath, P., & Sokal, R. (1973). Numerical Taxonomy WH Freeman and Co. *San Francisco*, 1-573.

Somerville, C., & Koornneef, M. (2002). A fortunate choice: the history of Arabidopsis as a model plant. *Nature Reviews Genetics, 3*(11), 883-889.

Soppe, W. J., & Bentsink, L. (2020). Seed dormancy back on track; its definition and regulation by *DOG1*. *New Phytologist*.

Souer, E., van Houwelingen, A., Kloos, D., Mol, J., & Koes, R. (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.

Spannagl, M., Mayer, K., Durner, J., Haberer, G., & Fröhlich, A. (2011). Exploring the genomes: from Arabidopsis to crops. *Journal of plant physiology, 168*(1), 3-8.

Steber, C. M., & McCourt, P. (2001). A role for brassinosteroids in germination in Arabidopsis. *Plant physiology, 125*(2), 763-769.

Stephenson, P., Stacey, N., Brüser, M., Pullen, N., Ilyas, M., O'Neill, C., . . . Østergaard, L. (2019). The power of model-to-crop translation illustrated by reducing seed loss from pod shatter in oilseed rape. *Plant reproduction, 32*(4), 331-340.

Stewart, R. R., & Bewley, J. D. (1980). Lipid peroxidation associated with accelerated aging of soybean axes. *Plant physiology, 65*(2), 245-248.

Stone, S. L., Braybrook, S. A., Paula, S. L., Kwong, L. W., Meuser, J., Pelletier, J., . . . Harada, J. J. (2008). Arabidopsis *LEAFY COTYLEDON2* induces maturation traits and auxin activity: implications for somatic embryogenesis. *Proceedings of the National Academy of Sciences, 105*(8), 3151-3156.

Su, H., Zhang, S., Yin, Y., Zhu, D., & Han, L. (2015). Genome-wide analysis of *NAM-ATAF1*, *2-CUC2* transcription factor family in *Solanum lycopersicum*. *Journal of plant biochemistry and biotechnology, 24*(2), 176-183.

Subbiah, V., & Reddy, K. J. (2010). Interactions between ethylene, abscisic acid and cytokinin during germination and seedling establishment in Arabidopsis. *Journal of biosciences, 35*(3), 451-458.

Sun, W. Q., & Leopold, A. C. (1994). Glassy state and seed storage stability: a viability equation analysis. *Annals of botany, 74*(6), 601-604.

Sun, W. Q., & Leopold, A. C. (1995). The Maillard reaction and oxidative stress during aging of soybean seeds. *Physiologia Plantarum, 94*(1), 94-104.

Suzuki, Y. J., Carini, M., & Butterfield, D. A. (2010). Protein carbonylation. *Antioxidants & redox signaling, 12*(3), 323-325.

Swarbreck, D., Wilks, C., Lamesch, P., Berardini, T. Z., Garcia-Hernandez, M., Foerster, H., . . . Ploetz, L. (2007). The Arabidopsis Information Resource (TAIR): gene structure and function annotation. *Nucleic acids research, 36*(suppl_1), D1009-D1014.

Takada, S., Hibara, K.-i., Ishida, T., & Tasaka, M. (2001). The *CUP-SHAPED COTYLEDON1* gene of Arabidopsis regulates shoot apical meristem formation. *Development, 128*(7), 1127-1135.

Tang, D.-S., Hamayun, M., Ko, Y.-M., Zhang, Y.-P., Kang, S.-M., & Lee, I.-J. (2008). Role of red light, temperature, stratification and nitrogen in breaking seed dormancy of *Chenopodium album* L. *Journal of Crop Science and Biotechnology, 11*(3), 199-204.

Tang, L. P., Zhou, C., Wang, S. S., Yuan, J., Zhang, X. S., & Su, Y. H. (2017). *FUSCA 3* interacting with *LEAFY COTYLEDON 2* controls lateral root formation through regulating *YUCCA 4* gene expression in *Arabidopsis thaliana*. *New Phytologist, 213*(4), 1740-1754.

Terskikh, V. V., Zeng, Y., Feurtado, J. A., Giblin, M., Abrams, S. R., & Kermode, A. R. (2008). Deterioration of western redcedar (*Thuja plicata* Donn ex D. Don) seeds: protein oxidation and *in vivo* NMR monitoring of storage oils. *Journal of experimental botany, 59*(4), 765-777.

Tesnier, K., Strookman-Donkers, H., Van Pijlen, J., Van der Geest, A., Bino, R., & Groot, S. (2002). A controlled deterioration test for *Arabidopsis thaliana* reveals genetic variation in seed quality. *Seed Science and Technology, 30*(1), 149-166.

Tian, H., Wang, X., Guo, H., Cheng, Y., Hou, C., Chen, J.-G., & Wang, S. (2017). NTL8 regulates trichome formation in Arabidopsis by directly activating R3 MYB genes *TRY* and *TCL1*. *Plant physiology, 174*(4), 2363-2375.

To, A., Valon, C., Savino, G., Guilleminot, J., Devic, M., Giraudat, J., & Parcy, F. (2006). A network of local and redundant gene regulation governs Arabidopsis seed maturation. *The Plant Cell, 18*(7), 1642-1651.

Tommasi, F., Paciolla, C., de Pinto, M. C., & Gara, L. D. (2001). A comparative study of glutathione and ascorbate metabolism during germination of *Pinus pinea* L. seeds. *Journal of experimental botany, 52*(361), 1647-1654.

Torada, A., Ikeguchi, S., & Koike, M. (2005). Mapping and validation of PCR-based markers associated with a major QTL for seed dormancy in wheat. *Euphytica, 143*(3), 251-255.

Tourrière, H., Chebli, K., & Tazi, J. (2002). mRNA degradation machines in eukaryotic cells. *Biochimie, 84*(8), 821-837.

Tran, L.-S. P., Nakashima, K., Sakuma, Y., Simpson, S. D., Fujita, Y., Maruyama, K., . . . Yamaguchi-Shinozaki, K. (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.

Turrens, J. F. (2003). Mitochondrial formation of reactive oxygen species. *The Journal of physiology, 552*(2), 335-344.

Uauy, C. (2017). Wheat genomics comes of age. *Current opinion in plant biology, 36*, 142-148.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., & Leunissen, J. A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic acids research, 35*(suppl_2), W71-W74.

van der Schaar, W., Alonso-Blanco, C., Léon-Kloosterziel, K. M., Jansen, R. C., van Ooijen, J. W., & Koornneef, M. (1997). QTL analysis of seed dormancy in Arabidopsis using recombinant inbred lines and MQM mapping. *Heredity, 79*(2), 190-200.

