Gene regulation in seeds - insights into translational dynamics

Bing Bai

Printing: GVO drukker & vormgevers B.V. | Ponsen & Looijen

Layout: Bing Bai
Cover design: Yinyin Liu

ISBN: 978-90-393-6513-7

The research described in this thesis was performed in the Molecular Plant Physiology group, Department of Biology, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH, Utrecht, and Laboratory of Plant Physiology, Department of Plant Sciences, Wageningen University, The Netherlands. The presented work was supported by the Netherlands Organization for Scientific Research (NWO)

Gene regulation in seeds – insights into translational dynamics

De regulatie van gen expressie in zaden –
inzicht in de dynamiek van de translatie

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G. J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 30 maart 2016

des middags te 12.45 uur

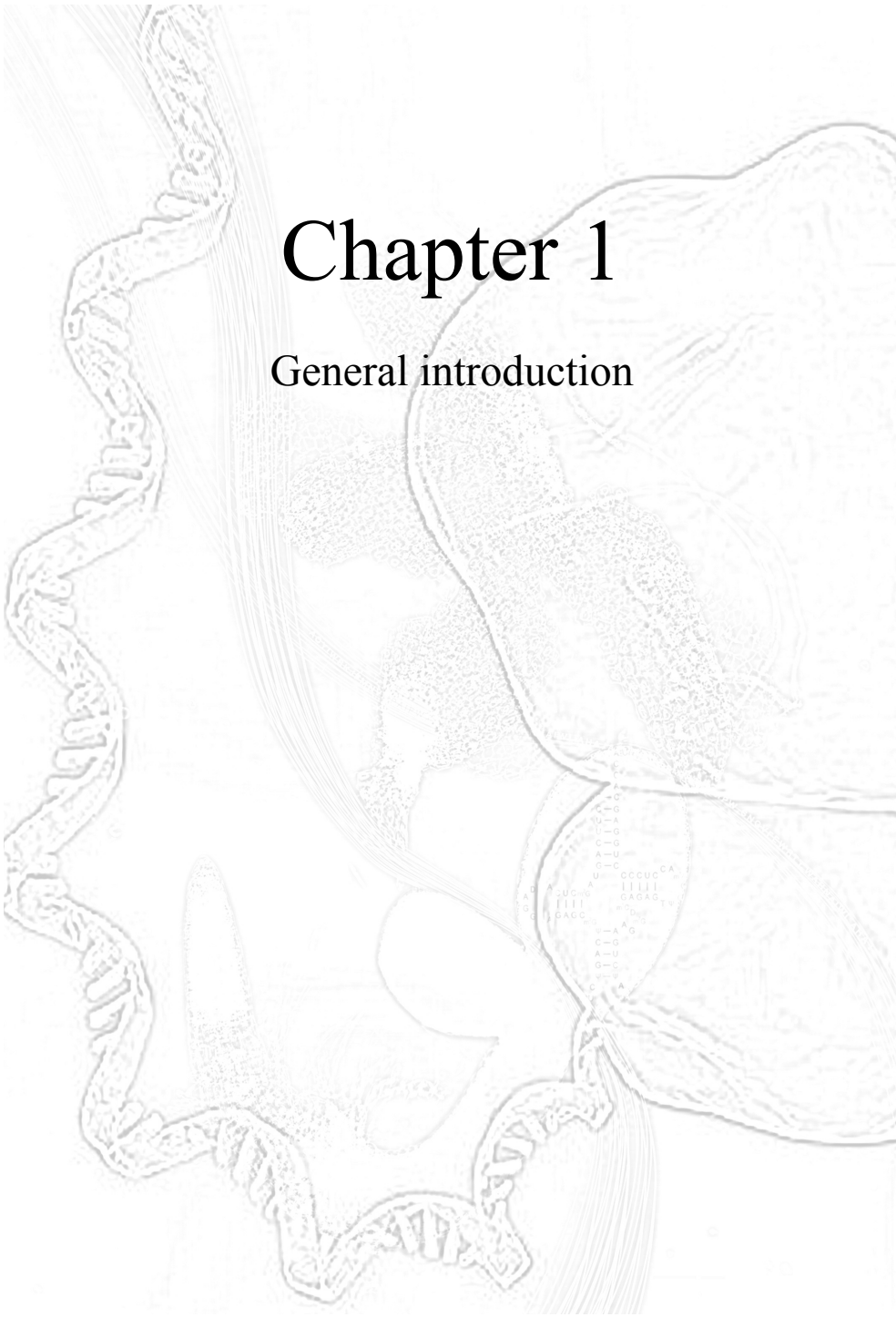
door

Bing Bai
geboren op 22 januari 1986 te Zhengzhou, China

Promotors:
Prof. dr. J.C.M. Smeekens
Prof. dr. Ir. H.J. Bouwmeester

Copromotors:
Dr. Ing. L. Bentsink
Dr. S.J. Hanson

Table of contents



Chapter 1

General introduction

The majority of the world population relies on seeds as staple food. Seeds are unique biological structures that allow the plant to survive unfavorable conditions and to spread their genetic material in time and space. Favorable seed traits are high nutritional value, low dormancy, high longevity, high seed vigor and robust seedling establishment. Thus, crop improvement strategies to improve these seed traits are of great agronomical importance. In this thesis, I studied processes related to these traits at both the transcriptional and translational level to further understand the regulation of seed germination, dormancy and longevity and to identify the underlying genes. In the introduction of this thesis, I focus on knowledge obtained using the model plant *Arabidopsis thaliana*.

1. Seed traits and genes

1.1 Seed dormancy

Seed dormancy is among the most important seed traits for agronomy and is defined as the failure of an intact, viable seed to complete germination under favorable conditions (Bewley, 1997). Seed dormancy is regulated at different developmental stages and controlled by distinct molecular pathways in interaction with environmental factors (Chen et al., 2014; Ding et al., 2014b; He et al., 2014). How these environmental signals are integrated into the seed developmental program is not fully understood.

Numerous mutants have been identified that are affected in seed dormancy. Among these are mutants in genes that play a role during seed maturation i.e. *ABA-INSENSITIVE3* (*ABI3*), *FUSCA3* (*FUS3*) and *LEAFY COTYLEDON* (*LEC1* and *LEC2*). These mutants are perturbed in seed dormancy as well as other processes such as accumulation of storage proteins and the acquisition of desiccation tolerance (Meinke, 1995; Parcy et al., 1997).

Mutants altered in phytohormone biosynthesis and signaling e.g. abscisic acid (ABA), gibberellin (GA), ethylene and auxin often have altered seed dormancy, as do mutants impaired in the maturation program. ABA and auxin enhance seed dormancy while GA and ethylene break seed dormancy and facilitate seed germination. ABA deficient mutants (*aba1* to *aba4*) (Koornneef et al., 1982; Karssen et al., 1983;

Leon-Kloosterziel et al., 1996a; Barrero et al., 2005; North et al., 2007) are selected by germination in the presence of paclobutrazol and tetcyclacis, which inhibit GA biosynthesis and thus allow seeds to germinate only when ABA synthesis is impaired (Leon-Kloosterziel et al., 1996). Mutants impaired in ABA signaling (*abi1 to abi5*) are selected in germination experiments in the presence of exogenously added ABA. Seeds of *abi1*, *abi2 and abi3* display a dramatic reduction in seed dormancy, similar to ABA deficient mutants (Koornneef et al., 1984; Finkelstein, 1994; Leung et al., 1997; Gosti et al., 1999), while no significant dormancy phenotypes are observed in the *abi4* and *abi5* mutants (Finkelstein, 1994). In contrast, *enhanced response to aba1* (*era1*) seeds are ABA hypersensitive and confer enhanced seed dormancy (Cutler et al., 1996).

Brassinosteroids (BR) and the more recently discovered strigolactones have also been implicated in regulating seed dormancy. BR synthesis and signaling genes are also involved in dormancy regulation (Steber and McCourt, 2001). *BRASSINOSTEROID SIGNALING KINASE* and *BRASSINOSTEROID INSENSITIVE 2* are differentially expressed in imbibed dormant seed compared to imbibed after-ripened seed (Chitnis et al., 2014). The activation of the BR synthesis gene At*DWF4* overcomes the ABA-induced inhibition of germination (Divi and Krishna, 2010). Strigolactones are originally identified as stimulant for parasitic seed germination (Cook et al., 1972). The role of strigolactone as germination stimulant is assumed to be conserved in Arabidopsis by its modulation of ABA and GA through *MORE AXILLARY GROWTH2* (*MAX2*) (Toh et al., 2012). *KARRIKIN INSENSITIVE1* (*KAI1*) encodes a karrikin receptor that is known to interact with *MAX2,* both the *kai1* and the *max2* mutants show increased seed dormancy (Nelson et al., 2011).

Hormones interact in the regulation of dormancy. Auxin application induces seed dormancy by enhancing *ABI3* expression. Furthermore the auxin biosynthesis YUCCA flavin monooxygenases mutant *yuc1yuc6* has reduced seed dormancy while endogenous addition of auxin enhanced the dormancy level. Auxin signaling is essential for the ABA inhibition of seed germination, seen that the reduced seed dormancy of the auxin response factor mutant *arf10arf16* cannot be overcome by

exogenous ABA application (Liu et al., 2013c). Another example is the Homeodomain containing protein mutant *GermoStatin Resistance locus 1*, which displays resistance to auxin mediated seed germination inhibition and reduced seed dormancy (Ye et al., 2015).

GA and ethylene antagonize ABA in regulating seed dormancy (Corbineau et al., 2014). The GA biosynthesis mutants *ga1, ga2* and *ga3* fail to germinate unless exogenous GA is applied. Seeds of ethylene insensitive mutants *ethylene resistant1* and *ethylene insensitive2* display enhanced primary dormancy. The ethylene hypersensitive mutant *constitutive triple responses1* has a slightly enhanced rate of germination (Bleecker et al., 1988; Beaudoin et al., 2000; Chiwocha et al., 2005; Subbiah and Reddy, 2010).

In Arabidopsis ecotypes substantial natural variation for seed dormancy is present (Alonso-Blanco et al., 2003; Bentsink et al., 2010). Quantitative Trait Loci (QTLs) affecting seed dormancy have been identified and were called *DELAY OF GERMINATION* (*DOG*). The effect of these QTLs have been confirmed using near isogenic lines (NILs) (Alonso-Blanco et al., 2003). One of the most dominant QTL, *DOG1* has been cloned (Bentsink et al., 2006) and the level of the DOG1 protein was proposed to function as a 'timer' for seed dormancy maintenance (Nakabayashi et al., 2005; Bentsink et al., 2010). These identified mutants and QTLs provide rich resources to investigate the gene regulatory networks underpinning seed dormancy and germination.

1.2 Seed germination

Seed germination is initiated from a quiescent dry state. Water uptake (imbibition) sequentially activates gene expression and metabolism to support reserve mobilization and seed germination. This sequential time course is closely associated with changes in physiological states, including seed hydration, testa rupture, radicle protrusion and root elongation followed by the greening and opening of the cotyledons (Nonogaki et al., 2010; Bewley et al., 2013). During early seed imbibition, germinating seeds can still be dried after applying a mild osmotic stress treatment with polyethylene

glycol (PEG). This ability is strictly dependent on the developmental stage of the germinated seeds and could in nature serve as a survival mechanism to cope with sudden dryness during early germination. At later stages, when root hair formation has started, seeds lose this capacity and become desiccation sensitive (Buitink et al., 2003; Maia et al., 2011; Costa et al., 2015; Dekkers et al., 2015a).

1.3 Seed storage – longevity

Seed longevity is another important seed trait as it determines how long seeds can be stored. Seed dormancy is overcome during dry storage, but after that seeds gradually deteriorate and show reduced germinability and eventually die. This deterioration process has been intensively studied, and several genes have been identified that are important for the process.

Reactive oxygen species (ROS) play a role in seed longevity. ROS is generated by oxygen and the most common ROS species are superoxide ($O_2^{-\cdot}$), hydrogen peroxide (H_2O_2), and hydroxyl radical ($HO^{\cdot}$). These highly reactive molecules trigger signal transduction networks to regulate the cellular redox homeostasis (Mittler et al., 2011). An imbalanced cellular ROS level could lead to oxidative stress, accelerating the oxidation of cellular components such as DNA, RNA and proteins resulting in the loss of seed viability. RNAs, especially poly(A)-mRNA, are rather fragile cellular components that are susceptible to damage (Kong and Lin, 2010). Oxidative modification of mRNA results in disturbance of the translational process and impairment of protein synthesis, which can cause cell deterioration or even cell death (Tanaka et al., 2007; Kong and Lin, 2010). To cope with the oxidative stress, the ROS scavenger system is extremely important. Superoxide dismutase (SOD), catalase (CAT) and ascorbate peroxidase (APX) are ubiquitous enzymes in the cell for redox balancing. Protein carbonylation triggers protein degradation mediated by the 20S ubiquitin-proteosome system (UPS) while antioxidants such as thioredoxin, peroxiredoxin, glutaredoxin or methionine sulfoxide reductase could buffer and reverse the oxidation of specific amino acids such as methionine and cysteine that could result in altered conformation and activity for many proteins (Wood et al.,

2003; Davies, 2005; Dos Santos et al., 2007).

Seed longevity positively correlates with antioxidant levels (Sattler et al., 2004; Rajjou and Debeaujon, 2008; Bailly and Kranner, 2011) and mutants involved in antioxidant systems have an altered longevity. Tocopherols (vitamin E) are lipophilic antioxidants abundant in seeds. Vitamin E deficient mutants (*vte1* and *vte2*) have a significant reduced seed longevity and increased oxidative damages such as lipid peroxidation (Sattler et al., 2004). DNA ligase provides initial repair of DNA damage in the embryo during dry storage. The importance of DNA repair for seed longevity is shown by the hypersensitivity to controlled ageing and abiotic stress of the DNA ligase mutants *lig4* and *lig4lig6* (Waterworth et al., 2010).

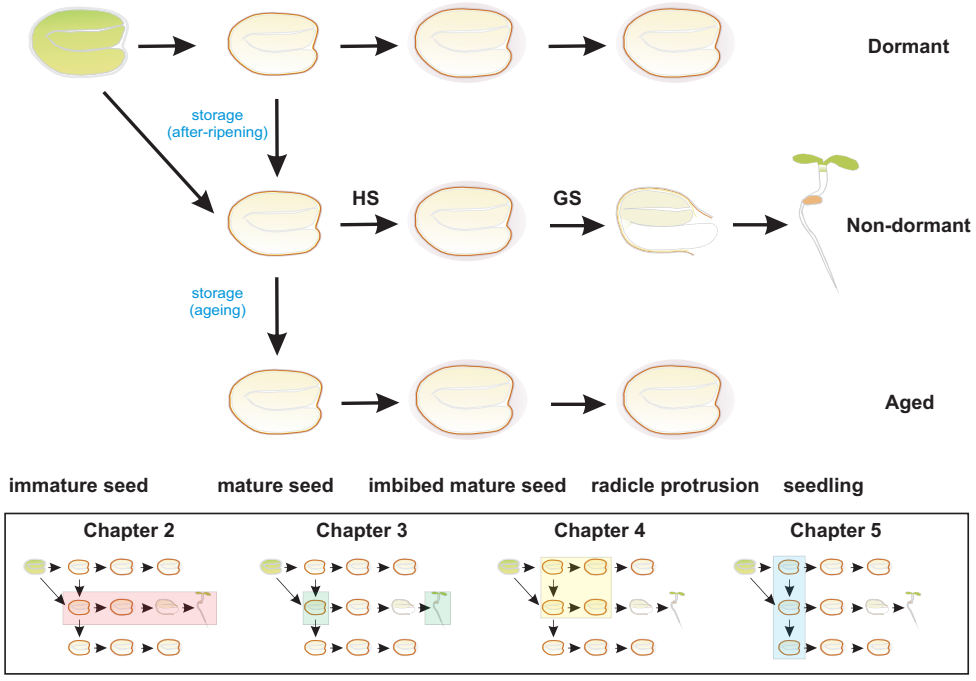


Figure 1. Schematic overview of seed stages. Seed dormancy is induced during seed maturation. Dormancy levels in Arabidopsis depend on the ecotype. Dormant seed cannot germinate even under optimal conditions. Following dry storage (after-ripening), seeds release dormancy. Non-dormant seeds can germinate and establish seedlings. Extended storage could lead to seed ageing. Seed viability decreases and thus cannot accomplish germination under optimal condition. The different colours below highlight the seed related process investigated in each Chapter. The figure is modified based on (Bentsink and Koornneef, 2008).

In all, seed dormancy and longevity provide two distinct but connected mechanisms for seeds to protect themselves against adverse environmental cues and thus safeguard the success for propagation through seed germination and seedling establishment (Figure 1).

2. Regulation of gene expression

Gene expression is the process by which information from genes is transferred to the synthesis of a functional gene product such as a protein or functional RNA. This work flow applies for all life forms including eukaryotes, prokaryotes and viruses.