Vellvé, R. (2013). *Saving the seed: genetic diversity and European agriculture*: Routledge.

Vidigal, D. S., Marques, A. C., Willems, L. A., Buijs, G., Méndez-Vigo, B., Hilhorst, H. W., . . . Alonso-Blanco, C. (2016). Altitudinal and climatic associations of seed dormancy and flowering traits evidence adaptation of annual life cycle timing in *Arabidopsis thaliana*. *Plant, cell & environment, 39*(8), 1737-1748.

Vik, Å., & Rine, J. (2000). Membrane biology: membrane-regulated transcription. *Current Biology, 10*(23), R869-R871.

Vranová, E., Inzé, D., & Van Breusegem, F. (2002). Signal transduction during oxidative stress. *Journal of experimental botany, 53*(372), 1227-1236.

Vroemen, C. W., Mordhorst, A. P., Albrecht, C., Kwaaitaal, M. A., & de Vries, S. C. (2003). The *CUP-SHAPED COTYLEDON3* gene is required for boundary and shoot meristem formation in Arabidopsis. *The Plant Cell, 15*(7), 1563-1577.

Walters, C. (1998). Understanding the mechanisms and kinetics of seed aging. *Seed Science Research, 8*(2), 223-244.

Walters, C., Hill, L. M., & Wheeler, L. J. (2005). Dying while dry: kinetics and mechanisms of deterioration in desiccated organisms. *Integrative and Comparative Biology, 45*(5), 751-758.

Walters, C., Reilley, A. A., Reeves, P. A., Baszczak, J., & Richards, C. M. (2006). The utility of aged seeds in DNA banks.

Walters, C., Wheeler, L. M., & Grotenhuis, J. M. (2005). Longevity of seeds stored in a genebank: species characteristics. *Seed Science Research, 15*(1), 1-20.

Wan, L., Ross, A. R., Yang, J., Hegedus, D. D., & Kermode, A. R. (2007). Phosphorylation of the 12 S globulin cruciferin in wild-type and *abi1-1* mutant *Arabidopsis thaliana* (thale cress) seeds. *Biochemical Journal, 404*(2), 247-256.

Wang, Q., Lin, Q., Wu, T., Duan, E., Huang, Y., Yang, C., . . . Xie, K. (2020). *OsDOG1L-3* regulates seed dormancy through the abscisic acid pathway in rice. *Plant Science, 298*, 110570.

Wang, S., Kwak, S.-H., Zeng, Q., Ellis, B. E., Chen, X.-Y., Schiefelbein, J., & Chen, J.-G. (2007). *TRICHOMELESS1* regulates trichome patterning by suppressing *GLABRA1* in Arabidopsis. *Development, 134*(21), 3873-3882.

Wang, Y., Li, L., Ye, T., Zhao, S., Liu, Z., Feng, Y. Q., & Wu, Y. (2011). Cytokinin antagonizes ABA suppression to seed germination of Arabidopsis by downregulating ABI5 expression. *The Plant Journal, 68*(2), 249-261.

Wareing, P. F. (1971). The control of seed dormancy. *Biochem J, 124*(2), 1P-2P. doi:10.1042/bj1240001p

Weigel, D., & Mott, R. (2009). The 1001 genomes project for *Arabidopsis thaliana*. *Genome biology, 10*(5), 1-5.

Weir, I., Lu, J., Cook, H., Causier, B., Schwarz-Sommer, Z., & Davies, B. (2004). *CUPULIFORMIS* establishes lateral organ boundaries in *Antirrhinum*. *Development, 131*(4), 915-922.

Wiebach, J., Nagel, M., Börner, A., Altmann, T., & Riewe, D. (2020). Age-dependent loss of seed viability is associated with increased lipid oxidation and hydrolysis. *Plant, cell & environment, 43*(2), 303-314.

Wienkoop, S., Baginsky, S., & Weckwerth, W. (2010). *Arabidopsis thaliana* as a model organism for plant proteome research. *Journal of proteomics, 73*(11), 2239-2248.

Wieser, R., Adam, G., Wagner, A., Schüller, C., Marchler, G., Ruis, H., . . . Bilinski, T. (1991). Heat shock factor-independent heat control of transcription of the *CTT1* gene encoding the Cytosolic Catalase T of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry, 266*(19), 12406-12411.

Wing, R. A., Purugganan, M. D., & Zhang, Q. (2018). The rice genome revolution: from an ancient grain to Green Super Rice. *Nature Reviews Genetics, 19*(8), 505-517.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., & Provart, N. J. (2007). An “Electronic Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS One, 2*(8), e718.

Wojtyla, Ł., Garnczarska, M., Zalewski, T., Bednarski, W., Ratajczak, L., & Jurga, S. (2006). A comparative study of water distribution, free radical production and activation of antioxidative metabolism in germinating pea seeds. *Journal of plant physiology, 163*(12), 1207-1220.

Wray, G. A., Hahn, M. W., Abouheif, E., Balhoff, J. P., Pizer, M., Rockman, M. V., & Romano, L. A. (2003). The evolution of transcriptional regulation in eukaryotes. *Molecular biology and evolution, 20*(9), 1377-1419.

Wu, A., Allu, A. D., Garapati, P., Siddiqui, H., Dortay, H., Zanor, M.-I., . . . Tohge, T. (2012). *JUNGBRUNNEN1*, a reactive oxygen species–responsive NAC transcription factor, regulates longevity in Arabidopsis. *The Plant Cell, 24*(2), 482-506.

Xiang, Y., Nakabayashi, K., Ding, J., He, F., Bentsink, L., & Soppe, W. J. (2014). *Reduced Dormancy5* encodes a protein phosphatase 2C that is required for seed dormancy in Arabidopsis. *The Plant Cell, 26*(11), 4362-4375.

Xiang, Y., Song, B., Née, G., Kramer, K., Finkemeier, I., & Soppe, W. J. (2016). Sequence polymorphisms at the *REDUCED DORMANCY5* pseudophosphatase underlie natural variation in Arabidopsis dormancy. *Plant physiology, 171*(4), 2659-2670.

Xie, K., Jiang, L., Lu, B., Yang, C., Li, L., Liu, X., . . . Wan, J. (2011). Identification of QTLs for seed dormancy in rice (*Oryza sativa* L.). *Plant breeding, 130*(3), 328-332.

Xie, Q., Frugis, G., Colgan, D., & Chua, N.-H. (2000). Arabidopsis *NAC1* transduces auxin signal downstream of *TIR1* to promote lateral root development. *Genes & development, 14*(23), 3024-3036.

Xie, Q., Sanz-Burgos, A. P., Guo, H., García, J. A., & Gutiérrez, C. (1999). GRAB proteins, novel members of the NAC domain family, isolated by their interaction with a geminivirus protein. *Plant molecular biology, 39*(4), 647-656.

Xu, M. Y., Dong, Y., Zhang, Q. X., Zhang, L., Luo, Y. Z., Sun, J., . . . Wang, L. (2012). Identification of miRNAs and their targets from *Brassica napus* by high-throughput sequencing and degradome analysis. *BMC genomics, 13*(1), 421.

Yamaguchi, S., Kamiya, Y., & Nambara, E. (2018). Regulation of ABA and GA levels during seed development and germination in Arabidopsis. *Annual Plant Reviews online*, 224-247.

Yamasaki, Y., Gao, F., Jordan, M. C., & Ayele, B. T. (2017). Seed maturation associated transcriptional programs and regulatory networks underlying genotypic difference in seed dormancy and size/weight in wheat (*Triticum aestivum* L.). *BMC plant biology, 17*(1), 1-18.

Yan, L. L., & Zaher, H. S. (2019). How do cells cope with RNA damage and its consequences? *Journal of Biological Chemistry, 294*(41), 15158-15171.

Yang, S.-D., Seo, P. J., Yoon, H.-K., & Park, C.-M. (2011). The Arabidopsis NAC transcription factor *VNI2* integrates abscisic acid signals into leaf senescence via the *COR*/*RD* genes. *The Plant Cell, 23*(6), 2155-2168.

Yang, Z.-T., Wang, M.-J., Sun, L., Lu, S.-J., Bi, D.-L., Sun, L., . . . Liu, J.-X. (2014). The membrane-associated transcription factor *NAC089* controls ER-stress-induced programmed cell death in plants. *PLoS Genet, 10*(3), e1004243.

Yatusevich, R., Fedak, H., Ciesielski, A., Krzyczmonik, K., Kulik, A., Dobrowolska, G., & Swiezewski, S. (2017). Antisense transcription represses Arabidopsis seed dormancy QTL *DOG1* to regulate drought tolerance. *EMBO Rep, 18*(12), 2186-2196. doi:10.15252/embr.201744862.

Yazdanpanah, F., Hanson, J., Hilhorst, H. W. M., & Bentsink, L. (2017). Differentially expressed genes during the imbibition of dormant and after-ripened seeds - a reverse genetics approach. *BMC Plant Biol, 17*(1), 151. doi:10.1186/s12870-017-1098-z.

Yoshida, T., Nishimura, N., Kitahata, N., Kuromori, T., Ito, T., Asami, T., . . . Hirayama, T. (2006). *ABA-hypersensitive germination3* encodes a protein phosphatase 2C (AtPP2CA) that strongly regulates abscisic acid signaling during germination among Arabidopsis protein phosphatase 2Cs. *Plant Physiol, 140*(1), 115-126. doi:10.1104/pp.105.070128.

Yu, B., Wang, Y., Zhou, H., Li, P., Liu, C., Chen, S., . . . Teng, S. (2020). Genome-wide binding analysis reveals that *ANAC060* directly represses sugar-induced transcription of *ABI5* in Arabidopsis. *The Plant Journal*.

Zhang, H., He, D., Yu, J., Li, M., Damaris, R. N., Gupta, R., . . . Yang, P. (2016). Analysis of dynamic protein carbonylation in rice embryo during germination through AP-SWATH. *Proteomics, 16*(6), 989-1000.

Zhang, M., Cao, X., Jia, Q., & Ohlrogge, J. (2016). *FUSCA 3* activates triacylglycerol accumulation in Arabidopsis seedlings and tobacco BY 2 cells. *The Plant Journal, 88*(1), 95-107.

Zhao, L., Wang, S., Fu, Y.-B., & Wang, H. (2020). Arabidopsis seed stored mRNAs are degraded constantly over aging time, as revealed by new quantification methods. *Frontiers in Plant Science, 10*, 1764.

Zhao, T., Holmer, R., de Bruijn, S., Angenent, G. C., van den Burg, H. A., & Schranz, M. E. (2017). Phylogenomic synteny network analysis of MADS-box transcription factor genes reveals lineage-specific transpositions, ancient tandem duplications, and deep positional conservation. *The Plant Cell, 29*(6), 1278-1292.

Zhao, T., & Schranz, M. E. (2019). Network-based microsynteny analysis identifies major differences and genomic outliers in mammalian and angiosperm genomes. *Proceedings of the National Academy of Sciences, 116*(6), 2165-2174.

Zhu, T., Nevo, E., Sun, D., & Peng, J. (2012). Phylogenetic analyses unravel the evolutionary history of NAC proteins in plants. *Evolution: International Journal of Organic Evolution, 66*(6), 1833-1848.

Zoeller, M., Stingl, N., Krischke, M., Fekete, A., Waller, F., Berger, S., & Mueller, M. J. (2012). Lipid profiling of the Arabidopsis hypersensitive response reveals specific lipid peroxidation and fragmentation processes: biogenesis of pimelic and azelaic acid. *Plant physiology, 160*(1), 365-378.

Zuckerkandl, E., & Pauling, L. (1965). Evolutionary divergence and convergence in proteins. In *Evolving genes and proteins* (pp. 97-166). Academic Press.

Summary

Seed dormancy and seed longevity are two critical seed traits contributing to seed quality. Seed dormancy is the phenomenon that fully developed seeds can temporarily not germinate. This inhibition of germination disappears naturally and gradually during dry seed storage. Seed longevity refers to the maximum time after harvest that seeds keep the ability to germinate. These two seed traits have similarities, both are closely related to seed maturation and together they constitute the seed life span. Moreover, there is no clear separation in the time between when seed dormancy is released and the point at which ageing starts. Knowledge on the regulation of seed dormancy and seed longevity would allow to have a better control of the timing of seed germination and to increase the life time of seeds during storage. Thus, improving seed quality and increase crop yields. The underlying network regulating these two traits are complex and still largely unexplored. The current knowledge on the regulators of seed dormancy and seed longevity has been introduced in **Chapter 1**.