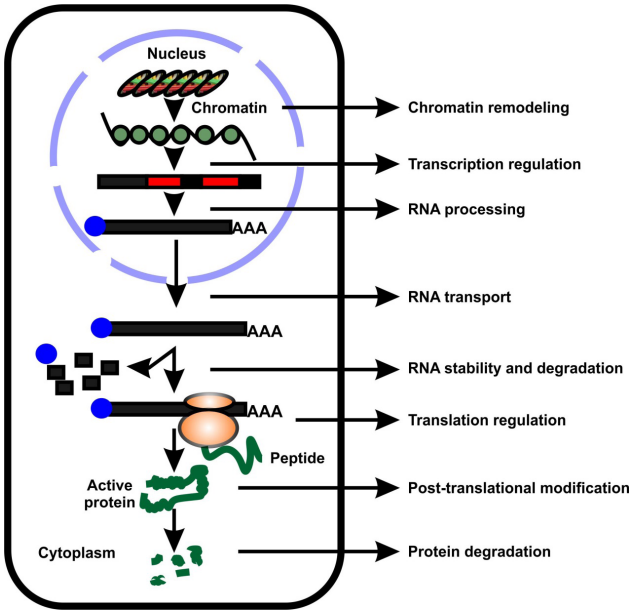


Figure 2. The overview of steps in the DNA-mRNA-protein pathway that gene expression can be regulated. Chromatin remodelling modifies the chromatin architecture which facilitates the access of transcriptional machineries to the condensed genomic DNA to control gene expression. With the increase accessibility of chromatin, transcription from DNA to RNA occurs mediated by the recognition of transcriptional machineries to the gene that need to be transcribed. This process is catalysed by RNA polymerase. Following transcription, primary RNA is converted into mature RNA through a series of processing events including 5' capping, 3' cleavage and polyadenylation, and alternative splicing to remove the introns. Mature RNAs are transported to cytosol through nuclear pores for protein translation in the cytosol. Translational regulation affects the ability for the ribosome to synthesize protein bases on a specific mRNA. After peptide synthesis, protein modification is essential for determining the protein activity. Protein activity can be reduced by modifications or by protein degradation.

Dissecting the regulation of gene expression is a central task in biology. Gene expression can be regulated at different levels including chromatin remodeling, transcriptional regulation, post-transcriptional regulation, translational regulation and post-translational regulation (Figure 2).

2.1 Regulation of RNA levels

2.1.1 mRNA synthesis and processing

Transcription is the first step of gene expression in which a DNA sequence is copied into RNA by three different RNA polymerases (Pol; Pol I, Pol II and Pol III) each of which has specific targets and activities. Pol I transcribes 18S, 5.8S and 28S ribosomal RNA while Pol II catalyzes the transcription of precursor mRNA, and the majority of small nuclear RNA (snRNA) and microRNA. In contrast, Pol III transcribes ribosomal 5S rRNA, tRNA and other small RNAs. The specificity of different polymerases provides independent mechanisms for transcriptional regulation. Transcriptional regulation is the most intensively studied gene expression regulatory mechanism.

In general, transcriptional regulation is controlled at three levels. First, transcription is controlled by the accessibility of the DNA for the RNA polymerases. This is affected by factors such as histone remodeling enzymes, transcription factors, activators and repressors. These factors bind specific DNA sequence elements to activate/repress transcription (Luo and Dean, 1999; Sassone-Corsi, 2002; Guo et al., 2015; Matharu and Ahituv, 2015). The second level is the rate of RNA elongation. This involves another set of factors to facilitate the progression of transcription from initiation to termination (Borukhov et al., 1992; Kim et al., 2007; Ui et al., 2015). Factors have been identified that control elongation and termination (Aranda and Proudfoot, 2001; Richard and Manley, 2009). These three levels work in concert to integrate signals for transcriptional adjustment in the cell. The activity of RNA polymerase is controlled by the phosphorylation status of the enzyme thereby adjusting the binding activity of various factors that together with RNA polymerase regulate the transcriptional efficiency (Komarnitsky et al., 2000; Phatnani and Greenleaf, 2006).

The translation of protein coding genes is only possible after the transcribed mRNA is transported to the cytosol. Post-transcriptional modification is closely linked to translation efficiency and represents a fascinating mechanism affecting plant development. Pre-mRNAs are subjected to a series of post-transcriptional events including RNA processing, editing and transportation. RNA modifications such as 5' capping, 3' polyadenylation and RNA splicing have important roles for mRNA stabilization and thus facilitate long-distance transport and translation. These post-transcriptional events dramatically enhance the genetic diversity and provide flexibility to adjust to the changing environment or developmental programs (Figure 2).

2.1.2 mRNA turnover

The steady state level of mRNAs is determined by the rates of both synthesis and degradation. Whether mRNAs are translated or degraded depends on the environmental and developmental context. This decision is made by mRNA surveillance mechanisms which safeguards the fidelity and quality of mRNA. To cope with the developmental and environmental changes, plants have to produce new mRNAs with different functions and such mRNAs are subject to decay. Knowledge on essential components in the decay pathways mostly comes from yeast and mammals. In plants, mRNA decay is an emerging research field. In general, there are two types of decay pathways, either 3'-5' exosome mediated decay pathways or 5'-3' exoribonuclease (XRN) mediated decay pathways. Both pathways are preceded by poly(A) tail shortening, known as deadenylation. Deadenylation involves three types of deadenylase complexes (Carbon Catabolite Repression 4 (Ccr4-Not), poly(A) ribonucleases (PARNs) and poly(A) binding protein dependent nuclease (PAN)). In yeast, the Ccr4-Not complex together with the Ccr4-Associated Factor 1 (Caf1) complex (Ccr4-Caf1) form the multi-subunit deadenylase complex. It contains two nucleases Ccr4p and Caf1p as well as several accessory proteins, Not1−Not5p, Caf4, Caf16, Caf40 and Caf130p (Tucker et al., 2001; Denis and Chen, 2003; Parker and Song, 2004). Not1 serves as the scaffold for the complex (Ito et al.,

2011), while Not4 is a chaperone that has E3 ubiquitin ligase activity in association with ribosomes (Panasenko et al., 2006). Selective deletion of components of the Ccr4-Not complex leads to different effects on the transcription profile, indicating the distinct function of different subunits in the complex (Cui et al., 2008; Azzouz et al., 2009). Ccr4-Caf1 is associated with NANOS2, a zinc-finger protein, to mediate mRNA decay (Suzuki et al., 2010; Suzuki et al., 2012). PARNs is a well characterized deadenylase in higher eukaryotes. It was first identified in mammals and is absent from yeast. Knock-out alleles of the Arabidopsis homolog AtPARN cause lethality prior to seed germination, indicating it is essential for embryo development (Chiba et al., 2004). The deadenylation activity of the PAN system has not been characterized in Arabidopsis, however at least one subunit of the PAN complex is encoded by the genome (Reverdatto et al., 2004). Arabidopsis uses different deadenylation systems to target distinct mRNA species, which is possibly mediated by the interaction of these deadenylation complexes with sequence features of the transcripts (Dehlin et al., 2000; Copeland and Wormington, 2001; Lai et al., 2003).

2.1.3 Processing bodies and stress granules

Following mRNA transport to the cytosol, not all mRNAs are translated. Some of the mRNAs are stored in specific cytosolic locations for translation repression or mRNA degradation. These discrete cytoplasmic domains are referred to as mRNA processing bodies (P-bodies). The decay of mRNAs by both pathways as described above (2.1.2) takes place in P-bodies as apparent by the presence of mRNA processing enzymes such as exoribonuclease XRN1 and mRNA-decapping enzyme 2 (DCP2) (Sheth and Parker, 2003; Cougot et al., 2004). P-bodies play a role in mRNA surveillance that ensures the translation of error-free mRNAs. mRNAs that contain premature translation-termination codons (PTCs) are degraded through nonsense-mediated mRNA decay (NMD) (Fasken and Corbett, 2005; Lejeune and Maquat, 2005). P-bodies also process the small interfering RNAs (siRNAs) and microRNAs (miRNAs) needed to induce gene silencing. The required decay components for miRNA targets are localized in P-bodies including argonaute proteins, GW182, the

Ccr4–Caf1–Not deadenylase complex, the decapping DCP1–DCP2 complex and XRN1 (Liu et al., 2005; Pillai et al., 2005; Sen and Blau, 2005; Behm-Ansmant et al., 2006; Chu and Rana, 2006).

Non-translating mRNAs can form another cytoplasmic granule referred to as stress granule (SG). The distinction between SGs and P-bodies is that SGs contain translation initiation factors and 40S ribosomal subunits, as well as cytoplasmic poly(A)-binding protein-1 (PABPC1) and the Ras-GTPase-activating SH3-domain-binding protein (G3BP) while DCP1, DCP2 and GW182 are P-body specific components. However, SG and P-bodies are rather similar and association and fusion events were observed between them, suggesting the exchange of components between the two granules (Kedersha et al., 2005; Wilczynska et al., 2005). Due to the presence of translation initiation components in SGs, mRNAs sequestered from the polysome during plant cell stress may be first directed to SGs. These SG-localized mRNAs are either redirected to the polysome for translation or to P-bodies for degradation during stress recovery (Kedersha et al., 2005).

2.2. Translational regulation

2.2.1 The ribosome

The ribosome consists of proteins (ribosomal proteins) and RNA (ribosomal RNA) and catalyzes protein synthesis. Ribosomal RNA (rRNA) contributes to approximately 60% of the ribosome weight. Ribosome biogenesis begins with the transcription of ribosomal DNA (rDNA) into a polycistronic precursor of approximately 45S in plants and yeast by RNA polymerase I (Pol I) in the nucleus. In contrast, the ribosomal protein (r-protein) genes are transcribed by RNA polymerase II. R-proteins, non-ribosomal proteins involved in processing and small nucleolar RNAs (snoRNAs) interact with the rRNA polycistronic precursor forming a 90S pre-particle. The 5S rRNA is transcribed by RNA polymerase III (Pol III) which associates with the 35S pre-rRNA (Zhang et al., 2007). The processing of the rRNA polycistronic precursor includes a series of cleavage and modification reactions involving about 200 different synthesis factors for rRNA maturation and subunit assembly. Together

with r-proteins, this process results in the pre-40 small subunit (SSU) and pre-60S large subunit (LSU) which are then transported to cytoplasm for ribosome assembly. The maturation of rRNA generates 18S, 5S, 5.8S and 23S rRNAs and together with the r-proteins give rise to the 80S ribosome, consisting of the 40S SSU including 18S rRNA and about 31 small subunit proteins and the 60S LSU including 23S, 5,8S, 5S rRNA and about 47 large subunit proteins. Structural sites of the ribosome essential for translation include the Acceptor site (A site), Donor site (P-site) and Exit site (E-site) with the peptidyltransfer center (PIC) for peptide bond formation.

2.2.2 Translation

Protein translation can be generally divided in three phases namely the initiation, elongation and termination (Figure 3). The initiation phase is the most rate-limiting step for protein translation, involving ternary complex formation and binding of the ribosome SSU to the mRNA . This ternary complex consists of the initiator tRNA (Met-tRNAiMet), eukaryotic initiation factor 2 (eIF2) and guanosine 5'-triphosphate (GTP). Together with eIF1, eIF1A, eIF2, eIF3 and eIF5 and the 40S SSU, a 43S pre-initiation complex (PIC) is formed. Promoted by the unwinding of the secondary structure in 5' leader, the scanning of the 43S PIC is facilitated and a poly(A) binding protein (PABP) mediated circular mRNA conformation is achieved, which stimulates the primary translation initiation and reinitiating events. Once the start codon AUG is recognized, the 43 PIC stops scanning and a 48 PIC is formed after the integration of the initiator tRNA base paired with the mRNA start codon in the P site. Followed by fusion of the 60S LSU and driven by GTP hydrolysis, the 80S ribosome complex is formed and the translation elongation leads to peptide synthesis and following termination to release of the protein and recycling of the ribosome.

mRNA levels do not always reflect the rate of protein synthesis. Numerous reports have revealed the discordance between the RNA and protein levels (Murrell et al., 1995; Chen et al., 2002; Beltaifa et al., 2005; Velez-Bermudez and Schmidt, 2014; Swindell et al., 2015). After processing, mRNAs are eventually recruited by the translational machinery (ribosome) for protein translation. Translation, which

represents one of the most energy-consuming processes in the cell, is heavily regulated.

2.2.3 Global translation regulation

Protein translation responds globally to environmental stimuli and developmental programs as evaluated by the polysome association and dissociation processes. Abiotic stress such as dehydration, hypoxia, heat, sucrose starvation in general lead to polysome dissociation and concomitant reduction of translation (Kawaguchi et al., 2004; Nicolai et al., 2006; Yanguez et al., 2013; Gamm et al., 2014; Juntawong et al., 2014). Translation initiation is the hot spot for translational control and is conserved in eukaryotic organisms (Browning, 2004; Van Der Kelen et al., 2009). The phosphorylation status of initiation factors and poly(A) binding protein (PABP) are key to regulate translation. Initiation factors eIF2, eIF4A, eIF4B, eIF4E, and eIFiso4E (an isoform of eIF4E) are hyper phosphorylated during some stages of development and in response to stresses, such as heat shock and hypoxia (Webster et al., 1991; Gallie et al., 1997). The phosphorylation status of PABP, eIF4B, and eIF4G is important in regulating the assembly of the initiation complex and 5' and 3' mRNA end interactions (Gallie et al., 1997; Le et al., 1997; Le et al., 1998, 2000). The caseine kinase II (CK2) has an important role in the phosphorylation and assembly multifactor complex (MFC) containing eIF1, eIF1A, eIF3, eIF5, and the ternary complex, which facilitate their interaction (Dennis and Browning, 2009; Dennis et al., 2009). eIF2 is another phosphorylation target that affects translation initiation. Phosphorylation of eIF2 blocks the eIF2B-catalyzed exchange of GDP for GTP, resulting in a decreased translation rate (Mathews et al., 2007). eIF2 is phosphorylated by the Arabidopsis eIF2 kinase homologue GCN2 upon stress (Lageix et al., 2008). Polysome profile analysis of *gcn2* demonstrated that GCN2 is important in reducing global protein synthesis as a mechanism for coping with a specific stress. Besides CK2, plants have two other central kinases, Snf1-related kinase 1 (SnRK1) and the target of rapamycin (TOR) kinase that link translation during growth and development to carbon nutrient and energy status (Smeekens et al., 2010; Robaglia et al., 2012).

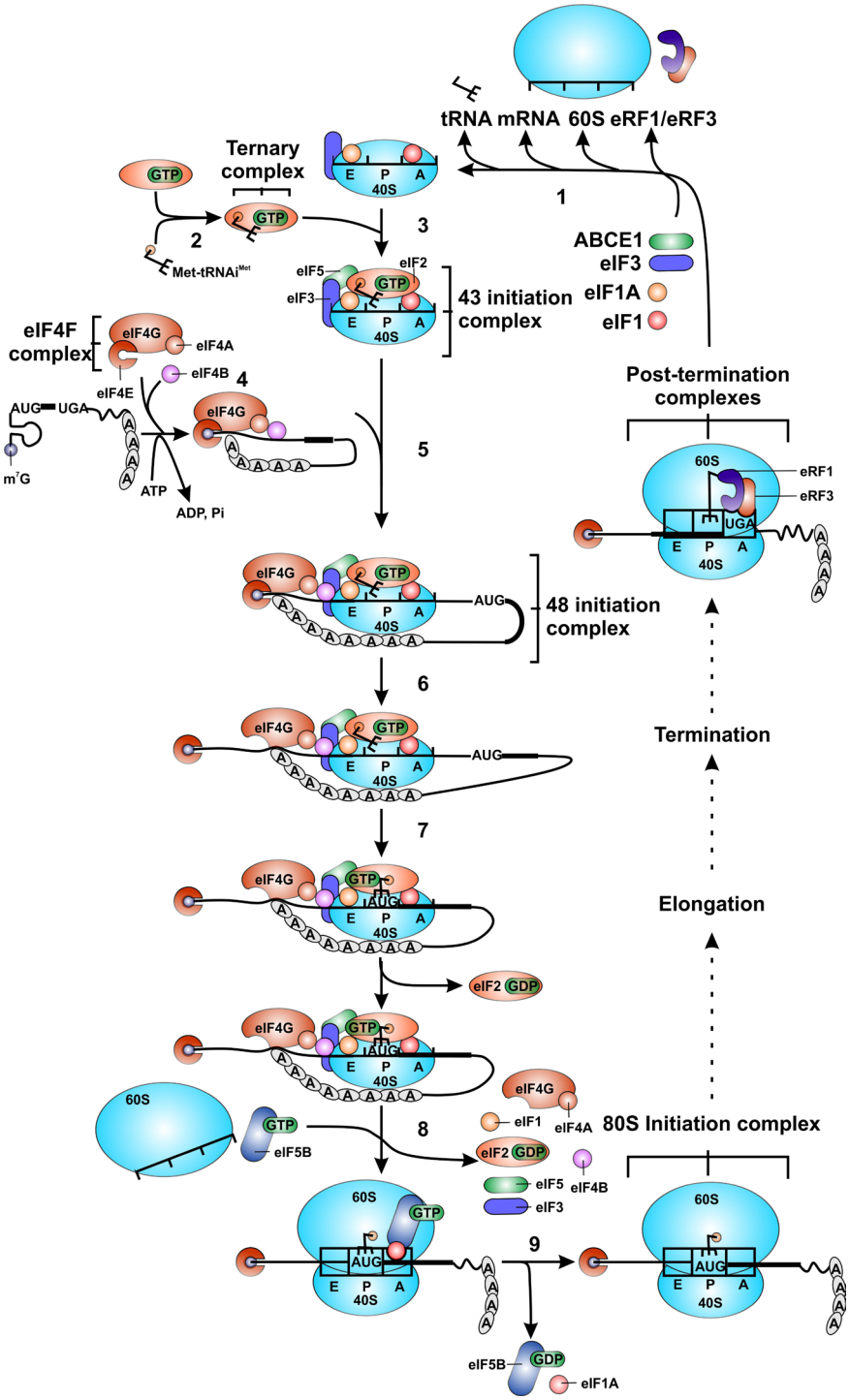

tRNA mRNA 60S eRF1/eRF3
1
Ternary complex
GTP
2
Met-tRNAi^Met
3
40S
ABCE1
eIF3
eIF1A
eIF1
eIF5
eIF2
eIF3
43 initiation complex
eIF4F complex
eIF4G
eIF4A
eIF4B
eIF4E
4
AUG
UGA
m7G
ATP
ADP, Pi
5
Post-termination complexes
60S
eRF1
eRF3
UGA
48 initiation complex
6
Termination
7
Elongation
eIF2 GDP
8
80S Initiation complex
eIF1
eIF5
eIF5B
9
GDP