Seed dormancy is an adaptive trait which is strongly influenced by the surrounding environment, as a result, there are accessions exhibiting different levels of seed dormancy. The presence of natural variation for seed dormancy in accessions of *Arabidopsis thaliana* has allowed the identification of dormancy quantitative trait loci. In **Chapter 2** the positional cloning and functional analyses of *DELAY OF GERMINATION 6,* a strong seed dormancy locus identified by the before mentioned quantitative trait loci, analyses are described. *DELAY OF GERMINATION 6* encodes *AT3G44290* which is a NAC transcription factor. NAC transcription factors are known to contain a conserved NAC domain and various C-terminals, which determine the presence of the transmembrane domain (TMD). *ANAC060* was sequenced in different accessions which revealed distinct polymorphisms. For instance, the Landsberg *erecta* (L*er*) *ANAC060* allele lacked the complete fourth exon that encodes the TMD, whereas Kashmir (Kas) allele contained a full length functional TMD. In addition, the Columbia-0 allele was found to be alternatively spliced, resulting in two forms of the cDNA that exist at the same time. Phenotypic analyses revealed opposing seed dormancy levels depending on the presence or the absence of the TMD. Ler_*ANAC060* seeds lacking the TMD had low seed dormancy whereas Kas-2_*ANAC060* seeds containing the TMD were highly dormant. As known, the TMD regulates transcription factors by influencing their localization. In order to check this, transgenic GFP lines containing the native *ANAC060* promoter and the respective *ANAC060* cDNAs cloned from L*er* and Kas-2 were made. Confocal

analyses showed that the ANAC060 protein without TMD was localized in nucleus, whereas the protein containing the TMD anchored at the nuclear membrane. PROTEIN PHOSPHATASE 2C class A proteins were identified as putative ANAC060 targets based on combining different studies: including public available data of a ChIP-Seq experiment performed in young seedlings after sugar treatment, gene expression analyses in seeds of L*er* and a *DOG6*-Kas near isogenic line and phenotypic analyses of double mutants of the *anac060* mutant and ABA related mutants. Moreover, through real-time quantitative polymerase chain reaction (RT-qPCR), *DOG1* expression was found to be partly regulated by *ANAC060*.

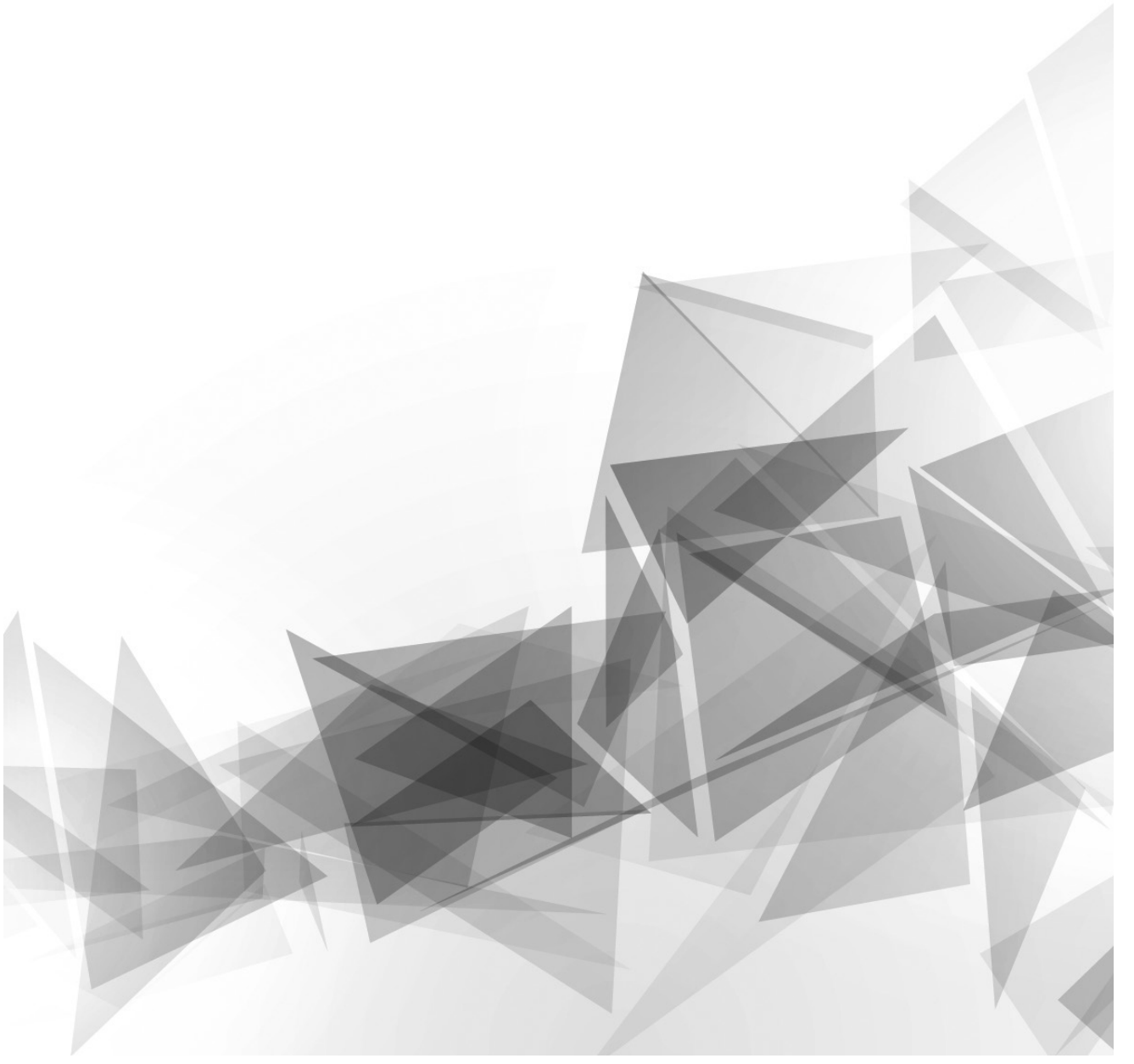
ANAC060 was founded to have two homologous genes, *ANAC040* and *ANAC089*. In **Chapter 3**, these three genes were further studied by comparing dormancy phenotypes and sequence similarities. It turned out that *ANAC040* and *ANAC089* did not regulate primary seed dormancy. Through a synteny network analysis, *ANAC040* was predicted to be the ancestor gene of *ANAC060* and *ANAC089*, which both were derived during the At-Alpha whole genome duplication event. Transgenic swapping experiments through swapping the promoters and coding regions, have been performed to identify possible redundancy between *ANAC060, ANAC040* and *ANAC089*. It was shown that the coding elements of *ANAC040* and *ANAC060* were exchangeable. The transgenic line containing the promoter of *ANAC040* and coding sequence of *ANAC060* was able to release seed dormancy in *anac060* mutant background. The functional difference between the full length *ANAC040* and *ANAC060* genes was hypothesized to be encoded in the promoters, leading to differential spatial and temporal expression patterns. In order to further identify the differences in the promoters of *ANAC060* and *ANAC040*, motif analyses were performed. Twelve shared motifs were exhibited in the promoter regions of both genes, four motifs were only present in promoter of *ANAC060* including W-box, AP-1, MYC and CAT-box. The motifs, STRE and the GA-motif, existed only in promoter of *ANAC040*. Whether these motifs determine the differences in expression still remains to be investigated.

Arabidopsis seeds can be stored over ten years in laboratory condition which is a too long time to obtain seeds with different ageing levels within a PhD project. In order to shorten the storage times elevated partial pressure of oxygen was utilized in **Chapter 4**. In total, nine different ageing states including non-EPPO treated seeds were collected. As a vulnerable target of oxygen, total RNA integrity was examined and this revealed that there was no correlation between RNA integrity and seed

ageing states. In order to investigated if the degradation pattern of mRNA correlated to changes in seed life span, three different seed ageing states were selected and used for degradome analyses. For this the 5' monophosphate mRNAs with an intact 3' poly(A) tail were isolated and sequenced. Cluster analyses revealed seven different degradation patterns. Combining these analyses with an earlier study, in which mRNAs that were bound to monosomes (single ribosome, that might play a role in mRNA protection in dry seeds) were identified, showed that most of the monosome bound mRNAs had up-regulated degradation patterns. To focus on specific transcripts, differential expression analyses was performed. During dormancy release, there were only seven transcripts up-regulated in the degradome, many more were down-regulated (133). Comparing the fully after-ripened non dormant seeds to the aged, non-germinating seed revealed 546 and 425 transcripts up-and down- regulated in the degradome respectively. The expression levels of twenty-three of these transcripts were further investigated by RT-qPCR in both dry and imbibed seeds. Four representative patterns were displayed that were described based on example genes. For each the expression pattern in dry and imbibed seeds was compared to patterns of mRNA translation and degradation. Overall a correlation between ageing states, mRNA degradation patterns and expression in aged imbibed seeds was identified, indicating that the failure of repairing certain degraded mRNA after imbibition might lead to seed death. Moreover, protein carbonylation patterns at the same three physiological seed stages and in seeds of seed storage mutants were observed, this confirmed the protective role of seed storage proteins during seed ageing.

In **Chapter 5**, the findings of this thesis are discussed in a broader context. I discussed the occurrence of natural genetic variation of the presence/absence of the transmembrane domain (TMD) and the possible regulatory mechanisms for the cleavage of TMDs. Also the potential of using the elevated partial pressure of oxygen method for understanding seed ageing was evaluated. Lastly, the barriers for transferring the knowledge from Arabidopsis to crops were discussed in respect of unsolved issues and the need to solve these in the near future.

Samenvatting



De zaadeigenschappen kiemrust en de bewaarbaarheid bepalen in grote mate de kwaliteit van zaden. Kiemrust is het fenomeen dat volledig ontwikkelde zaden tijdelijk niet kunnen kiemen. Deze inhibitie van kieming verdwijnt geleidelijk wanneer zaden droog bewaard worden. De bewaarbaarheid van zaden refereert naar de maximum tijd dat zaden na oogst bewaard kunnen worden. Deze twee eigenschappen hebben overeenkomstig dat ze gerelateerd zijn aan zaad rijping en dat ze samen de levensduur van een zaad bepalen. Daarnaast is er geen duidelijke scheiding tussen de fase waarin kiemrust gebroken wordt en wanneer de veroudering van een zaad start. Kennis over de regulatie van kiemrust en bewaarbaarheid van zaden kan zorgen voor een betere timing van kieming en een langere levensduur tijdens bewaring. Dit leidt tot een verbeterde kwaliteit van gewaszaden en een hogere opbrengt. Het netwerk van genen dat beide eigenschappen reguleert is complex en nog grotendeels onontgonnen. De huidige kennis over de regulatoren van kiemrust en bewaarbaarheid wordt in **Hoofdstuk 1** geïntroduceerd.

Kiemrust is een adaptieve eigenschap die sterk beïnvloed wordt door de omgeving, met als resultaat een grote variatie in kiemrust niveaus. Variatie voor kiemrust in accessies van *Arabidopsis thaliana* heeft het identificeren van kiemrust loci (quantitative trait loci; QTL) mogelijk gemaakt. In **Hoofdstuk 2**, is de precieze kartering en klonering van *DELAY OF GERMINATION 6,* een sterk kiemrust locus geïdentificeerd in de hierboven genoemde genetische analyses, beschreven. *DELAY OF GERMINATION 6* codeert de NAC transcriptie factor *AT3G44290*. NAC transcriptie factoren bevatten een geconserveerd NAC domein en een meer variabel C-uiteinde waarin het transmembraan domein (TMD) ligt. *ANAC060* sequentie analyses in verschillende accessies heeft verscheidene polymorfismen geïdentificeerd. Bijvoorbeeld, het Landsberg *erecta* (L*er*) *ANAC060* allel mist het volledige vierde exon dat codeert voor het TMD, terwijl het Kashmir (Kas) allel een volledig TMD bevat. Daarnaast blijkt het Columbia-0 allel twee verschillende splicings varianten te kennen, die beide tegelijkertijd voorkomen. Fenotypische analyses hebben tegengestelde kiemrust niveaus aangetoond, afhankelijk van het wel of niet aanwezig zijn van het TMD. Het Ler_*ANAC060* allel waarin het TMD

ontbreekt had een lage kiemrust, terwijl Kas-2_*ANAC060* zaden die het TMD bevatten behoorlijk dormant zijn. Zoals bekend reguleert het TMD, transcriptie factoren door hun lokalisatie te beïnvloeden. Om dit te bevestigen zijn er transgene GFP lijnen gemaakt waarbij de L*er* en Kas-2 ANAC060 cDNAs gekloneerd zijn onder de gen eigen promotor. Confocale analyses laten zien dat het ANAC060 eiwit zonder het TMD in de kern van de cel zit, terwijl het eiwit dat de TMD bevat aan de kernmembraan gekoppeld is. PROTEIN PHOSPHATASE 2C class A eiwitten worden vermoedelijk gereguleerd door ANAC060, deze conclusie is gebaseerd op verschillende resultaten o.a. een publiekelijk beschikbaar CHIP-Seq experiment waarin jonge zaailingen met suiker behandeld zijn, gen expressie data van L*er* en de DOG6-Kas bijna isogene lijn en fenotypische analyses van dubbel mutanten van de *anac060* mutant en ABA gerelateerde mutanten. Verder wees een kwantitatieve PCR (RT-qPCR) uit dat *DOG1* expressie gedeeltelijk door *ANAC060* gereguleerd wordt.