Figure 3. Model of the canonical pathway of eukaryotic translation initiation. The canonical pathway of eukaryotic translation initiation is divided into nine stages (1–9). These stages follow the recycling of post-termination complexes (1) to yield separated 40S and 60S ribosomal subunits, and result in the formation of an 80S ribosomal initiation complex, in which Met-tRNAMeti is base paired with the initiation codon in the ribosomal P-site and which is competent to start the translation elongation stage. These stages are: eukaryotic initiation factor 2 (eIF2)–GTP–Met-tRNAMeti ternary complex formation (2); formation of a 43S preinitiation complex comprising a 40S subunit, eIF1, eIF1A, eIF3, eIF2–GTP–Met-tRNAMeti and probably eIF5 (3); mRNA activation, during which the mRNA cap-proximal region is unwound in an ATP-dependent manner by eIF4F with eIF4B (4); attachment of the 43S complex to this mRNA region (5); scanning of the 5' UTR in a 5' to 3' direction by 43S complexes (6); recognition of the initiation codon and 48S initiation complex formation, which switches the scanning complex to a 'closed' conformation and leads to displacement of eIF1 to allow eIF5-mediated hydrolysis of eIF2-bound GTP and Pi release (7); joining of 60S subunits to 48S complexes and concomitant displacement of eIF2–GDP and other factors (eIF1, eIF3, eIF4A, eIF4B and eIF5) mediated by eIF5B (8); and GTP hydrolysis by eIF5B and release of eIF1A and GDP-bound eIF5B from assembled elongation competent 80S ribosomes (9). Translation is a cyclical process, in which termination follows elongation and leads to recycling (1), which generates separated ribosomal subunits. The Figure is modified based on (Jackson et al., 2010) Figure 3. Model of the canonical pathway of eukaryotic translation initiation. The canonical pathway of eukaryotic translation initiation is divided into nine stages (1–9). These stages follow the recycling of post-termination complexes (1) to yield separated 40S and 60S ribosomal subunits, and result in the formation of an 80S ribosomal initiation complex, in which Met-tRNAMeti is base paired with the initiation codon in the ribosomal P-site and which is competent to start the translation elongation stage. These stages are: eukaryotic initiation factor 2 (eIF2)–GTP–Met-tRNAMeti ternary complex formation (2); formation of a 43S preinitiation complex comprising a 40S subunit, eIF1, eIF1A, eIF3, eIF2–GTP–Met-tRNAMeti and probably eIF5 (3); mRNA activation, during which the mRNA cap-proximal region is unwound in an ATP-dependent manner by eIF4F with eIF4B (4); attachment of the 43S complex to this mRNA region (5); scanning of the 5' UTR in a 5' to 3' direction by 43S complexes (6); recognition of the initiation codon and 48S initiation complex formation, which switches the scanning complex to a 'closed' conformation and leads to displacement of eIF1 to allow eIF5-mediated hydrolysis of eIF2-bound GTP and Pi release (7); joining of 60S subunits to 48S complexes and concomitant displacement of eIF2–GDP and other factors (eIF1, eIF3, eIF4A, eIF4B and eIF5) mediated by eIF5B (8); and GTP hydrolysis by eIF5B and release of eIF1A and GDP-bound eIF5B from assembled elongation competent 80S ribosomes (9). Translation is a cyclical process, in which termination follows elongation and leads to recycling (1), which generates separated ribosomal subunits. The Figure is modified based on (Jackson et al., 2010).

Both kinases could adjust global translation status by affecting ribosome biogenesis and activity (Baena-Gonzalez et al., 2007; Deprost et al., 2007; Ren et al., 2012). SnRK1 is activated by nutrient stress resulting in the inhibition of growth, while TOR is activated when metabolites are abundant and TOR stimulates translation and growth (Lastdrager et al., 2014), suggesting a different regulatory role for TOR and

SnRK1 in global translation and plant development.

2.3.4 Gene specific translational control

Defects in components of the translation machinery can affect translation in a gene specific manner. Studies on eukaryotic translation initiation (eIF) factors revealed that eIF-5A affects biological switches determining cell fate by differentially positioning its different isoforms in dividing cell or dying cells (Thompson et al., 2004; Feng et al., 2007). Elongation factor eEF-1B beta 1, a guanine nucleotide exchange factor playing a role in translation elongation and has also been reported recently as an important growth regulator of cell wall synthesis (Hossain et al., 2012).

Ribosomal proteins are associated with specific plant growth phenotypes. Ribosomal proteins do not only functional as scaffold for the ribosome, and studies have revealed their multi-functional role in regulating processes such as plant pathogen interaction (Yang et al., 2009). Ribosomal protein mutants in Arabidopsis often show a pointed leave edge which may indicate their convergent effect on plant growth and the possible cross-talk with auxin, one of the essential phytohormones. However, different ribosomal proteins may affect translation and plant development in a specific manner. Ribosomal protein S18 (RPS18) is expressed in the meristematic tissue and affects the meristem specifically when it is mutated (Van Lijsebettens et al., 1994). Tobacco plants defective in Ribosomal Protein L3 (RPL3) have a reduced cell number/size ratio indicating that RPL3 positively affects cell division. The decrease of pre-rRNA in gene silenced plants established that RPL3 is involved in ribosome biogenesis (Popescu and Tumer, 2004). Ribosomal Protein L10 (RPL10) was shown to affect plant growth in a paralog specific way (Falcone Ferreyra et al., 2010). The insertion in RPL10A caused lethality whereas a mutation in RPL10B only caused abnormal growth. Different paralogs also showed a different abiotic stress sensitivity, which reveals a functional non-redundant role of ribosome protein in plant development and stress response.

Another aspect of gene specific translational regulation lies in sequence features of the regulated mRNA. Mature mRNA consists of a 5’ cap, a 5’ untranslated region

(5'UTR), a start codon, a coding region, a stop codon, a 3' UTR and a poly(A) tail. Transcript lengths, GC content, initiation codon context, codon usage frequency, RNA binding protein (RBP) binding sites, and mRNA structure and upstream open reading frames (uORFs) affect ribosome behavior on the mRNA and thus the translation efficiency in a transcript specific manner (Crick, 1966; Hershberg and Petrov, 2008, 2009; Qu et al., 2011a; Valleriani et al., 2011; Liu et al., 2012b). A 7-methylguanylate cap (m^7G) located at the 5' end of the mRNA consists of the guanine nucleotide connected to the mRNA via the unusual 5' to 5' triphosphate linkage. The special nucleotide structure provides the 5' cap with unique functions such as regulation of nuclear export, resistant to degradation, promotion of translation and promotion of 5' proximal intron excision. The 5' cap is necessary for translation since eIF-4E, a translation initiation factor, is competing with the decapping enzyme to bind the cap thus to avoid mRNA degradation from the 5' end. Codon bias refers to the different frequency of the occurrence of synonymous codons (referred to as codon degeneracy). The preference of each codon reflects a balance between mutational bias and natural selection for translation optimization. A model has been raised for the co-evolution between the expression of a tRNA and codon usage although the causative relationship between the two is still not clear (Behura and Severson, 2011). GC content is defined as the percentage of guanine (G) and cytosine (C) in a given sequence. The three hydrogen bonds in the GC pair compared with two hydrogen bonds in the AU pair bring increased thermal stability and heat tolerance for RNAs with higher GC content. Evidence has shown that the GC content is proportional to the length of the coding region and the stop codon has the bias towards the A and T nucleotides. Therefore shorter mRNAs tend to have a higher AT bias (Wuitschick and Karrer, 1999).

Recently, mRNA structure, which is defined by specific base pairing interactions encoded within the primary sequence, has been linked to a role in gene expression regulation (Mauger and Weeks, 2010; Zheng et al., 2010; Li et al., 2012; Liu et al., 2012a; Burrill et al., 2013; Ding et al., 2014a; Rouskin et al., 2014; Gosai et al., 2015). RNA structure affects pre-mRNA splicing by masking or exposing the splicing site

(Raker et al., 2009; Warf and Berglund, 2010). Also, it provides extensive interaction changes between the ribosome and RNA binding proteins through which mRNA translation and metabolism can be dramatically affected (Klasens et al., 1998; Zarudnaya et al., 2003; Li et al., 2012).

2.3.5 RNA binding proteins

RNA binding proteins (RBPs) are emerging as important translational regulators involved in different developmental processes. RBPs regulate mRNA metabolism and translation by specifically recognizing mRNA sequence motifs or mRNA recognition elements present in their target mRNAs, or by functioning as a scaffold and assembly platform for recruiting proteins to act synergistically (Cech and Steitz, 2014). RBPs are structurally diverse and include many distinct classes based on their binding motif structure such as the RNA recognition motif (RRM), K homology (KH) domain, DEAD motif, double-stranded RNA-binding motif (DSRM) or zinc-finger domain (Gerstberger et al., 2014), which may play multi-functional roles in diverse types of post-transcriptional regulation (Ray et al., 2013).

The Arabidopsis genome encodes more than 200 different RBPs, most of which are plant specific and are therefore likely to perform plant-specific functions (Lorkovic, 2009). DEAD box proteins are largest helicase family characterized by the presence of an Asp-Glu-Ala-Asp (DEAD) motif that uses ATP to bind or remodel RBPs (Linder and Jankowsky, 2011). PUF proteins, a novel family of RBPs named after the *Drosophila Pumilio (Pum)* protein which is evolutionarily highly conserved from yeast to humans and plants. PUF is characterized by a C-terminal RNA-binding domain, composed of eight tandem repeats (Spassov and Jurecic, 2003), which bind to specific recognition sequences in the UTRs of mRNAs and control their stability and translation. Increased mRNA structure negatively correlates with RBPs since the exposed RBP binding sites including motifs for alternative splicing (AS) and polyadenylation are limited. Furthermore, mRNAs important in certain processes such as cell death/apoptosis and postembryonic and organ development, response to desiccation, abscisic acid and cold, stress response, posttranslational modification

and mRNA processing, could have similar RBP binding motifs (Gosai et al., 2015), providing evidence on the existence of posttranscriptional operons in modulating cellular functions.

3. Gene regulation in seeds - Insight into translational dynamics

3.1 The role of the ribosome during seed development

The studies on ribosomes are pioneered since 1950s by Palade and colleagues (Palade, et al., 1955). The ribosome was described at that time as particulate component of small dimensions (100 to 150 Å) and high density in association with the membrane of the endoplasmic reticulum. The discovery of this cellular component and the characterization of its structure and functions were later awarded with Nobel Prizes in 1974 and 2009 respectively. The study of the ribosome in relation to seed mRNA storage started in the 1960s. These studies showed that mRNA translation during early imbibition is not depending on transcription but occurred from stored mRNAs and ribosomal components (Marcus and Feeley, 1964; Dure and Waters, 1965; Chen et al., 1968). The stored form of these mRNAs was later identified in association with informosomes, which nowadays are recognized as messenger ribonucleoprotein (mRNP). Sucrose and CsCl gradients were widely applied to investigate the dynamics of total and *de novo* synthesized RNPs (Woese, 1961; Marcus et al., 1966; Lawford, 1969; Rinaldi and Monroy, 1969; App et al., 1971; Fountain and Bewley, 1973). During wheat and chickpea ripening, there is a gradual disappearance of polysomes (complex of an mRNA molecule with two or more ribosomes) and an accumulation of free monosomes (singular ribosomes). This reflects a decrease of protein synthesis rate and accumulation of stored mRNA in form of monosomes during the transition to the ripe seed (Ajtkhozhin et al., 1976; Nair and Koundal, 1993), indicating that translation is finely regulated according to the seed developmental program.

3.2 Ribosome and protein synthesis during seed germination

Following harvest, seeds normally maintain variable levels of dormancy depending on the genotype. Thus seeds may not reach 100% germination although the seeds are completely matured. During a period of dry storage, called after-ripening, seeds lose dormancy after which they can achieve a higher germination rate. Extensive studies have addressed this process regarding the dynamics of ribosome and proteins synthesis (Marcus et al., 1966; Marre, 1967; App et al., 1971; Fountain and Bewley, 1973; Siwecka et al., 1973; Speigel and Marcus, 1975; Subramanian, 1978; Iskakov and Aitkhozhin, 1979). Seed germination is affected by stored mRNAs and proteins. The dynamic changes in these stored mRNA protein complexes during the germination process have been described (Ottoleng.S et al., 1973; Ajtkhozhin et al., 1976). Protein synthesis during seed germination was studied in peanut (Marcus and Feeley, 1964). The authors of this study concluded that the machinery for protein synthesis exists in the dry seeds, however the mRNA is limiting, which can be overcome by imbibition. Later experiments showed that polysome formation is important for mRNA translation, which distinguishes the imbibed from the dry seed (Marcus and Feeley, 1965). This process is independent from transcription and polyadenylation but prevented by translation inhibition (Speigel and Marcus, 1975). The studies on ribosome behavior during seed germination peaked in the 1970s. The reduced resistance of the ribosome to dissociation during seed germination was reported (App et al., 1971). This physiological process was presumed independent of energy and protein synthesis as well as RNase activation. In addition, the dissociation of the 80S ribosome might be the initial event in activating protein synthesis during seed germination. Contradictory results were reported in lettuce seed germination regarding the capability to produce polysomes. Mitchell and coworkers reported that polysome formation is inhibited when lettuce seeds are imbibed in darkness, which prevents seed germination (Mitchell and Villiers, 1972), however a similar experiment conducted by Fountain later showed that lettuce seeds imbibed in darkness could produce polysomes at a similar level as non-dormant seeds (Fountain and Bewley, 1973). The utilization of stored versus newly transcribed mRNAs can

be evidenced from the poly(A) content at early imbibition (Delseny et al., 1977). At later imbibition stages poly(A) content increases indicating that new mRNAs are synthesized. Consensus on ribosome behavior and the importance of stored mRNAs was later established (Subramanian, 1978; Aspart et al., 1984; Suzuki and Minamikawa, 1985). The stored mRNAs encode ribosomal proteins (Beltran-Pena et al., 1995), heat shock and LEA proteins (Almoguera and Jordano, 1992).

3.3 RNA binding protein in seed research

mRNA binding proteins are closely associated with mRNAs and potentially regulate translation. A large number of studies have revealed the regulatory role of DEAD box RNA helicase on plant development and environmental responses such as seed development and seedling growth (Kanai et al., 2013), glucose and ABA signaling during seed germination (Hsu et al., 2014), restoration of seed translation upon rehydration after dehydration (Nayak et al., 2013) and stress response (Gong et al., 2002; Gong et al., 2005; Barak et al., 2014; Khan et al., 2014). PUF proteins have recently been proposed to play a role in germination. Two genes, *APUM9* and *APUM11* showed dramatically increased transcript levels in the *reduced dormancy 5-1 (rdo5-1)* mutant background. Transgenic plants with enhanced or reduced expression levels of *APUM9* and *APUM11* showed a decrease and increase, respectively in seed dormancy levels (Xiang et al., 2014).

3.4 The relevance of mRNA decay in seeds

mRNA decay pathways determine the fate of mRNA and are important for seed germination. Transcripts encoding Ccr4-Not complex components have been detected associated with the polysome in the non-dormant seed during germination, indicating its distinct role during seed germination (Layat et al., 2014). The Ccr4-Not complex is also associated with the endosperm weakening process during seed germination (Voegele et al., 2011). In addition to Ccr4-Not complex, PARN and PAN are another active deadenylases. PARN is conserved deadenylase belonging to RNase D family of nucleases (Wilusz et al., 2001). Abscisic Acid (ABA) Hypersensitive

Germination 2 (AHG2) is an essential component of PARN. *ahg2* mutant seeds show delayed germination, hypersensitivity to ABA and pleiotropic phenotypes during embryo development (Chiba et al., 2004; Reverdatto et al., 2004; Nishimura et al., 2005). The deletion of AtPARN affects the deadenylation of a subset of embryonic transcripts including Cruciferin, Oleosin, LEC1 and LEC1-like genes, indicating the specificity of the AtPARN (Reverdatto et al., 2004).

3.5 New techniques applied in seed research

Transcriptomics and proteomics have been widely used for the investigation of total mRNA (Dekkers et al., 2013; Meimoun et al., 2014), protein changes (Chibani et al., 2006; Galland et al., 2014; Staszak and Pawlowski, 2014) and modifications (Oracz et al., 2007; Bazin et al., 2011; El-Maarouf-Bouteau et al., 2013) in related to seed dormancy and germination. The general conclusion drawn from these studies was that there are limited changes in total mRNA pool and extensive oxidation events at both RNA and protein level in dry seeds. This results in a different proteome in after-ripened seeds upon imbibition. Genome-wide studies of seed stored mRNAs and proteins have been conducted providing evidence on the epigenetic and genetic regulation of transcription in seed and the function of stored and neosynthesized mRNAs could be distinguished during seed imbibition (Rajjou et al., 2004; Nakabayashi et al., 2005). The more recent efforts have been on elucidating the functional role of mRNA recruited with polysome which represents the actively translated mRNA (Layat et al., 2014; Basbouss-Serhal et al., 2015).

Outline of the thesis

Seeds are a unique structure in the plant life cycle. The variation in timing of seed maturation, dispersion, and the establishment of seed dormancy and longevity, increases the chances of plant survival and enlarge the distance that plants could disperse in the natural habitat. As a structure that can be stored for a long period and thus travel through time and space for dispersion and propagation, a seed is referred to as a time capsule.

In this thesis, I use *Arabidopsis thaliana* seeds as a model system to investigate the translational dynamics during different seed developmental phases. This study is facilitated by the newly developed translatome profiling technique, in which ribosomes are fractionated based on their engagement in translation followed by the high throughput profiling of the mRNAs associated to these ribosomes.

In **Chapter 2**, the translational dynamics from a quiescent dry seeds to a seedling are described. The levels of polysome-associated mRNAs (translatome) are compared with the levels of the mRNAs in the total mRNA pool (transcriptome) in consecutive stages during seed germination. In **Chapter 3** I study seed stored mRNAs and their association with ribosomes (monosome/polysome). The function of these seed stored mRNAs is addressed regarding their role in seed storage and germination.

Fully matured seeds can only germinate after a period of dry storage called after-ripening. Is gene transcription or translation the driving force for the changes in germinability during dry storage? In **Chapter 4** we explore this question by investigating genes that are differentially transcribed and translated during the imbibition of NIL*DOG1* (a near isogenic line carrying an introgression fragment of the Cape Verde Islands accession in the Landsberg *erecta* genetic background) dormant and non-dormant seeds.

In **Chapter 5** we identified a novel ATP Binding Cassette gene (*ABCI20*), the knockout of which dramatically affects seed dormancy, longevity and ABA sensitivity. With the knowledge gained in **Chapter 2-4**, I performed transcriptome, translatome and ribosome proteome analysis on *abci20* seeds. The molecular processes that contribute to the distinct phenotype of *abci20* are discussed.

Finally, in **Chapter 6**, I integrate the findings on translational dynamics in the various aspects of seed physiology. Furthermore the potential for translational regulation for future research in seed biology and plant breeding are discussed.

Chapter 2

Extensive translational regulation during seed germination revealed by translational profiling

Bing Bai[1,2], Alessia Peviani[3], Sjors van der Horst[3], Magdalena Gamm[1], Berend Snel[3], Leónie Bentsink[2], Johannes Hanson[1,4]

[1] Department of Molecular Plant Physiology, Utrecht University, 3584 CH Utrecht, The Netherlands

[2] Wageningen Seed Laboratory, Laboratory of Plant Physiology, Wageningen University, 6708 PB Wageningen, The Netherlands

[3] Theoretical Biology and Bioinformatics, Utrecht University, 3584 CH Utrecht, The Netherlands

[4] Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden

Abstract

This work investigates the importance and extent of translational regulation during seed germination. The translational efficiency of each gene was determined by genome-wide profiling of total mRNA and polysome associated mRNA. This revealed extensive translational regulation during *Arabidopsis thaliana* seed germination. The translational efficiency of thousands of individual mRNAs strongly changes during the germination process. Intriguingly these changes are restricted to two phases, seed hydration and germination, referred to as the Hydration and Germination Shift. Sequence features, such as transcript length, secondary structures and the presence and location of specific motifs correlated with translational regulation. However, these differed significantly between the two shifts, which indicates that different mechanisms regulating translation operate during seed germination. This study reveals substantial translational dynamics during seed germination and identifies development dependent sequence features and cis elements that correlate with the translation control, uncovering a novel and important layer of gene regulation during seed germination.

Keywords: *Arabidopsis thaliana*, germination, ribosome, RNA structure, seedling establishment, translatomics

Introduction

Seed germination represents the start of a new plant life cycle. It involves the switch from a quiescent (dry seed) state to a metabolic active embryo which breaks through the encapsulating structures (endosperm and testa) to establish a young seedling. These early stages are critical for plant establishment and crop production. Arabidopsis seed germination is characterized by two visible events. First the testa (seed coat) ruptures exposing the underlying endosperm layer and secondly the endosperm ruptures which occurs when the root tip protrudes through the endosperm thereby completing germination sensu stricto. This is followed by growth and establishment of the seedling for autotrophic growth (Bewley, 1997).

Seed germination is triphasic starting with fast water uptake (imbibition, phase I). Genes encoding ribosomal proteins (r-proteins) are not transcribed at this developmental stage (Jimenez-Lopez et al., 2011). The first phase ends with a plateau phase (phase II) for water uptake featured by the activation of a series of metabolic processes facilitating energy production and reserve mobilization. During this process, ribosomal protein gene expression and ribosomal activity increases dramatically, facilitating the de novo synthesis of proteins important for seed germination (Fu et al., 2005; Dekkers et al., 2013; Galland et al., 2014). The last stage (phase III) is characterized by testa and endosperm rupture (germination) followed by radicle protrusion associated with post-germination events controlling seed to seedling transition. In the end, germinated seeds are equipped with the machineries and substrates necessary for autotrophic growth (Bewley, 1997). Intensive studies have elucidated the molecular changes during early seed imbibition and seed to seedling transition including transcriptome (Yu et al., 2014), proteome (Gallardo et al., 2001) and, metabolome (Fait et al., 2006) analyses. However, the steady state mRNA pool may not reflect the protein output due to lack of the linearity between transcription and translation (Gibon et al., 2006; Baerenfaller et al., 2008; Fernie & Stitt, 2012). Translational profiling employing sucrose gradient based fractionation allows the separation of mRNAs based on their association to polysomes and thus identifying mRNAs actively involved in translation. With high throughput mRNA

profiling techniques such as microarray analysis and RNA-sequencing, thousands of translated mRNAs could be quantified (Mustroph et al., 2009; Layat et al., 2014; Lin et al., 2014; Vragovic et al., 2015). Generating datasets of both the total mRNA as well as the polysomal bound mRNA allows the calculation of the ratio between the abundance of an individual mRNA in total mRNA fraction and the abundance in the polysomal mRNA fraction. Changes in this ratio between time points or between different treatments indicate changes in the translational efficiency (TE) showing that a certain mRNA is under translational regulation. This system has been successfully applied to investigate translational control in *Saccharomyces cerevisiae* (Arava et al., 2003; Halbeisen & Gerber, 2009; Ingolia et al., 2009), *Aspergillus fumigatus* (Krishnan et al., 2014), a mammalian cell line (de Klerk et al., 2015) and *Arabidopsis thaliana* (Jiao & Meyerowitz, 2010; Liu et al., 2012; Liu et al., 2013; Juntawong et al., 2014; Basbouss-Serhal et al., 2015).

In order to investigate the degree and dynamics of translational regulation as well as to identify gene sets under translational regulation during germination, transcriptome (total mRNA) and translatome (polysomal mRNAs) data sets were generated using microarray analysis of five consecutive stages during Arabidopsis seed germination. By combining polysomal and total RNA data the changes in translational efficiency on a genome-wide basis from the dry seed stage to seedling establishment were investigated. Thousands of individual mRNAs whose translational efficiency was affected during the germination process were identified. Intriguingly, translational regulation was not uniformly present throughout the germination process but was restricted to two temporal phases, one during seed hydration and one encompassing seed germination. Using bioinformatics analysis we were able to correlate the translational regulation to mRNA structure and the presence of sequence motifs present in these mRNAs. Thus, next to the strong transcriptional regulation observed previously during Arabidopsis seed germination, this study identified large sets of genes that are regulated on the translational level revealing an additional layer of gene expression regulation and its dynamics during germination.

Materials and methods

Plant material and growth conditions

Seeds of the *Arabidopsis thaliana* accession Columbia-0 were used for all assays described (NASC N60000). The timing of testa and endosperm rupture and seedling greening of fully after-ripened was performed as described previously (Joosen et al., 2010). In brief, two layers of Blue blotter paper (Anchorpaper company, www.seedpaper.com) were equilibrated with 48 ml demineralized water in plastic trays (15 x 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 (143 µmol $m^{-2}s^{-1}$). Germination parameters were manually counted.

For ribosome analyses, dry seeds were imbibed as mentioned above. Seeds and seedlings were harvested at each physiological state during seed to seedling transition, frozen in liquid nitrogen followed by freeze-drying. The dry material was stored at -80°C until further analyses.

Isolation of total RNA and polysomal RNA and polysome analysis

For the isolation of polysomal RNA, about 400 mg of freeze-dried tissue was extracted with 8 ml of polysome extraction buffer, PEB (400 mM Tris, pH 9.0, 200 mM KCl, 35 mM $MgCl_2$, 5 mM EGTA, 50 µg/mL Cycloheximide, 50 µg/mL Chloramphenicol) modified from (Subramanian, 1978). The extracts were loaded on top of a sucrose cushion (1.75 M sucrose in PEB) and centrifuged (18h, 90,000 g) using a Beckman Ti70 rotor for 18 h (Beckman Coulter, Brea, USA). The resulting pellet was resuspended in wash buffer (200 mM Tris, pH 9.0, 200 mM KCl, 0.025 M EGTA, 35 mM $MgCl_2$, 5 mM DTT, 50 µg/mL Cycloheximide, 50 µg/mL Chloramphenicol) and loaded on a 20-60% linear sucrose gradient, centrifuged at 190,000 g for 1.5 h at 4°C using Beckman SW55 rotor (Beckman Coulter). After ultracentrifugation, the gradients were fractionated into 20 fractions using a Teledyne Isco Density Gradient Fractionation System (Teledyne Isco. Lincoln, USA) with

online spectrophotometric detection of (254 nm). The fractions corresponding to the polysome region in the ribosome profile were pooled for future analysis. The ribosome abundance is reflected by the area under the curve and was calculated after subtracting the baseline obtained by measuring a blank gradient and normalizing to total area under the curve to account for possible uneven loading of the gradients.

Isolation and analysis of RNA species

The quantification of rRNA species in the unfractionated total RNA was done by isolating total RNA from dry seeds using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). An aliquot of 300 µl of total extract and 800 µl pooled polysome fraction was spiked with a mix of the four Eukaryotic Poly(A) RNA including *lys*, *phe*, *thr*, *dap* (Affymetrix, Ambion, P/N900433) with relative final concentration 1:100,000, 1:50,000, 1:25,000, 1:6,667 and purified with TriPure Isolation Reagent (Roche, Basel, Switzerland), further purified using RNeasy Mini spin columns (Qiagen, Hilden, Germany) and dissolved in RNase-free H_2O for further analyses. Ribosome fractions were calculated by determining the relative amounts of the small subunits of cytosolic, plastid and mitochondrial rRNAs by qRT–PCR assuming no presence of naked rRNA species (Piques et al., 2009).

cDNA was synthesized using the iScriptTM cDNA synthesis kit (Bio-Rad, Hercules, USA) according to the manufacturer's protocol. iQ-SYBRGreen-Supermix (Bio-Rad) was used for gene expression analysis on an MyIQ RT-qPCR machine (Bio-Rad). After DNase1 treatment (Thermo Scientific). Quantitative real-time PCR was performed using Power SYBR Green (Applied Biosystems, Waltham MA, USA) in a 5 µl reaction using the standard program of ViiA™ 7 instrument (Applied Biosystems). Data was analyzed using ViiA™ 7 Software v1.1 (Applied BioSystems). Primer amplification efficiency was calculated using LinRegPCR (Ruijter et al., 2009). The quantitation of the Poly(A) RNA control spikes were used to normalize real time qRT-PCR data. All primers used are provided in the Appendix.

Data analysis

Affymetrix Arabidopsis Gene 1.1 ST Arrays (Affymetrix, Santa Clara, USA) were hybridized using the GeneChip® 3› IVT Express kit (cat. # 901229) according to instructions from the manufacturer. Hybridization data were analyzed and gene specific signal intensities were computed using the R statistical programming environment (www.R-project.com) and the BioConductor package affy (Gautier *et al.*, 2004) and the Brainarray cdf file ver. 17.1.0 (http://brainarray.mbni.med.umich.edu/). DNA microarray data are available in the Gene Expression Omnibus (GEO) repository (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE65780. The limma and affy package were used for RMA normalization (Irizarry et al., 2003). Probe set intensity signals that never exceeded the noise threshold (logExprs<4 in all samples) were removed. A linear model and Empirical Bayes methods were applied for assessing differential expression (Smyth, 2004). Correlation between RMA normalized biological replicates averaged 0.96 (Pearson's correlation) and ranged between 0.93 and 0.98 (Figure S6). Relative RNA levels were validated with qRT-PCR experiments with four spikes as internal standard for the normalization (Vandesompele et al., 2002). DNA sequences and efficiencies of primer pairs used for qRT-PCR experiments and comparison of relative mRNA levels determined in GeneChip and qRT-PCR experiments are given in Table S1b. Principle component analysis (PCA) was performed using TM4 (Saeed et al., 2003).

Analysis of identified mRNA sequences

Gene trail (http://genetrail.bioinf.uni-sb.de/) and Revigo (http://revigo.irb.hr/) were used for over-representation analysis using default parameters to characterize the dominant transcriptional and translational processes related to seed germination. Geneset enrichment analysis, was performed in the Plant Geneset Enrichment Analysis toolkit (http://www.broadinstitute.org/gsea/index.jsp) by Fisher test followed by Hochberg FDR at the significance level of 0.05. GO-term enrichment analysis was performed for the translational shift gene sets using the topGO package (Alexa & Rahnenfuhrer, 2010).

Sequence feature analysis

Genes with significantly increased and decreased translational efficiency at each developmental shift were compared to the microarray background for several sequence features using custom scripts. The distributions of sequences length and GC content were evaluated separately for CDS, 5'UTR, 3'UTR, and full transcript. CDSs were also analysed for GC3 content, after removing sequences missing the start codon and/or containing premature stop codons; CDSs shorter than 100 codons were further removed for the codon bias analysis, measured using the Effective Number of Codons (Nc) index (Sun et al., 2013). The same analyses were performed separately for the CDS of protein-coding genes having both or no annotated UTR (UTRs called present when having length > 1 nt). Given the non-normality of the values distributions, a Wilcoxon signed-rank test was adopted for all statistical comparisons (median as test statistic).

RNA structural analysis

Experimentally determined structure scores per nucleotide, as provided by (Li et al., 2012), were used to calculate average structure scores of the genes with significantly increased and decreased ribosomal association at each developmental switch. Relative scaling was achieved by averaging the structure scores per region (5'UTR, CDS and 3'UTR) in 100 bins. Standard errors and Student's t-tests were performed using the Python SciPy module (http://www.scipy.org/).

Motif analysis

DNA motif analyses were performed using the MEME suite (Bailey et al., 2009), for full transcript, 5'UTR, CDS and 3'UTR sequences, extracted from the TAIR10 database (http://www.arabidopsis.org/). The minimum and maximum motif width was set to 6 and 10, respectively. If a gene had multiple isoforms, only the TAIR10 representative splice form was used. Background dinucleotide frequencies were provided separately for each sequence type. To test specificity of the resulting motifs, FIMO (Bailey et al., 2009) was used to scan all genes represented on the microarray

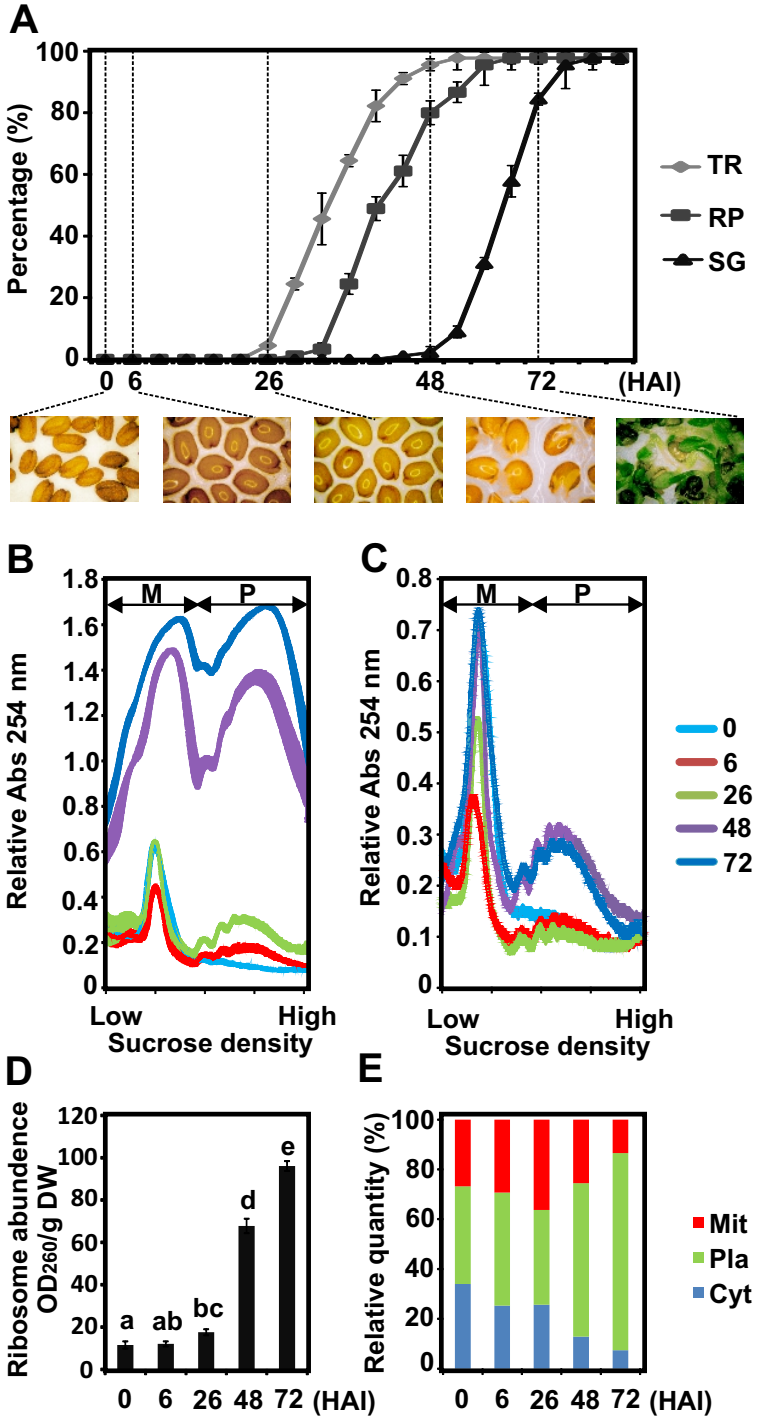
for motif hits in the corresponding sequence type. Motifs with *P*-value <= 0.001 were considered significant hits. Obtained motif counts were used to compute the enrichment *P*-value for the gene lists versus the background by means of a one-tailed Fisher's exact test, performed with a custom script and the R software package (http://www.r-project.org/). The occurrences for Table 1 were derived from the same FIMO outputs. For each motif, the positions on the transcripts, as provided by the FIMO output, were used to calculate the relative number of motifs per (relative) position along the mRNA. Relative scaling was performed in a similar fashion as for the structure scores.