ANAC060 heeft twee homologe genen, *ANAC040* en *ANAC089*. In **Hoofdstuk 3** zijn deze genen nader bestudeerd door kiemrust fenotypes en sequenties te vergelijken. Hieruit bleek dat *ANAC040* en *ANAC089* geen primaire kiemrust reguleren. Synteny netwerk analyses voorspellen dat *ANAC060* en *ANAC089* afstammen van *ANAC040*, een resultaat van de At-Alpha genoom verdubbeling. Het uitruilen van promotoren en coderende sequenties heeft het mogelijk gemaakt de redundantie tussen *ANAC060, ANAC040* en *ANAC089* inzichtelijk te maken. De coderende elementen van *ANAC040* en *ANAC060* bleken uitwisselbaar. De transgene lijn met de promotor van *ANAC040* en de coderende sequentie van *ANAC060* was in staat de kiemrust van de *anac060* mutant op te heffen. Het functionele verschil tussen *ANAC040* en *ANAC060* genen wordt verondersteld te worden gecodeerd door hun promotoren, welke leiden tot verschillen in waar en wanneer de genen tot expressie komen. Om de verschillende in de *ANAC060* en *ANAC040* promotoren te identificeren is een analyse naar regulerende patronen uitgevoerd. Twaalf patronen kwamen voor in beide promotoren, terwijl vier patronen specifiek waren voor de *ANAC060* promoter, waaronder W-box, AP-1, MYC en CAT-box. De patronen STRE

en het GA-patroon kwamen alleen in de *ANAC040* promotor voor. Of deze patronen ook daadwerkelijk de verschillen in expressie bepalen dient nog verder te worden uitgezocht.

Arabidopsis zaden kunnen in laboratorium condities langer dan tien jaar bewaard worden, wat te lang is om binnen een AIO project zaden met verschillende verouderingsniveaus te verkrijgen. Om de verouderingstijd te verkorten is er in **Hoofdstuk 4** gebruik gemaakt van verhoogde zuurstofdruk. In totaal zijn er negen verschillende verouderingsstadia, inclusief onbehandelde zaden, verzameld. De RNA integriteit was bepaald omdat RNA vatbaar is voor oxidatie, echter er bleek geen correlatie te zijn tussen RNA integriteit en de verschillende verouderingsstadia. Om te bepalen of het degradatie patroon van het mRNA correleert met veranderingen tijdens de levensduur van zaden zijn drie verouderingsstadia gebruikt voor “degradome” analyses. Hiervoor zijn 5’ monofosfaat mRNAs met een intacte 3’ poly(A) staart geïsoleerd en gesequenced. Cluster analyses laten zeven verschillende degradatie patronen zien. Gecombineerd met een eerdere studie, waarin monosoom (een enkel ribosoom, welke mogelijk mRNAs beschermd) gebonden mRNAs geïdentificeerd zijn, laat deze data zien dan monosoom gebonden mRNA op-gereguleerde degradatie patronen hebben. Om specifieke transcripten te bekijken is differentiële expressie analyse uitgevoerd. In de fase waarin kiemrust afneemt waren er slechts zeven transcripten op gereguleerd, veel meer waren er neerwaarts gereguleerd (133). De vergelijking tussen volledig na gerijpte niet dormante zaden en verouderde zaden resulteerde in 546 en 425 transcripten die respectievelijk omhoog of neerwaarts gereguleerd waren in het “degradome”. De expressie van drieëntwintig transcripten zijn met behulp van RT-qPCR verder geanalyseerd in zowel droog als geïmbibeerd zaad. Vier representatieve patronen zijn beschreven middels een voorbeeld gen. Voor elk is het expressie patroon in droog en geïmbibeerd zaad vergeleken met de patronen van mRNA translatie en afbraak. In algemene zin is er een correlatie gevonden tussen verouderingsstadia, mRNA afbraak patronen en de expressie in geïmbibeerd zaad, wat suggereert dat een defect in de reparatie van mRNAs er toe leid dat een

zaad niet meer kiemt. Verder zijn in zaden met dezelfde drie verouderingsstadia en in zaden van mutanten die een defect hebben in het opslaan van reserve eiwitten de carbonylatie profielen van eiwitten bekeken, deze resultaten bevestigden dat opslageiwitten een beschermende rol hebben tijdens veroudering.

In **Hoofdstuk 5** worden de bevindingen van dit proefschrift in een bredere context bediscussieerd. Ik bespreek de natuurlijke variatie voor de aan- of afwezigheid van het transmembraan domein (TMD) en de mogelijke regulatie mechanismen die betrekking hebben op het verwijderen van het TMD. Verder evalueer ik het gebruik van de verhoogde zuurstofdruk methode voor het begrijpen van veroudering in zaden. Tenslotte, bediscussieer ik de barrières in het vertalen van kennis vergaard in Arabidopsis naar gewassen, in relatie tot openstaande vragen die in de recente toekomst opgelost dienen te worden.

Acknowledgements

It's time to thank to everyone who has helped me in my PhD and even my life, without anyone of you, this thesis could never be accomplished. Here I have a few pages to describe my appreciation.

First of all, my supervisor Leónie, I could never start and finish this PhD project without you. I still remember the first skype meeting between us, your serious face made me feel you would be a strict supervisor. But later on, it turns out you are a really warm and easy-going supervisor, and the best one who I ever have. You understood that I was having the language problem when I just started here, you would write really long email to describe the scientific points during our meeting as I might not understand totally. You always encouraged me in a positive way and kept telling me the most important was to self-development when I complained to you about experimental failure. Honestly, this was strange to me, I thought I always needed to have good results otherwise I would never graduate. Nevertheless, trying best to make experiment to be successful is always good. The private project discussion meeting started from my third year really helped me a lot, from you, I learned plenty of seed related knowledge and your creative minds also inspired me how to think. I also appreciated that you guided me not only scientifically but also personally. You had listened to my quite some complaints from my personal life and gave me suggestions about how to face and solve them. For finishing this thesis, you also put huge efforts and worked during weekend for my thesis before deadline, without your concrete help I could not image how I could finish this on time. You are an excellent seed scientist who also sets an example to me, I wish I could be like you in seed research in the future. There are so many other things still need to be mentioned, but I will leave it here, I'm happy to be one of your PhD students.