Results

Translational activation precedes ribosome biogenesis during Arabidopsis seed germination

Monitoring the seed to seedling transition of fully after-ripened Arabidopsis seeds was performed by scoring testa rupture (TR), radicle protrusion (RP) and seedling greening (SG) over time. TR started around 26 hours after imbibition (HAI). RP was first observed 35 HAI and at 48 HAI 80% of the seeds showed radicle protrusion. By 72 HAI, 80% of the seedlings reached the SG stage and at 82 HAI all the seedlings had turned green (Figure. 1A). The time-points (0, 6, 26, 48 and 72 HAI) that mark different physiological stages (dry seeds, early imbibition, the initiation of TR, 80% RP, and 80% SG respectively) were selected for ribosome profiling based on both equal dry weight (Figure. 1B) and equal RNA loading (Figure. 1C). Ribosome profiles changed dramatically during the seed to seedling transition. In dry seeds, ribosomes were mainly present in the monosome form (Figure 1B, C). Following imbibition ribosome profiles changed. This was first visible by an increase in the polysome peak from dry to 6 HAI (Figure 1B) concurrent with a decrease of the monosome peak, followed by an increased total area that represents the increase in ribosome abundance (Figure 1B-D). These newly synthesized ribosomes mostly represent organellar ribosomes, especially plastid ribosomes after 48 HAI as shown by the relative quantity of ribosomal RNA specific to each organelle (Figure 1E).

Figure 1. Translatome profiling during seed germination. (A) Time points during Col-0 seed germination included dry seeds, 6, 26 (testa rupture (TR) initiation), 48 (80% radicle protrusion (RP)) and 72 hours after imbibition (HAI) (80% seedling greening (SG)). The data is presented as mean ± SD of three independent replicates. (B) Absorbance profiles of sucrose density gradient fractionated ribosomes for the five time-points during seed to seedling transition. The data is presented as mean ± SD of three independent replicates SD is indicated by the width of the line. The monosome and polysome region on the profile are labelled as M and P respectively. Gradient loading according to equal dry weight. (C) Representative absorbance profiles of sucrose density gradient fractionated ribosomes for the five time-points during seed to seedling transition, gradient loading according to identical RNA loading. (D) Mean ribosome abundance is represented by the OD260 unit normalized to the dry weight. (E) Relative abundance in of different ribosomes (cytosolic (Cyt), plastidic (Pla) and mitochondrial (Mit)). Error bars ± SD and the letters above each bar indicate the significance (P-value<0.05).


A
Percentage (%)
100
80
60
40
20
0
0 6 26 48 72 (HAI)
TR
RP
SG
B
Relative Abs 254 nm
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
M
P
Low
High
Sucrose density
C
Relative Abs 254 nm
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
M
P
Low
High
Sucrose density
0
6
26
48
72
D
Ribosome abundence
OD260/g DW
120
100
80
60
40
20
0
a
ab
bc
d
e
0 6 26 48 72 (HAI)
E
Relative quantity (%)
100
80
60
40
20
0
0 6 26 48 72 (HAI)
Mit
Pla
Cyt

Transcriptional changes are reflected in polysomal mRNA levels during seed germination

Total mRNA (T) and the polysomal mRNA (P) were analysed to investigate the translational dynamics during the seed to seedling transition. The change in mRNA abundance between each stage was determined by comparing RNA levels to the preceding stage during seed germination. The expression of several thousand genes changed to a similar extent in both the total mRNA and the polysomal mRNA in terms of number of differential expressed genes (Figure S1) and gene functions (Figure S2). The up-regulated genes were over-representing processes involved in protein localization, oxygen and reactive oxygen species metabolic process, ribosome biogenesis, translation, stress responses, cell wall organization, photosynthesis and lipid transport and localization (Figure S2B). In contrast, chitin response, abscisic acid response, defence, secondary metabolism, seed development, RNA processing and ribosome biogenesis were sequentially over-represented in the down-regulated gene set (Figure S2C). Generally, a correlation between transcription and translation across seed germination was observed.

Polysomal profiling reveals two phases of translational control

To identify the genes that are under translational control, we assessed the translational efficiency (TE) of each mRNA species, which is defined as the ratio between the mRNA in the polysome pool and the total mRNA (Bailey-Serres, 1999; Branco-Price et al., 2005; Branco-Price et al., 2008). By comparing the TE between each stage and the preceding time point, we identified two temporal phases with extensive changes in translational control: between dry seeds and 6 HAI seeds and between 26 and 48 HAI seeds (Figure 2). We refer to these phases as the Hydration and the Germination Shifts. In total 1204 genes were down-regulated in the Hydration Shift (hydration down) and 435 genes were up-regulated (hydration up). For the Germination Shift the numbers were 717 (germination down) and 195 (germination up). Minor significant translational changes were identified between 6 and 26 HAI and between 48 and 72 HAI (Figure 2B). To visualize the transcriptional and translational dynamics of

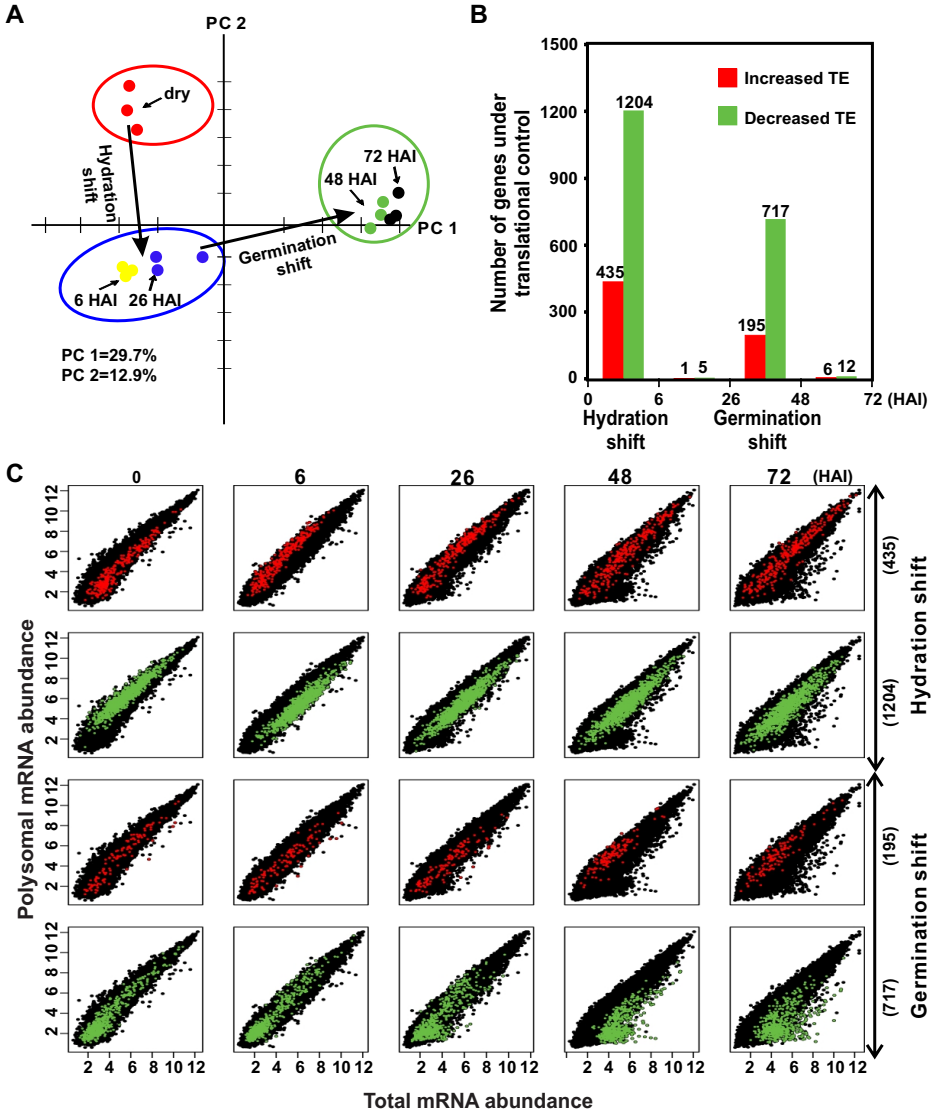


Figure 2. Arabidopsis seed germination is characterized by two translational shifts. (A) Principal component (PC) analysis of translational efficiency (TE) changes (Polysomal mRNA levels / total mRNA levels) during seed germination. The first two components (PC1 and PC2 explain 29% and 13% of the total variation, respectively. B) The number of mRNAs with changed TE at the two translational shifts at 6 HAI (Hydration Shift) and 48 HAI (Germination Shift). (C) Dynamics of genes underlying the two translational shifts. The levels of total (X-axis) and polysomal (Y-axis) mRNAs following seed imbibition (dry seeds, 6, 26, 48 and 72 hours after imbibition (HAI) are plotted (black). Genes identified with increased (red dots) or decreased (green dots) TE are indicated for both the Hydration Shift (upper panels) and the Germination Shift (lower panels).

the genes under translational control the four different gene sets were highlighted in correlation plots of the different time-points (Figure 2C). This showed that the Hydration up genes were relatively lowly expressed and similarly weakly associated to the polysomes in dry seeds. At 6 HAI these genes are associated to polysomes at higher levels than expected based on their expression, followed by similar levels in both pools during the later imbibition phases. The opposite pattern is shown for the Hydration down genes. These genes were generally highly associated to the polysome in dry seeds, and decreased in polysome association at 6 HAI. The Germination up gene set is specifically highly associated to the polysomes at 48 HAI and this continues in the later time point. The Germination down gene set is represented by mRNAs associated with polysomes at levels corresponding to total mRNA levels at early time points but specifically non-associated with polysomes at the two later time points. In principle changes in TE can be attributed to changes in both the transcriptome or translatome (or combinations thereof). By comparing these effects separately different patterns emerge. The Hydration down group of genes seems primarily affected negatively on the translatome level, while most of the Germination down genes are characterized by a dramatic upregulation on the transcriptome level (Figure S7). The microarray procedure used is determining relative changes. However, the striking different pattern between the down-regulated genes in the two shifts indicate different regulatory mechanisms at different stages during seed germination.

Transcript features correlate with translational regulation

To investigate whether transcript features correlate with translational regulation we determined transcript length and GC content of the translationally regulated mRNAs. It is established that short transcripts and transcripts with low GC content are in general more efficiently translated than long ones (Qu et al., 2011; Valleriani et al., 2011; Liu et al., 2012). For the Hydration Shift we found significantly longer genes in the down regulated set compared to the up-regulated and background gene sets. However, the Germination Shift showed an opposite pattern (Figure 3A, Figure

S3). This indicates that translation at the shifts is regulated by distinct mechanisms. Significantly higher GC contents were identified in the 5' UTR and 3' UTR of the Hydration Shift and in the coding sequence (CDS) of the Germination Shift down regulated genes, which correlates with suppression of the translation efficiency. Due to redundancy in the genetic code most amino acids are encoded by several synonymous codons, however it is thought that some codons are translated more efficiently than others. This codon bias is calculated based on the active number of codons and extremely biased genes only use one codon per amino acid (Nc of 20) (Hershberg & Petrov, 2008; Hershberg & Petrov, 2009). Codon degeneracy is nearly completely related to the third-base position (Crick, 1966) and highly accommodated by the GC content at this position. Interestingly, TE during seed hydration correlates negatively with effective number of codons (Nc) and positively with GC3 (Guanine and Cytosine content in the third codon position; synonymous sites) (Figure 3, Figure S3). This has been observed in several species (Ikemura, 1985; Bulmer, 1987; Akashi, 1994; Duret, 2000; Drummond & Wilke, 2008; Shabalina et al., 2013).

To further validate the identified sequence features, we mapped the translational regulated genes to the dataset used for investigating the genome-wide mRNA decay profiles and associated sequence features (Narsai et al., 2007). Interestingly, the genes translationally regulated in the Hydration Shift significantly differed in transcript stability. Translationally down-regulated genes of the shift are significantly less stable than expected by random and the up-regulated genes are more stable with significantly high number of introns especially in the CDS and 3'UTR, while other mRNA characteristics of the genes are not different from expected values (Table S1E). The role of the mRNA's secondary structure on translational control was tested by investigating an experimentally derived structural score defined by Li et al. (2012). This score is an indicator of transcript complexity in which high structure scores are equivalent to more double stranded (ds) than single stranded (ss) RNA at a certain position in a transcript and *vice versa* for low structure scores. Average structure scores were plotted over the 5'UTR, CDS and 3'UTR for both translational shifts (Figure 4A, B). In general, the background shows a steep decrease in structure

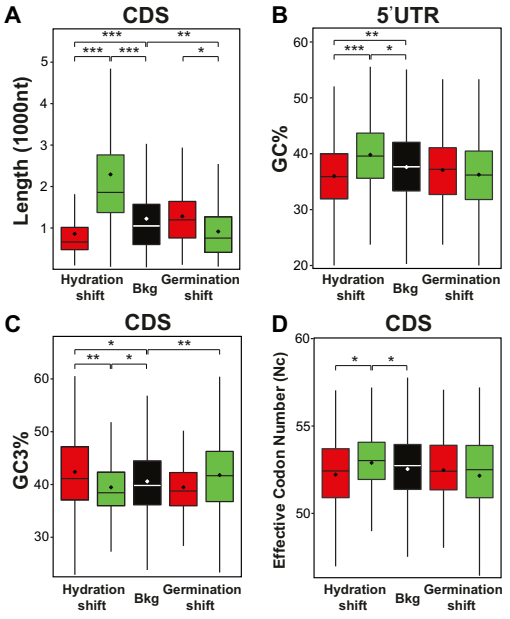
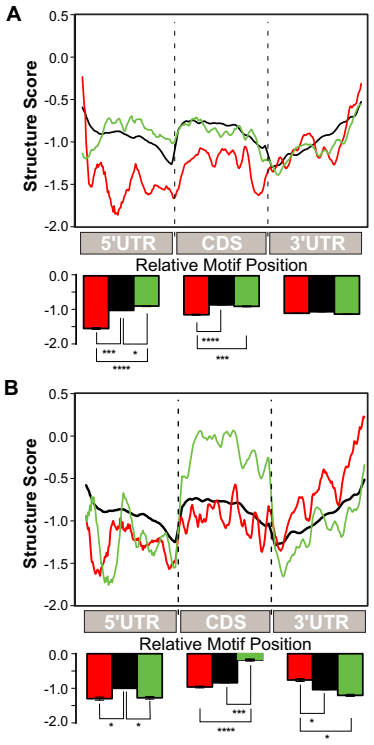


Figure 3. Translationally controlled mRNAs are characterized by distinct sequence features. (A) Length of coding sequence (CDS), (B) GC content of 5'UTR, (C) GC3 content and (D) Effective codon number (Nc) of CDS are shown. Colour scheme: black: microarray background (Bkg), red or green: translationally up- or down-regulated genes at the Hydration and Germination Shift respectively. Error bars ± SE (**P*-value<E-10, ** P-value<E-20, *** P-value<E-50. P-values are calculated by a Student's t-test.

at the start and stop codon, which is a conserved structural feature for eukaryotes facilitating the accessibility of ribosome for translation (Kozak, 2005; Kertesz et al., 2010; Li et al., 2012). At the Hydration Shift up regulated mRNAs are less structured in the 5'UTR and CDS than those down regulated, which suggests that mRNAs with lower structure scores are translationally favoured over more structured transcripts. At the Germination Shift down regulated mRNAs have an overall higher structure score in the CDS, and the opposite trend was found in the 3'UTR. The high structure of the CDS may attenuate the progression of the ribosomes and thereby inhibit translation of these mRNAs.

Motif analysis was performed on four regions (5'UTR, CDS, 3'UTR and the whole transcript) for both shifts (Figure S4). In total 5 significantly (*P*-value<E-5) enriched motifs were detected (Figure 4), all for the Hydration Shift. One was present in the 5'UTR of the hydration up transcripts, and four in the whole transcript RNA sequence of the hydration down set. The five motifs fall in two groups; UC (motif 1b, 1c and 2b) or GA (motif 3b and 3c) repeats (Figure S4). The UC-rich motifs (1b, 1c and 2b) are overrepresented in the entire transcript (Figure S4). These motifs possibly



C

Motif name	Motif logo	Number of genes with motif: Entire mRNA HS Down (1,204)	Entire mRNA Backgr. (27,827)	P-value
Motif 1b (0.0E000)		1,041	19,835	9.88E-54
Motif 2b (8.7E-060)		1,073	20,418	2.63E-61
Motif 2b (5.3E-152)		1,089	20,735	2.82E-65
Motif 3b (8.1E-216)		1,064	21,490	8.53E-39

D

Motif name (E-value)	Motif logo	Number of genes with motif: 5' UTR HS Up (343)	5' UTR Backgr. (19,159)	P-value
Motif 3c 3.8E-113		197	8,858	2.99E-06

Figure 4. Secondary structures and motif features correlate with polysome association. The average structure score is plotted over the 5'UTR, CDS and 3'UTR of the (A) Hydration Shift and (B) Germination Shift affected transcripts. Background; all transcripts on the array (black line), translational up (red line) or translational down transcripts (green line). Significant enriched motifs are detected across (C) the whole transcript length (cDNA) and (D) 5'UTR. The bar plots represent the mean structure scores per transcript region (5'UTR, CDS and 3'UTR, respectively). Error bars ± SE (*P-value<E-10, ** P-value<E-20, ***P-value<E-50). P-values are calculated by a Student's t-test.

attenuate the progression of ribosomes and thus inhibit translation elongation, e.g. by binding RNA binding proteins.

Motif 3c is significantly enriched in the 5' UTR of translationally enhanced transcripts in the Hydration Shift. This adenosine enriched poly(A) tract is mainly localized in the 50 nt region upstream of the start codon which could potentially bind to translation initiation factors and enhance translation initiation (Xia et al., 2011).