Thanks to Christa, you have brought an energetic and professional to Plant physiology in the middle of my PhD. I also made great friends with some of them and enjoy all the great and lovely atmosphere in our big group. Henk, thanks to the nice conservations when we were in coffee break room. Wilco, you are humorous and really helpful. You were always patient to solve my software problem. Richard, thanks to you for being my external advisor, you gave great suggestions during my

project meeting and offered chance for me to learn and perform ChIP experiment in your group. Charlotte, I like the teaching days which we spent together, I learned from you about how to be a good teacher. You also wrote a warm letter to me with so many complements on our group 'Valentine's day', thank you so much for your kindness. Sander, thanks to your creative suggestions during project meeting, presentation and literature study. Harm, I enjoyed working with you even in a such a short time, without your help, degradome analyses can never continue so smoothly. Thanks to Rina for organizing everything about documents.

Leo, you are number one technician who helped me in my full PhD life, you were always patient and willing to answer my 'silly' questions. You helped me a lot for my experiment and gave me valuable suggestions. Also, thanks for your feedbacks for my thesis writing. Although you are really mean to me sometimes, but I mean back to you always, so we are equal. Jessica, you are such a nice friend. You always helped me when I needed you in lab except you did not know neither. I was so happy to be invited to your place and really enjoyed your cooking. I also thanks to both Leo and Jessica for being my great paranymphs. Juriaan, you were teaching me how to do gateway cloning in the very first of my PhD, I learned a lot of practical lab skills from you. Thanks to Jan for helping to prepare the EPPO experiment setup. Jacqueline, even though you moved back to bioscience, you still were so kind to answer my questions when I met you in the corridor of lab. Froukje, thanks to you for teaching me how to perform ChIP experiment and being patient to discuss and answer my questions.

Bing, thanks to your valuable help in my PhD. You gave me advises not only on scientific area but also future career directions. You are like a colleague 'brother' to me. I was also happy to be invited to join nice and fun parties in your big house. Yanxia, you are like Bing to give to me great suggestions, colleague 'sister', I always had nice discussions if I had some questions when you passed my seat, or I met you in lad when you were free. Rumyana, thanks to your great suggestions on my degradation chapter and also on my other projects every time after my presentation. Bas, thanks to teach me the first experiment which I did in my PhD and also for the

primers, seeds you left to me. Anirban, you are a funny person, thanks for your help about making awesome confocal pictures and I also enjoyed the beer sessions which we had in front of forum. Thanks to Tao for the evolutionary analyses for *ANAC060*, I learned a lot from you.

Elise, thanks to your explanation about how to analysis germination test result, I still have the note which you taught me of each step. Mariana, thanks to your nice explanation about selection of transgenic lines in my every first of my PhD. Nikita, thanks to you for your kindness and happiness which you brought, also appreciated the help from you to my student Annika. Jasper, thanks to you for moving to the silence room, we always had short and commercial talk in the morning, and also really nice discussions when we had time. You were also super patient to explain to me the scientific questions if you could help, you know that “useless” was a joke, I appreciated all from you. Eva, you are a happy and lovely girl in our group. Patricija, thanks for your delicious hand make cakes on almost every Monday. Hongfei, Yutao, Zhenbiao, thanks to your kindness, we had nice talks and hang outs, I had great time with you guys. Baojian, you were so nice to help me when I needed something in lad and also explained to me patiently when I asked you questions. Rufang, you are a hard worker and an example to me. We also had nice walk and talks together. Last but not least, Kees, my soulmate, I definitely want to say a lot about you, but I prefer to keep it low key here, hhh.

Thanks to all my students, Iona, you were my first student, thanks to your great attitude. Pallavi, thanks for your hard work, I hope you would get great PhD position. Annika, you are a lovely, independent and energetic student. I also thank to your hand made hair bands which you sent to me as gifts, they are so beautiful although I have short hair now. But I will keep them forever and may use them in the future. I feel so happy for you to get PhD position in seed lab, it’s such a pity that we cannot be colleagues. Lot, you are a so sunny girl, you were considerate about experiment and paid attention to small but important details. You were also a hard worker, even you had exam in between, you still tried to make two hours to work on the project. I hope you could enjoy the coming master life. Ao, you were not only a student to me,

but you also helped so much and worked a lot on my project. You are a nice friend and family member to me, your mom and stepfather also supported at any time when I need them. We had so many happy travels together in different countries. Also, with your parents, we had so nice road trip and I enjoyed real French Christmas, it was so interesting to me. I appreciated about everything you have brought to me. I hope you could get a good PhD position and see you soon. Max, you are my last master student during my PhD, I'm really happy to work with you. Thanks for your great work on the bioinformatic analyses, I believe you will have an interesting PhD project in the near future.

I also thank to my family members in China, especially my parents, without your support, I could never be here. 愿你们身体安康，笑口常开. I also thanks to my previous bachelor and master supervisors, they encouraged me to pursue the chance to study abroad.

Thanks to China scholarship council to offer me living fee to study abroad for four years. In the end, thanks to myself for getting the opportunity and finishing this PhD project.

About the author

Shuang Song was born on April 25, 1991, in Henan province, China. She grew up in a small village producing wheat and corn every year and sometimes she went to field to help her parents harvest corns in Autumn. Since there, she was associated with seeds. She chose to learn horticulture in her bachelor study in Henan University of Science and Technology, Luoyang, in 2010. After four years, she decided to continue in science. She became to be a master student and studied in Nanjing Agricultural University, Nanjing, in 2014. In the beginning of the master study, she knew she wanted to study abroad to be a PhD, so she tried her best to apply for scholarship from China Scholarship Council. In the meantime, she contacted with Leónie Bentsink. She managed to be a PhD candidate in Plant Physiology, Wageningen University, the Netherlands, in 2016. She did the research about transcriptional regulation of seed dormancy and seed longevity, the work performed during her PhD is described in this thesis.

Education Statement of the Graduate School

Experimental Plant Sciences

The Graduate School EXPERIMENTAL PLANT SCIENCES

Issued to: **Shuang Song**
Date: **23 March 2021**
Group: **Plant Physiology**
University: **Wageningen University & Research**

1) Start-Up Phase	*date*	*cp*
► **First presentation of your project**		
Functional analyses of the *DELAY OF GERMINATION 6*	12 Jun 2017	1.5
► **Writing or rewriting a project proposal**		
Functional analyses of the *DELAY OF GERMINATION 6*	05 May 2017	3.0
► **MSc courses**		
Subtotal Start-Up Phase		4.5

2) Scientific Exposure	*date*	*cp*
► **EPS PhD student days**		
EPS PhD student days 'Get2Gether', Soest, NL	09-10 Feb 2017	0.6
EPS PhD student days 'Get2Gether', Soest, NL	11-12 Feb 2019	0.6
► **EPS theme symposia**		
EPS Theme 3 Symposium 'Metabolism and Adaptation', Wageningen University	14 Mar 2017	0.3
EPS Theme 1 Symposium 'Developmental Biology of Plants', Wageningen University	30 Jan 2018	0.3
EPS Theme 3 Symposium 'Metabolism and Adaptation', Wageningen University	13 Mar 2018	0.3
EPS Theme 3 Symposium 'Metabolism and Adaptation', Radboud University	21 Oct 2019	0.3
EPS Theme 4 Symposium 'Genome Biology', Wageningen University	13 Dec 2019	0.3
EPS Theme 3 Symposium 'Metabolism and Adaptation', online	30 Oct 2020	0.2
► **Lunteren Days and other national platforms**		
Annual Meeting 'Experimental Plant Sciences', Lunteren, NL	10-11 Apr 2017	0.6
Annual Meeting 'Experimental Plant Sciences', Lunteren, NL	09-10 Apr 2018	0.6
Annual Meeting 'Experimental Plant Sciences', Lunteren, NL	08-09 Apr 2019	0.6
► **Seminars (series), workshops and symposia**		
5th Dutch Seed Symposium, Wageningen, NL	04-05 Oct 2016	0.6
STW committee meeting: How does nature regulate seed life span?	08 Nov 2016	0.2
Seminar of Dr. Sanjay Kapoor 'Regulators of Reproductive Development in Rice', Wageningen University	29 Aug 2017	0.1
6th Dutch Seed Symposium, Wageningen, NL	03 Oct 2017	0.3
Seminar of Frantisek Marec 'From sex chromosomes to sex determination in Lepidoptera', Wageningen University	25 Oct 2017	0.1
7th Dutch Seed Symposium, Wageningen, NL	02 Oct 2018	0.3
Seminar of Prof. Wolf Frommer 'Logistics: Allocation of carbon and energy for yield and pathogen resistance', Wageningen University	17 Jan 2019	0.1
8th Dutch Seed Symposium, Wageningen, NL	01 Oct 2019	0.3
Seminar of Prof. Julia Bailey 'A mixed message: Making, holding, decoding and destroying mRNAs for stress resilience', Wageningen University	23 Jan 2020	0.1
Flying seminar 'Plant growth and development'	20 Nov 2020	0.1
► **Seminar plus**		
► **International symposia and congresses**		
10th European Plant Science Retreat (EPSR), Utrecht, NL	03-06 Jul 2018	1.0
6th Workshop Molecular Aspects of Seed Dormancy and Germination, Volendam, NL	22-25 Sep 2019	0.9
► **Presentations**		
Talk: Functional analyses of the *DELAY OF GERMINATION* 6, 6th Workshop Molecular Aspects of Seed Dormancy and Germination, Volendam, NL	22-25 Sep 2019	1.0
Talk: Functional analyses of the *DELAY OF GERMINATION 6*, 8th Dutch Seed Symposium, Wageningen, NL	01 Oct 2019	1.0
Talk: Functional analyses of the *DELAY OF GERMINATION 6*, Annual EPS Theme 3 Symposium 'Metabolism and Adaptation', Radboud University, NL	21 Oct 2019	1.0
Poster: Functional analyses of the *DELAY OF GERMINATION 6*, Lunteren meetings, NL	2018-2019	1.0
► **3rd year interview**		
► **Excursions**		
Syngenta	23 Jan 2020	0.3
Subtotal Scientific Exposure		13.1

3) In-Depth Studies	*date*	*cp*
► **Advanced scientific courses & workshops**		
14th International Master Class on Seed Technology, Wageningen, NL	15-19 Apr 2018	1.2
Transcription Factors and Transcriptional Regulation, Wageningen, NL	10-12 Dec 2018	1.0
► **Journal club**		
Literature discussion in Plant Physiology	2016-2019	3.0
► **Individual research training**		
Subtotal In-Depth Studies		5.2

CONTINUED ON NEXT PAGE

4) Personal Development	date	cp
► **General skill training courses**		
Data Management Planning, Wageningen, NL	06 Feb 2017	0.4
Project and Time Management, Wageningen, NL	07 Sep - 26 Oct 2017	1.5
WGS PhD Workshop Carousel, Wageningen, NL	25 May 2018	0.3
Scientific Writing, Wageningen, NL	18 Sep - 06 Nov 2018	1.8
► **Organisation of meetings, PhD courses or outreach activities**		
Organisation of PhD symposium 'Science with impact'	01 Mar - 25 Oct 2019	1.5
► **Membership of EPS PhD Council**		
Subtotal Personal Development		5.5

5) Teaching & Supervision Duties	date	cp
► **Courses**		
Plant cell and tissue culture	18 Mar - 26 Apr 2019	3.0
► **Supervision of BSc/MSc students**		
Supervision of BSc Iona van den Berg	18 Mar - 05 Jul 2019	
Supervision of MSc Pallavi Shakya	2020	
Supervision of MSc Annika Liefferink	2020	3.0
Supervision of BSc Lot Lubberts	06 Jan - 13 Mar 2020	
Supervision of MSc Ao Jiao	02 Mar -18 Aug 2020	
Supervision of MSc Max Irena Theodorus	2021	
Subtotal Teaching & Supervision Duties		6.0

TOTAL NUMBER OF CREDIT POINTS*	**34.3**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS with a minimum total of 30 ECTS credits.

** A credit represents a normative study load of 28 hours of study.*

This research was performed at the Laboratory of Plant Physiology, Wageningen University & Research, financially supported by Dutch Technology Foundation STW TTW (STW project 12951), which is part of the Netherlands Organisation for Scientific Research (NWO) and the China Scholarship Council (CSC).

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design: Shuang Song

Printed by: proefschriftmaken.nl