Seed germination is temporally featured by selective translation of evolutionarily old and young genes

To determine whether the genes of the two translational shifts are biased for evolutionary events, we classified the genes of the Hydration and Germination Shift using a phylostratigraphic approach. This approach has earlier identified an evolutionarily conserved transcriptome during embryo development and seed germination (Quint et al., 2012; Dekkers et al., 2013) and may be important to the spatiotemporal organization and differentiation during evolution of multicellular life. Here we used three phylostrata classes (PS), containing (1) genes that arose before plant evolution (PS1 and PS2 combined), (2) genes that arose during early plant evolution (algae and non-seed-bearing plants; PS3–PS5), and (3) the evolutionarily youngest genes (which evolved in seed bearing plants; PS6–PS12) (Quint et al., 2012). We evaluated whether mRNA transcripts during seed germination have an evolutionary preference for their translation. By this criterion, genes translationally down regulated at the two shifts were observed to be differentially enriched in different phylostrata classes (Figure 5). Genes decreased in TE at the Hydration Shift are enriched in PS 1-2. In contrast, genes decreased in TE at Germination Shift were enriched in PS 3-5 and PS 6-12.

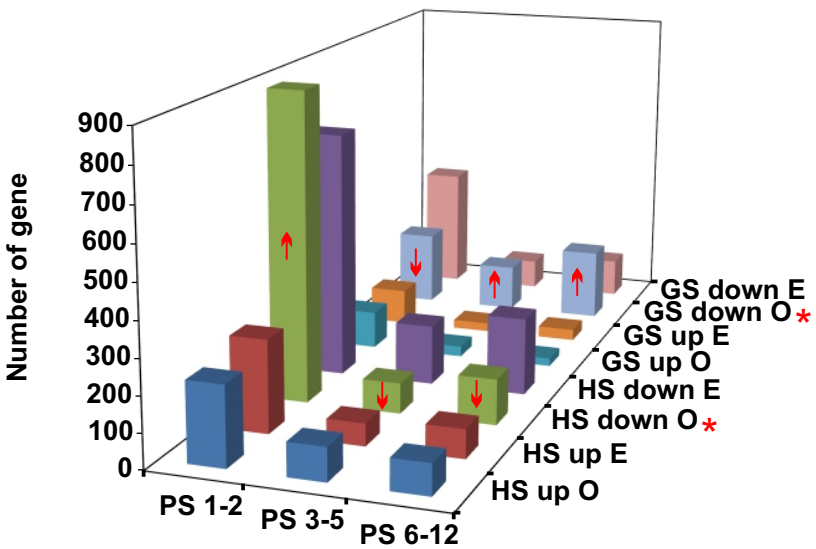


Figure 5. Evolutionary status of translationally regulated genes during Arabidopsis seed to seedling transition. The observed (O) translationally regulated gene number in Hydration Shift (HS) and Germination Shift (GS) in each phylostrata classes (PS) are compared with the expected (E) number based on the total gene set by a Chi-square test (*P*-value<0.01), gene sets for which the phylostrata organisation significantly differed from the expected are labelled with red asterisks. The red arrow indicates the direction of significant observed genes compared with the expected number of genes.

Discussion

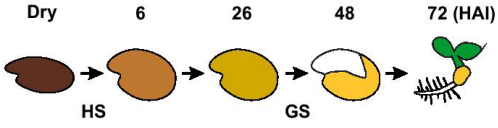
Translation is essential for seed germination (Rajjou et al., 2004; Narsai et al., 2011; Galland et al., 2014; Layat et al., 2014). However, an accurate temporal evaluation of the extent of translational regulation during seed germination is missing. Here we present a time course study on the translational dynamics, which reveals that translational activation precedes ribosome biogenesis during seed to seedling transition. Translatome profiling identified thousands of genes differentially transcribed and translated with temporal resolution. Thus, translatome profiling efficiently assists in identifying genes regulated during seed germination. Our data demonstrated a large overlap of differentially regulated genes on both the transcriptional and translational level, a natural consequence of mRNA dependent translation (Figure 2B, C). However, by analysing the translational efficiency, defined as the ratio of polysome associated and total mRNA levels, across the germination time course we found the two phases where the TE of thousands of mRNAs changes. Translational regulation during seed germination is extensive in comparison to other studies (Nicolai et al., 2006; Gamm et al., 2014; Lin et al., 2014). The two major shifts are here referred to as Hydration and Germination Shift respectively. These shifts occur in temporal correlation with key stages of seed to seedling transition and might refer to physiological control points. Comparison of the four groups of genes under translational control revealed hardly any overlap (Figure S5A). Thus, translational control during seed germination is stage specific and possibly represents different development dependent regulatory mechanisms as revealed by the different secondary structures and motifs identified in the two shifts. Translational regulation is partially conserved based on the overlap with genes previously shown to be translationally regulated under different environmental cues and plant growth

conditions (Figure S5B-F) (Nicolai et al., 2006; Branco-Price et al., 2008; Gamm et al., 2014; Juntawong et al., 2014; Lin et al., 2014; Sorenson & Bailey-Serres, 2014). The overlap is the largest with hypoxia translationally regulated genes (~25% of the total identified genes under translational control during seed germination), indicating that these translationally regulated genes are involved in modulating the molecular network linked to two processes. Thus, seeds may translationally modulate oxygen responses during the seed to seedling transition with similar regulatory mechanisms as during hypoxia stress.

We have identified sequence features that correlate with translational regulation (Figure 4). Different features correlate to the different translational shifts. For the Hydration Shift we found that reduced transcript length, GC content and secondary structure correlate with an up-regulation of translation. The different level of secondary structure of the translationally regulated mRNAs indicates the possibility that structural features are important for this regulation. Since the translationally regulated transcripts identified differ between the two shifts, there should be additional factors that affect the translational regulation. One can envisage a model in which the sensitivity to structure at different stages of seed to seedling transition is mediated by differential activity of RNA helicases, which are dedicated to unpacking the annealed nucleic acid strands such as secondary structures of RNA complex. RNA helicases have mostly been described to be responsive to abiotic stresses. However the Arabidopsis RNA helicase LOS4 (LOW EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 4) has been linked to the regulation of seed germination and other developmental phase changes (Gong et al., 2002; Gong et al., 2005). Whether LOS4 or other helicases play a role in the translational regulation of the two shifts remains to be investigated. Other factors that might affect translation are RNA binding proteins specifically interacting with the identified elements in translationally regulated mRNAs. The Arabidopsis genome encodes hundreds of RNA binding proteins and for one class, PUF (proteins, which are characterized by the presence of a conserved Pumilio homology domain) proteins, a role in germination has recently been proposed (Xiang et al., 2014). The pyrimidine (UC) and purine (GA) enriched

motifs identified among the transcripts in the Hydration Shift may represent binding sites for polypyrimidine tract-binding protein, PTB (Singh et al., 1995; Perez et al., 1997; Oberstrass et al., 2005). Motif 3b GAAGAAGAAG is similar to the target sequence (GAAGAAGAAGCUC) of SERINE/ARGININE-RICH PROTEIN SPLICING FACTOR 40 which acts as exon enhancer mediated by a complex of nuclear proteins (Yeakley et al., 1996). The Arabidopsis SR paralog SERINE/ ARGININE-RICH SC35-LIKE SPLICING FACTOR 33 has been identified and play a role in regulating alternative splicing (Thomas et al., 2012). Although not investigated here, splicing might play a role in the Hydration Shift as introns are specifically more frequent in Hydration Shift mRNAs (Table S1E).

Overall, our data reveal a model of changing translational regulation during seed germination and seedling establishment (Figure 6). The extensive translational regulation during germination and the changes therein are unlikely regulated by a single mechanism. The diversity of sequence features identified favours a multifactorial model. Further research will focus on how these identified features are recognized and thus mediate the translation control. The Arabidopsis genome harbours hundreds of mRNA binding proteins of which a large majority have no assigned function. Likely the regulators of translation during seed germination are to be found in this group of highly interesting proteins.



Sequence features correlate with translational shifts

HS	GS
5'UTR length	CDS length
5'UTR GC content	CDS GC content
GC3 content	CDS structure
Effective codon number	
5'UTR and CDS structure	
UC, GA enriched motif poly(A) enriched motif	

Figure 6. Summary of the sequence features in mRNAs identified as translationally regulated during germination. The Hydration Shift (HS) and Germination Shift (GS) are two translational shifts identified during seed germination where large sets of mRNAs are under translational control. The features correlated with translational shifts are largely non-redundant as listed.

Acknowledgements

This work was supported by the Netherlands Organization for Scientific Research (project number 821.02.010). Financial support came from the Marie Curie ITN project MERIT, Grant Agreement number 264474 (AP) and we thank Bio4Energy, a Strategic Research Environment appointed by the Swedish government, for supporting this work. We thank Service XS, Leiden, The Netherlands for performing the microarray hybridizations and Hajk-Georg Drost and Marcel Quint of the Leibniz Institute of Plant Biochemistry, Halle, Germany for providing information on the phylogenetic classes. We also thank Bas Dekkers for proof reading and suggestions on the manuscript.

Supporting Information

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

Figure S1. Transcriptional and translational expression shift following seed germination.

Figure S2. Temporal differences between transcriptome and translatome using over-representation analysis (ORA).

Figure S3. Comparison of sequence features between genes regulated at the Hydration and Germination Shift and the background

Figure S4. Spatial distribution of enriched motifs.

Figure S5. Dataset comparison of the shift genes during seed germination.

Figure S6. Gene 1.1 ST GeneChip quality assessment and reproducibility.

Figure S7. Gene expression profiling of translational regulated genes.

Supplemental Table S1. Gene set enrichment analysis for the translational shift genes.

Chapter 3

Fate specific mRNA storage in seeds

Bing Bai[1,2], Sjors van der Horst[2], Jan H.G. Cordewener[4,5,6], Antoine H.P. America[4,5,6], Leónie Bentsink[2], Johannes Hanson[1,3]

[1] Department of Molecular Plant Physiology, Utrecht University, 3584 CH Utrecht, The Netherlands

[2] Wageningen Seed Laboratory, Laboratory of Plant Physiology, Wageningen University, 6708 PB Wageningen, The Netherlands

[3] Umeå Plant Science Center, Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden

[4] BU Bioscience, Plant Research International, P.O. Box 619, 6700 AP, Wageningen, The Netherlands

[5] Centre for BioSystems Genomics, P.O. Box 98, 6700 AB, Wageningen, The Netherlands

[6] Netherlands Proteomics Centre, P.O. Box 80082, 3508 TB, Utrecht, The Netherlands

Abstract

Dry seeds contain large amounts of mRNAs. These transcripts accumulate during seed maturation and are thought to be important for seed germination. How these mRNAs are stored and protected from, for example, oxidation during seed storage is unknown. Here we investigated the functional relevance of stored mRNAs in Arabidopsis seeds. Hereto, we performed ribosome profiling and investigated the monosome (mRNAs with one ribosome attached) and polysome (mRNAs with more than one ribosome attached) fractions of dry seeds and seedlings for their associated mRNAs and proteins. We show that most stored mRNAs are associated to ribosomes. mRNAs co-purify with monosomes and approximately 30% of these monosome associated mRNAs are translationally up regulated upon seed imbibition. Seed mRNAs that are associated to polysomes are mostly leftovers from translation during late seed maturation, as determined by their expression pattern and ontology. These are associated with processing bodies (P-bodies) and to a large extent translationally down regulated upon seed hydration. The monosome associated mRNAs in seeds are characterized by aberrant GC content, secondary structures and specific motifs in the 5'UTR and before the start codon. Moreover, monosomes appeared to be associated with several types of mRNA binding proteins, and GLUTATHIONE S-TRANSFERASES, we propose a role for these proteins in translational regulation and in the protection of the seed stored mRNAs to oxidation during seed dry storage.

Keywords: translatome profiling, ribosome, seed, seedlings, *Arabidopsis thaliana*

Introduction

Seeds represent a unique stage during the plant life cycle and are characterized by their quiescent state and low metabolic activity. Dry seeds are heterotrophic and contain all the components that are required for germination and seedling establishment until the seedling reaches the autotrophic state and can photosynthesize. In seeds, all activities including cell-division, transcription, translation and general metabolism are put on hold (Marcus & Feeley, 1964; Dure & Waters, 1965; Bewley et al., 2013). Nevertheless, the seed is filled with bioactive molecules such as proteins and mRNAs. These stored mRNAs accumulate during seed development (Sano et al., 2015) and retain their function until germination (Hughes & Galau, 1989; Comai & Harada, 1990; Galau et al., 1991). Stored mRNAs have been detected in seeds and microarray analysis has identified thousands of stored mRNAs in Arabidopsis, rice, wheat and barley (Nakabayashi et al., 2005; Sreenivasulu et al., 2008; Howell et al., 2009; Yu et al., 2014; Sano et al., 2015). The translation of these stored mRNAs is selective in time by prioritizing the translation of proteins with specific function during seed germination and post-germination seedling growth (Sano et al., 2012; Galland et al., 2014).

Usually mRNAs are closely associated with mRNA ribonucleoprotein complexes (mRNPs) and subjected to stringent regulation of translation, degradation or storage. Three types of mRNPs are the ribosomes, the processing bodies (P-bodies) and the stress granules (SG). Often the interaction between mRNA and its binding proteins is dependent on environmental stimuli or developmental programs (Balagopal & Parker, 2009; Buchan & Parker, 2009; Decker & Parker, 2012). mRNPs contain large amounts of mRNA binding proteins (RBPs) that are involved in different aspects of mRNA metabolism including RNA biogenesis, processing, alternative splicing, polyadenylation, transport, stabilization, translation and decay. The binding targets of these RBPs are specific RNA structures such as RNA-recognition motifs (RRM), double-stranded RNA-binding motifs (dsRBMs) and Zinc finger binding motifs. Also in developing seeds, RBPs have been shown to be involved in RNA processing, translation, signalling and metabolism (Sami-Subbu et al., 2001; Doroshenk et

al., 2009). The accumulation pattern of some RBPs is associated with a distinct physiological state during seed development such as desiccation and dormancy establishment (Sano et al., 2013).

The mechanisms for mRNA storage in dry seeds remain to be described in detail, although reports have shown the existence of temporarily inactive mRNAs in seed embryos that are part of mRNPs (Marcus & Feeley, 1964; Dure & Waters, 1965; Chen et al., 1968; Ihle & Dure, 1970; Poulson & Beevers, 1973; Hammett & Katterman, 1975; Ajtkhozhin et al., 1976; Harris & Dure, 1976).

Here we investigated the functional relevance of stored mRNAs in Arabidopsis seeds. In this manuscript we provide an understanding on mRNA storage in dry seeds and how this differs from later stages such as seedlings. Specific structures and features of seed stored mRNAs are discussed in relation to seed development, germination and seedling growth and with the identification of seed specific mRBPs co-purified with stored mRNAs, a model including two pools of seed transcripts with differential activities is postulated.

Results

mRNAs are bound to ribosomes in both dry seeds and seedlings

Seeds contain mRNAs that have accumulated during seed maturation, however it has remained unclear how they are stored. To investigate if the stored mRNAs are bound to large protein complexes or present freely in the cell, size fractionation over a sucrose cushion was performed. mRNAs bound to large complexes, such as ribosomes, pass through the cushion while free RNAs are left in the supernatant. RNAs were isolated from after-ripened dry seeds (hereafter referred to as seed) and seedlings after 72 hours of imbibition, and centrifuged over a sucrose cushion. Total RNA and mRNA were isolated from both the supernatant and pellet and mRNA enrichment was estimated based on the ratio between mRNA and total RNA. This revealed a significant enrichment (~3 fold) of mRNA in the pellet compared with the supernatant (Figure 1A), which suggests that the mRNAs in both tissue types are mainly associated with large protein complexes. The large complexes from the pellet

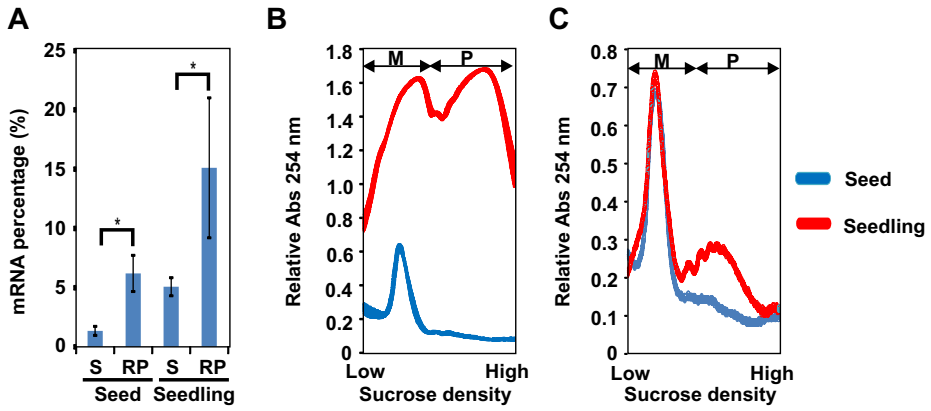


Figure 1. RNA content and ribosome profiles of seeds and seedlings. (A) mRNA fraction from the total RNA in both supernatant (non-ribosome) fraction (S) and ribosome pellet (RP) fraction in seed and seedlings. (B) Ribosome profiles of seed and seedling based on equal dry weight. Monosome (M) and polysome (P) region in the ribosome profile. (C) Ribosome profile of seed and seedling based on equal RNA loading. Data are mean ± SD. *P*-value<0.05 as indicated by asterisks. *P*-values are calculated by a Student's t-test (n=3)

were further purified by sucrose gradient centrifugation and this showed that seeds mainly contain monosomes and no visible polysomes, whereas seedlings contain a large quantity of monosomes as well as polysomes, which are indicative of active translation. (Figure 1B, C).

Monosome mRNAs prime seeds for germination

To identify the mRNAs that are specifically bound to the monosomes, microarray analysis was employed and the seed monosome (M) associated mRNAs were compared to those of seedling monosomes and seed and seedling polysomes (P). Among the 27,827 genes investigated on the array, 17,189 were identified as present in either seeds or seedlings. By differential expression analysis we identified the genes that were preferentially located in either the monosome or polysome fractions. In both seeds and seedlings monosome specific transcripts were identified (643 in seeds and 555 in seedlings) of which only about 10% were shared between the tissue types. Similarly 1418 polysome specific transcripts were identified in seeds and 1168 in seedlings, of which 192 are shared (Figure 2A). A large proportion of the monosome

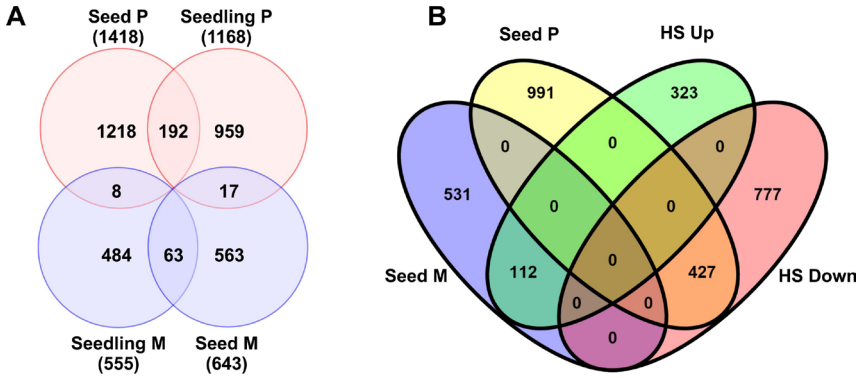


Figure 2. The comparison of monosome and polysome associated transcripts in seeds and seedlings. (A) Venn diagram for transcripts with specific ribosome association states (monosome (M) or polysome (P) in seeds and seedlings. (B) Venn diagram comparing transcripts that are specific to the monosome (M) or polysome (P) in seeds with those identified as translationally regulated during the Hydration Shift (HS) during seed germination (Chapter 2).

specific transcripts in seeds are later found to be specifically translated during germination, while the polysome specific transcripts in seeds are translationally down regulated during germination (Figure 2B, Chapter 2). This suggests that the seed specific mRNAs needed during germination associate specifically to monosomes while the polysome associated mRNAs might be leftovers of transcripts needed during seed maturation. In support of this hypothesis it was found that monosome associated mRNAs include those related to chromatin assembly such as *HISTONE1-3, 4, B2 B9*, *B11*, *R12*, *HIGH MOBILITY GROUP B4, B5*; abiotic stress response such as *heat shock proteins* (*HSPs*), *late embryogenesis abundant proteins* (*LEAs)* and *DEHYDRIN, PEROXIDASE*, *GLUTATHIONE S-TRANSFERASE7* and *26*, *THIOREDOXINS*, *SNF1-RELATED PROTEIN KINASE2.4*, *FLOWERING LOCUS C* and ribosome biogenesis related mRNAs. Polysome specific transcripts include those with functions in post-embryonic seed development such as over 20 *EMBRYO DEFECTIVE* (*EMB*) mRNAs, *EMBRYO-DEFECTIVE-DEVELOPMENT*, *MATERNAL EFFECT EMBRYO ARREST5*, *44, 65, EMBRYONIC FACTOR1*, *DEFECTIVE KERNEL1* and *TORMOZEMBRYO DEFECTIVE* and mRNAs that were identified as related to embryo development such as *AUXIN RESPONSE FACTOR2, 4, 6, 8* and auxin efflux transporter *ATP*

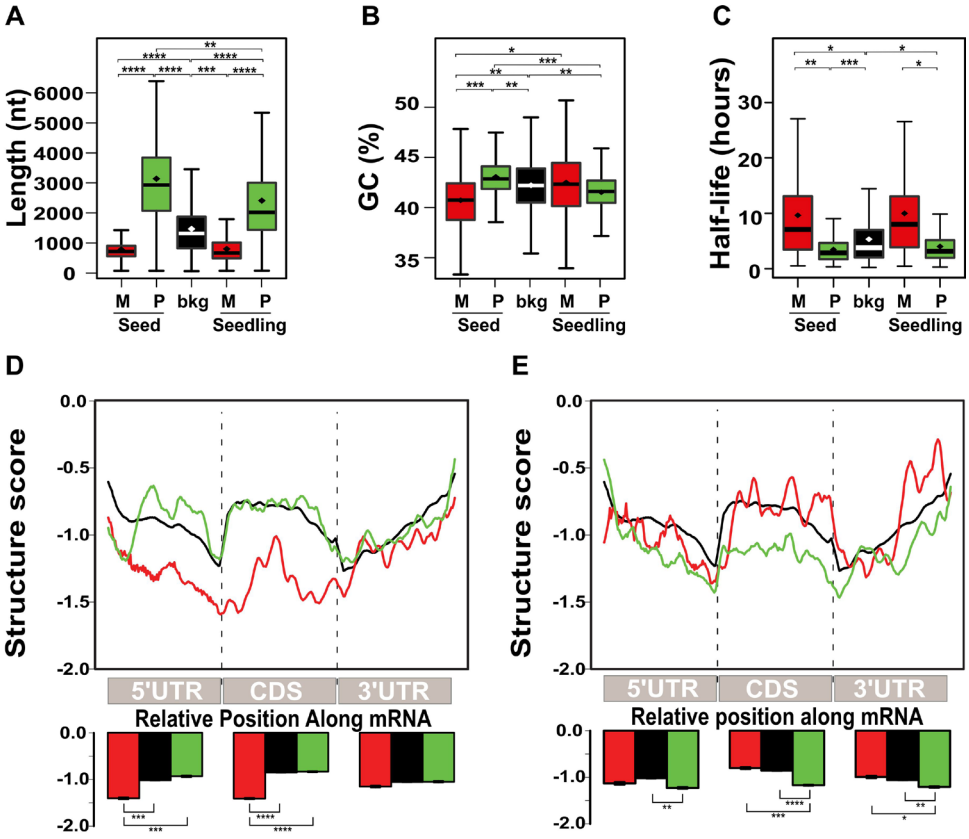
BINDING CASSETTE SUBFAMILY B19, *TARGET OF RAPAMYCIN* and *TREHALOSE-6-PHOSPHATE SYNTHASE* (Supplemental Table 1, 2).

Sequence characteristics of seed stored mRNAs

The sequence of monosome and polysome specific mRNAs in seeds and seedlings was investigated for length and GC content (GC%) of the full transcript and specific regions (5'UTR, 3'UTR and coding sequence, CDS), GC content in the third codon position (GC3%) and Effective Number of Codons Nc (a codon bias index) (Supplemental Figure 1). Monosome associated transcripts were significantly shorter in both seeds and seedlings, likely due to the difficulty of small mRNAs to accommodate several ribosomes. Moreover, monosome associated transcripts were low in GC content in seeds (Figure 3A, B). To reveal whether the mRNA stability of seed monosome transcripts differed from the other transcripts, the transcript half-life time was evaluated. For this, published mRNA half-life times of a study performed by Narsai et al (2007) were used. For 375 of the 643 seed monosome mRNAs half-life time data were available. Interestingly, we observed differential half-life times between the monosome and polysome associated transcripts in both tissues (Figure 3C). Monosome associated transcripts had a significant longer half-life time than polysome associated transcripts, indicating a higher stability of monosome associated transcripts compared to polysome associated transcripts.

Whether the stability of monosome mRNAs could be explained by their secondary structure was investigated using structural data from (2012) as an indicator of transcript complexity. High structure scores are equivalent to more double stranded RNA (dsRNA) than single stranded (ssRNA) at a certain position in a transcript and vice versa for low structure scores. Average structure scores were plotted over the 5'UTR, CDS and 3'UTR for both monosome- and polysome-specific transcripts and compared to all expressed transcripts in seeds or seedlings (Figure 3D and E). Seed monosome transcripts were significantly less structured in both 5'UTR and CDS compared to any of the other ribosome associated transcripts (Figure 3D).

Seeds appear to contain specific mRNAs that are associated with monosomes until



F

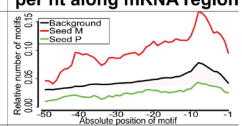
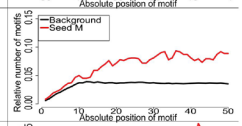
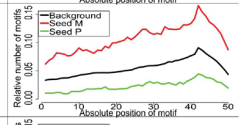
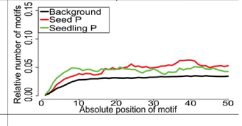
Motif name Dicovery region	Motif logo	Enriched region on mRNA	Gene list	*P*-value	FC	Relative number of motifs per nt along mRNA region
Motif 1a **Seed M** **(5'UTR)**	E-value: 6.4e-199	upstream start codon 5'UTR CDS 50 nt	Seed M Seed P Seedling M Seedling P	6.3e-21 4.0e-11 1.3e-1 1.1e-2	1.92 0.59 1.13 0.84	
		downstream 5' end Transcript 50 nt	Seed M Seed P Seedling M Seedling P	1.3e-17 2.4e-4 1.5e-1 8.7e-2	1.90 0.78 1.11 1.10	
Motif 6b **Seedling P** **(3'UTR)**	E-value: 1.6e-18	upstream start codon 5'UTR CDS 50 nt	Seed M Seed P Seedling M Seedling P	3.6e-14 1.1e-17 5.4e-2 1.1e-3	1.69 0.50 1.17 0.80	
Motif 3b **Seed P** **(5'UTR)**	E-value: 2.9e-116	downstream 5' end Transcript 50 nt	Seed M Seed P Seedling M Seedling P	1.2e-2 2.6e-14 5.1e-1 1.7e-8	0.79 1.49 1.00 1.39	

Figure 3. Sequence features of transcripts associated with the monosome (M) or polysome (P) in seeds and seedlings. (A) The length of the transcripts with specific ribosome association in seeds and seedlings (B) The GC% of the transcripts with specific ribosome association in seeds and seedlings. (C) mRNA half-life of monosome and polysome specifc transcripts in seeds and seedlings. *P-value<E-10, ** P-value<E-20, *** P-value<E-50 **** P-value<E-100. (D) Secondary structure of transcripts specifically associated with the monosome (red line) or polysome (green line) in seeds. (E) Secondary structure of transcripts specifically associated with the monosome (red line) or polysome (green line) in seedlings (*P-value<E-10, **P-value<E-20, ***P-value<E-50 ****P-value<E-100). P-values are determined by t test. (F) Motifs identified in transcripts specifically associated with monosome and polysome in seeds. Motif name and discovery region, Motif logo with E-value, differentially enriched motif position along trasncript, gene lists, P-value of the difference compared to the background signal and fold change (FC) enrichment compared with background signal and the relative number of motifs per nucleotide (nt) along the mRNA are shown. FC labled with green and red color indicate that the motif is significantly over- or under-represented in the transcript set.

conditions allow seed germination. Possibly specific RNA sequence motifs play a role in the recognition of these mRNAs. Therefore, a motif analysis was performed on different regions of the transcript (5'UTR, CDS, 3'UTR and the whole transcript) (Figure 3F). Two motifs specific for seed monosome mRNAs were identified. In the 5'UTR and before the start codon of the seed monosome specific transcripts, a motif with the consensus GAAGAAGAA (motif 1a) was significantly overrepresented (Figure 3F). It peaks around 9 bases prior to the start codon. Moreover, the same motif at the same position was underrepresented in seed polysome transcripts. Similarly, motif AAGAAAAAGA (Motif 6b) was enriched in the seed monosome transcripts and also this motif was underrepresented in seed polysome specific transcripts.

The dry seed ribosomal proteome

mRNAs specifically translated during early seed imbibition are apparently often bound to monosomes in the seed. To investigate the protein composition of these monosome complexes proteomics was performed. Protein extracts from the seed monosome and polysome sucrose gradient fractions were digested and the resulting peptide mixtures were analysed using high-resolution LC-MS. Although the polysome peak was hardly visible in dry seeds (Figure 1 B, C), proteins and RNA could be isolated from both fractions (Supplemental Figure 2). In total, 1,398 proteins were identified in at least one ribosomal fraction (monosome or polysome)

in seeds (Supplemental table 6). These ribosomal fractions may contain, in addition to the actual ribosomal proteins, ribosome associated proteins as well as (partial) proteins being translated on the active ribosomes. In the ribosome fraction, 868 proteins (62%) were detected in both the monosome and polysome fractions of seeds. The identified seed ribosomal proteome covers a wide range of protein functions and the most prominent functional groups are identified as ribosomal proteins (209 identified), which represent 86.4% of the 242 ribosomal proteins annotated in Arabidopsis; proteins with RNA binding capacity (271 identified) including 108 RBPs and 163 ribosomal proteins (Supplemental Figure 3A, Supplemental table 7, 9 and 12). Other translation related factors were also significantly enriched, such as translation initiation factors and translation elongation factors, confirming the fractions to mainly consist of translational complexes. Comparing the proteins identified from monosome and polysome fractions identified 363 and 374 proteins with a biased association to either the monosome or polysome fraction, respectively, including 108 and 99 proteins that were only identified in association with either the monosome or polysome fraction (Supplemental Figure 3B, Supplemental table 10). 19 and 33 of the identified RBPs were significantly higher in abundance in either the monosome or polysome fraction, respectively (Table 1). The monosome specific RBPs include five glutathione S-transferases and two translation elongation factors (Supplemental table 13).

Dry seeds are translationally inactive but upon imbibition monosome mRNAs are specifically translated, whereas polysome mRNAs appeared to be remnants of seed maturation (Figure 2B). Seeds contain two types of mRNA processing machineries, stress granules (SG) and P-bodies (Balagopal & Parker, 2009). Stress granules normally function in temporary mRNA storage and translation initiation while P-bodies are involved in mRNA degradation and translational repression. Since not many components of plant stress granules and P-bodies have been identified we made use of what is known from studies in mammals. This revealed that both the monosome and the polysome fraction contain proteins putatively associated with stress granules or protein bodies (Table 2). P-body related proteins seem to be over-

Table 1. Putative RNA binding proteins over-represented in seed monosome and polysome.

RBPs	Protein Symbol	Protein Name
Monosome		
AT1G59700	ATGSTU16	Glutathione S-Transferase Tau 16
AT1G78380	ATGSTU19	Glutathione S-Transferase Tau 19
AT1G69410	ELF5A-3	Eukaryotic Translation Initiation Factor 5A-3
AT2G30860	ATGSTF9	Glutathione S-Transferase Phi 9
AT2G47730	ATGSTF8	Glutathione S-Transferase 6
AT3G55620	EMB1624	Embryo Defective 1624
AT5G11200	DEAD/DEAH BOX 15	Dead-Box ATP-Dependent RNA Helicase 15
AT2G29450	ATGSTU5	Glutathione S-Transferase 103-1A
AT1G18070	EF1A	Translation Elongation Factor EF1A
AT1G57720	EF1B	Elongation Factor 1-Gamma 2
AT1G68010	HRP	Hydroxypyruvate Reductase
AT2G21660	ATGRP7	Glycine Rich Protein 7
AT2G33410	RRM/RBD/RNP	RNA-Binding (RRM/RBD/RNP Motifs) Family Protein
AT5G07350	TUDOR1	Tudor-Sn Protein 1
AT5G54900	ATRBP45A	RNA-Binding Protein 45A
AT5G61780	TUDOR2	Tudor-Sn Protein 2
AT5G66190	FNR1	Ferredoxin--NADP Reductase 1
AT1G29880	AT1G29880	Glycyl-tRNA Synthetase 1
AT5G16840	BPA1	Binding Partner Of ACD11 1
Polysome		
AT1G80070	SUS2	Abnormal Suspensor 2
AT3G50670	U1-70K	U1 Small Nuclear Ribonucleoprotein 70 Kda
AT3G15590	TPR-LIKE	Pentatricopeptide Repeat-Containing Protein
AT5G41770	AT5G41770	Crooked Neck Protein
AT3G05060	NOP56-LIKE	Probable Nucleolar Protein 5-2
AT5G60790	ATGCN1	General Control Non-Repressible 1
AT4G31180	AT4G31180	At4G31180
AT2G29140	APUM3	Pumilio 3
AT1G53280	AT1G53280	At1G53280
AT3G20250	APUM5	Pumilio 5
AT1G54270	EIF4A-2	Eukaryotic Initiation Factor 4A-2
AT1G64550	ATGCN3	General Control Non-Repressible 3
AT5G47010	LBA1	Regulator Of Nonsense Transcripts 1 Homolog
AT2G33340	MAC3B	Pre-mRNA-Processing Factor 19 Homolog 2
AT3G02720	AT3G02720	At3G02720
AT3G02760	AT3G02760	At3G02760
AT1G56110	NOP56	At1G56110
AT5G53440	AT5G53440	At5G53440
AT1G76810	EIF-2	Eukaryotic Translation Initiation Factor 2
AT1G55310	SR33	SC35-Like Splicing Factor 33
AT1G07360	AT1G07360	Zinc Finger CCCH Domain-Containing Protein 4
AT1G32790	CID11	CTC-INTERACTING DOMAIN 11
AT1G67680	AT1G67680	SRP72 RNA-Binding Domain
AT2G18510	EMB2444	Embryo Defective 2444
AT2G39780	RNS2	Ribonuclease 2
AT2G40660	AT2G40660	OB-Fold-Like Protein
AT3G14450	CID9	CTC-Interacting Domain 9
AT3G55460	SCL30	SC35-Like Splicing Factor 30
AT3G61860	RSP31	Arginine/Serine-Rich-Splicing Factor Rsp31
AT5G52040	ATRSP41	Arginine/Serine-Rich-Splicing Factor Rsp41
AT1G76010	AT1G76010	Alba DNA/RNA-Binding Protein
AT2G37340	RSZ33	Arginine/Serine-Rich Zinc Knuckle-Containing Protein 33
AT3G26420	ATRZ-1A	At3G26420

Table 2. Putative stress granule (SG) and P-body (PB) components identified in association with the seed monosome and polysome fractions. Monosome&Polysome indicates that the protein is present equally in the monosome and polysome fractions.

Putative stress granule (SG) and P-body (PB) components identified in association with seed monosome and polysome							
AGI	**Ribosome association**	**Granule Type**	**Protein Symbol**	**Protein Name**	**Oganism**	**Homolog**	**Reference**
AT2G33410	Monosome	SG/PB	RRM/RBD/RNP	RNA-binding (RRM/RBD/RNP motifs) family protein	Mammals	Musashi	Kawahara et al, 2008
AT5G10450	Monosome	SG	GRF6	14-3-3-Like Protein Gf14 Lambda	Mammals	14-3-3 protein	Stoecklin et ak, 2004
AT5G16050	Monosome	SG	GRF5	14-3-3-Like Protein Gf14 Upsilon	Mammals	14-3-3 protein	Stoecklin et ak, 2004
AT1G54270	Polysome	SG	EIF4A-2	Eukaryotic Initiation Factor 4A-2	Mammals	EIF4A	Kedersha et al, 2000
AT2G17870	Polysome	SG	AT2G17870	Cold Shock Domain Protein 3	Mammals	Lin28	Balzer and Moss, 2007; Polesskaya et al, 2007
AT2G29140	Polysome	SG	APUM3	Pumilio 3	Mammals	PUM1/2	Vessey et al, 2006;
AT3G20250	Polysome	SG	APUM5	Pumilio 5	Mammals	PUM1/2	Vessey et al, 2006;
AT4G36020	Polysome	SG/PB	CSDP1	Cold Shock Domain Protein 1	Mammals	Lin28	Balzer and Moss, 2007; Polesskaya et al, 2007
AT5G47010	Polysome	SG/PB	LBA1/UPF1	Low-Level Beta-Amylase 1	Arabidopsis	UPF1	Arciga-Reyes et al, 2006;
AT1G02080	Polysome	PB	AT1G02080	CCR4-Not complex component, Not1	Mammals, yeast	CCR4	Sheth and Parker, 2003; Andrei et al, 2005
AT1G03790	Polysome	PB	SOM/TZF4	Zinc Finger CCCH Domain-Containing Protein 2	Arabidopsis	TZF4	Bogamuwa and Jang, 2013;
AT1G76810	Polysome	PB	EIF-2	Eukaryotic Translation Initiation Factor 2	Mammals	EIF-2	Kimbal et al, 2003
AT2G17870	Polysome	PB	CSP3	Cold Shock Domain Protein 3	Mammals	Lin28	Balzer and Moss, 2007; Polesskaya et al, 2007
AT1G14170	Monosome&Polysome	SG/PB	AT1G14170	RNA-Binding KH Domain-Containing Protein	Mammals	PCBP2	Fujimura et al, 2008
AT1G48410	Monosome&Polysome	SG/PB	AGO1	Protein Argonaute	Arabidopsis	AGO1	Morel et al, 2002
AT2G21060	Monosome&Polysome	SG/PB	ATGRP2B	Glycine-Rich Protein 2B	Mammals	Lin28	Balzer and Moss, 2007; Polesskaya et al, 2008
AT3G13300	Monosome&Polysome	PB	VCS	Varicose	Arabidopsis	VCS	Goeres et al. 2007; Xu and Chua, 2009
AT1G22300	Monosome&Polysome	SG	GRF10	14-3-3-Like Protein Gf14 Epsilon	Mammals	14-3-3 protein	Stoecklin et ak, 2004
AT1G54080	Monosome&Polysome	SG	UBP1A	Oligouridylate-Binding Protein 1A	Mammals, yeast	TIA-1	Kedersha et al, 1999; Buchan et al, 2008
AT1G56340	Monosome&Polysome	SG	CRT1	Calreticulin-1	Mammals	Calreticulin	Decca et al, 2007
AT1G78300	Monosome&Polysome	SG	GRF2	14-3-3-Like Protein Gf14 Omega	Mammals	14-3-3 protein	Stoecklin et ak, 2004
AT2G27040	Monosome&Polysome	SG	AGO4	Argonaute 4	Mammals	AGO2	Leung et al, 2006
AT4G09000	Monosome&Polysome	SG	GRF1	14-3-3-Like Protein Gf14 Chi	Mammals	14-3-3 protein	Stoecklin et ak, 2004
AT5G44500	Monosome&Polysome	SG	AT5G44500	Small Nuclear Ribonucleoprotein Family Protein	Mammals	SMN	Bechade et al, 1999; Hua and Zhou, 2004; Liu and Gall, 2007
AT5G47210	Monosome&Polysome	SG	AT5G47210	Hyaluronan / mRNA Binding Family	Mammals	SERBP1	Goullet et al, 2008
AT5G65430	Monosome&Polysome	SG	GRF8	14-3-3-Like Protein Gf14 Kappa	Mammals	14-3-3 protein	Stoecklin et ak, 2004
AT3G18130	Monosome&Polysome	SG	RACK1C	Receptor for activated C kinase 1C	Mammals	RACK1	Arimoto et al, 2008

represented in the polysomes.

Discussion

Dry seeds contain stored mRNAs. Seeds of some species (such as palm and lotus) are capable of tolerating dry storage for more than thousand years without losing the ability to germinate and grow (Shenmiller *et al.*, 1995; Sallon *et al.*, 2008). How these mRNAs are stored, protected and later fulfill a function in supporting seed germination and seedling growth has remained largely elusive. The investigations presented here provide insight into the role of the seed stored mRNAs in Arabidopsis and reveals possible mechanisms for their protection by ribosomes.

In this study we have first shown that seed stored mRNAs are mostly bound to large protein complexes as is the case for mRNAs in actively metabolizing cells (Figure 1A). Ribosome profiling using a sucrose gradient and UV absorbance detection shows that these protein complexes resemble the molecular weight of the 80S monosome and no complexes of other sizes could be clearly detected in seeds (Figure 1B, C). However proteomic analysis of the fractions corresponding to the polysome fraction on the dry seed revealed the presence of polysomes even though they were not detectable based on the UV absorbance signal of the sucrose gradient. Detailed analyses of the monosome protein complex revealed that, next to numerous ribosomal proteins, proteins known to be involved in cellular detoxification processes were abundantly present in the isolated seed monosomes. Among these were five GLUTATHIONE S-TRANSFERASES (GSTs). GSTs are important for protection against oxidative stress (Devos et al., 1994; DePaula et al., 1996; Rhazi et al., 2003). Oxidative stress is threatening mRNAs during dry storage, since reactive oxygen species (ROS) and oxidized glutathione accumulate during seed desiccation and seed storage (Devos et al., 1994; DePaula et al., 1996; De Gara et al., 2003; Rhazi et al., 2003). Also other identified monosome associated proteins are related to redox balancing, including FERREDOXIN NADP REDUCTASE and HYDROXYPYRUVATE REDUCTASE, supporting the importance of the protection of monosome associated mRNAs against oxidation (Supplemental Table 1).

It has been suggested that seed stored mRNAs are important for seed germination by supporting protein translation (Harris & Dure, 1976; Aspart et al., 1984; Suzuki & Minamikawa, 1985; Beltran-Pena et al., 1995; Nakabayashi et al., 2005). Using whole genome transcriptome profiling we were able to identify the transcripts that are associated with monosomes in the dry state and that are translated (associated to polysomes) upon seed imbibition. Here we show that 30% of the mRNAs that are bound to monosomes are actually translationally regulated during germination.

In contrast, mRNAs that are bound to polysomes in dry seeds appear to be remnants of translation during seed maturation. These mRNAs are translationally down-regulated upon seed imbibition. Examples are genes that are important for embryo development like the *EMB* genes. The polysome fractions included several proteins that have been related to P-bodies, which is another indication that mRNAs attached to these ribosomal complexes are being broken down since P-bodies have been associated with non-translating mRNAs, degrading mRNAs, and factors involved in translation repression, mRNA decay, RNA interference, and non-sense mediated mRNA decay (NMD) (Balagopal & Parker, 2009; Decker & Parker, 2012). The monosome associated mRNAs that are not translationally enhanced during germination are possibly the results of artefactual enrichment of very short mRNAs, that like the mRNAs of the polysome fraction are that are translationally down regulated during germination.

It is presently unknown how monosome-specific mRNAs are retrieved in the monosome complexes, marking them for translation during imbibition. Our data provide several indications such as a possible specific mRNA structure or the presence of specific sequence motifs in the selected mRNAs. Interestingly, we identified sequence motifs specifically in the 5' region and before the start codon. Possibly, mRBP binding inhibits translation in the seeds, or inhibits progression of the ribosome from the translational initiation sites, and primes these mRNAs for translation during imbibition. Perhaps independent factors exist that determine the association of these mRNAs with monosomes. The proteomic investigation of the monosome fraction supports this suggestion as several proteins annotated as having

mRNA binding function are specifically present in this fraction. It remains to be determined whether specific interactions exist between mRNA binding proteins and the selected mRNAs but the finding of specific sites present in the mRNAs and the enrichment of specific mRBPs favors this model.

In conclusion, in this manuscript we show that selected mRNAs synthesized during late seed development, are specifically stored to function during seed germination. We also substantiate these findings by characterizing the mode of storage as the selected mRNAs co-purify with monosome and specific mRNA binding proteins, and proteins previously shown to be part of stress granules.

Materials and methods

Plant material and growth conditions

Seeds of the *Arabidopsis thaliana* accession Columbia-0 were used for all assays described. The timing of testa and endosperm rupture and seedling greening of fully after-ripened was performed as described previously (Joosen *et al.*, 2010). In brief, two layers of Blue blotter paper (Anchorpaper company, www.seedpaper.com) were equilibrated with 48 ml demineralized water in plastic trays (15 x 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 (143 μmol $m^{-2}s^{-1}$).

For ribosome analyses, seeds and green seedlings were harvested and frozen in liquid nitrogen followed by freeze-drying. The freeze dried material was stored at -80°C until further analyses.

Isolation of total RNA and polysomal RNA and polysome analysis

For the isolation of polysomal RNA, approx. 400 mg of freeze-dried tissue was extracted with 8 ml of polysome extraction buffer, PEB (400 mM Tris, pH 9.0, 200 mM KCl, 35 mM $MgCl_2$, 5 mM EGTA 50 μg/mL Cycloheximide, 50 μg/mL Chloramphenicol) modified from (Subramanian, 1978). The extracts were loaded on

top of a sucrose cushion (1.75 M sucrose in PEB) and centrifuged (18h, 90,000 g) using a Beckman Ti70 rotor (Beckman Coulter, Brea, USA). The resulting pellet was resuspended in wash buffer (200 mM Tris, pH 9.0, 200 mM KCl, 0.025 M EGTA, 35 mM $MgCl_2$, 5 mM DTT, 50 µg/mL Cycloheximide, 50 µg/mL Chloramphenicol) and loaded on a 20-60% linear sucrose gradient, centrifuged at 190,000 g for 1.5 h at 4°C using Beckman SW55 rotor (Beckman Coulter). After ultracentrifugation, the gradients were fractionated into 20 fractions using a Teledyne Isco Density Gradient Fractionation System (Teledyne Isco. Lincoln, USA) with online spectrophotometric detection of (254 nm). The fractions were pooled into samples containing monosome and polysomal complexes. The ribosome abundance was measured by the area under the curve and was calculated after subtracting the absorbance obtained from blank gradient and normalizing to total area under the curve to account for possible uneven loading of the gradients.

Quantification of mRNA

The quantification of mRNA species from the ribosome fraction and the supernatant is based on a total RNA isolation using TriPure Isolation Reagent (Roche, Basel, Switzerland), further purification was performed RNeasy Mini spin columns (Qiagen, Hilden, Germany) and the RNA was dissolved in RNase-free H_2O for further analyses. The mRNA species from each fraction are then purified by Dynabeads® Oligo(dT)25 and both the total and mRNA fraction were quantified by Qubit® RNA HS Assay Kit (ThermoFisher, Waltham, USA). The mRNA content is represented by the ratio between the mRNA and total RNA in each fraction.

Microarray data analysis

Affymetrix Arabidopsis Gene 1.1 ST Arrays (Affymetrix, Santa Clara, USA) were hybridized using the GeneChip® 3› IVT Express kit (cat. # 901229) according to instructions from the manufacturer. Hybridization data were analyzed and gene specific signal intensities were computed using the R statistical programming environment (www.R-project.com) and the BioConductor package Affy (Gautier et

al., 2004) and the Brainarray cdf file ver. 17.1.0 (http://brainarray.mbni.med.umich.edu/). Our DNA microarray data are available in the Gene Expression Omnibus (GEO) repository (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE76905. The limma and affy package were used for RMA normalization (Irizarry et al., 2003). Probe set intensity signals that never exceeded the noise threshold (logExprs<4 in all samples) were removed. A linear model and Empirical Bayes methods were applied for assessing differential expression (Smyth, 2005; Diboun et al., 2006). Correlation between RMA normalized biological replicates averaged at 0.96 (Pearson's correlation) and ranged between 0.93 and 0.98 (Supplementary Figure 6). Relative RNA levels were validated with qRT-PCR experiments with four spikes as internal standard for the normalization (Vandesompele et al., 2002).

Sequence feature analysis

The four different gene sets (Seed monosome, Seed polysome, Seedling monosome and Seedling polysome) were compared to each other and the microarray background for several sequence features using custom scripts. The distributions of sequences length, GC content and nucleotide content were evaluated separately for CDS, 5'UTR, 3'UTR, and full transcript. CDSs were also analyzed for GC3 content, after removing sequences that were missing the start codon and/or containing premature stop codons; CDSs shorter than 100 codons were further removed for the codon bias analysis, measured using the Effective Number of Codons (Nc) index (Sun et al., 2013). The same analyses were performed separately for the CDS of protein-coding genes having both or no annotated UTR (UTRs called present when having length > 1 nt). p values are calculated by ranked sum wilcoxon test (length and codon bias) or t test (GC%, GC3% and nt content).

RNA structure and stability analysis

Experimentally determined structure scores per nucleotide, as provided by (Li et al., 2012), were used to calculate average structure scores of the genes with significantly increased and decreased ribosomal association at each gene set (seed or seedling and

monosome or polysome). Relative scaling was achieved by averaging the structure scores per region (5'UTR, CDS and 3'UTR) in 100 bins. Standard errors and Student's t-tests were performed using the Python SciPy module (http://www.scipy.org/). Transcript stability was evaluated by retrieving data from (Narsai et al., 2007). Student's t-tests were performed using R (https://www.r-project.org/).

Motif analysis

DNA motif analyses were performed using the MEME suite (Bailey et al., 2009), for full transcript, 5'UTR, CDS and 3'UTR sequences of each gene set (DS or GS and monosome or polysome), extracted from the TAIR10 database (http://www.arabidopsis.org/). The minimum and maximum motif width was set to 6 and 10, respectively. If a gene had multiple isoforms, only the TAIR10 representative splice form was used. Background dinucleotide frequencies were provided separately for each sequence type. This analysis resulted in 42 motifs with an E-value<1E-5, which were used by FIMO (Bailey et al., 2009) to retrieve all locations of the motif of all the transcripts in the background (all transcripts present on the microarray). This information was used to calculate the average amount of times a motif is present at a specific position (relative number of motifs) for five distinct regions for each gene set and the background. These regions were at the 5' end of the transcript and downstream, from the start codon upstream, in the middle of the Coding DNA Sequence (CDS), from the stop codon downstream end and from the 3'end upstream, each region being 50 nucleotides in length (as illustrated in supplementary Figure 4). These results were used to create graphs and calculated enrichment or depletion of the motif at certain positions for each gene set including p-values and fold changes (supplementary Figure 4) using a custom script (python) and the R software package (http://www.r-project.org/). Motifs with a P-value < 1E-10 at any position were considered significant, which resulted in 10 motifs which could be subdivided into four groups (Figure 6/table 1 and supplementary Figure 4). All gene sets and positions with a *P*-value<1E-7 are shown in Figure 6/table 1.

Ribosomal protein extraction, tryptic digestion and peptide purification

The ribosome fractionation was performed as described above. The fractions corresponding to monosome and polysome regions were separately pooled and 1.1 ml of methanol was added to and 0.6 ml of each pooled ribosome extract and mixed thoroughly. After adding 0.3 ml of chloroform, vortexing and centrifugation for 15 minutes at 16,000 g, the supernatant was carefully removed with a capillary pipette. The air-dried protein pellet was dissolved in 100 µl of 8 M urea and an aliquot was taken for protein estimation using the Qubit® protein HS Assay Kit (ThermoFisher, Waltham, USA). Between 4-15 ug of protein dissolved in 100 ul of 8 M urea was used for digestion after reduction (5 mM DTT; 30 min at 37°C) and alkylation (15 mM iodoacetamide; 30 min at RT in the dark) of the cysteines. After adding 75 ul of 0.1 M ammonium bicarbonate buffer (pH8.3), 25 µl of Trypsin/Lys-C Mix (0.02 ug/µl, Promega, Madison, USA) was added for digestion (3 hours at 37°C). After dilution with 200 ul of 0.1 M ammonium bicarbonate, 4 ul of modified trypsin (0.05 µg/µl, Promega, Madison, USA) was added for overnight incubation at 37°C. Digestion was terminated by adding 0.1% trifluoroacetic acid (TFA). The tryptic digests were cleaned by reverse phase SPE (solid phase extraction) on an Oasis HLB Elution Plate (Water). Peptides were eluted with 0.1 ml of 50% acetonitrile (ACN), 0.1% TFA and dried by vacuum centrifugation. For 2D-LC-MS analysis the peptides were dissolved in 20 µl of 0.1 M ammonium formate (pH 10).

Mass spectrometry measurement

For high resolution separation of the ribosomal protein digests, a nanoAcquity 2-D UPLC system (Waters Corporation, Manchester, UK) was used employing orthogonal reverse phase separation at high and low pH, respectively. With this 2-D set up, the pool of peptides was eluted from the first dimension XBridge C18 trap column (in 20 mM amononium formate pH10) in two steps of 20% and 65% ACN. For the second dimension we used a BEH C18 column (75 µm ×25 cm, water UK) eluting with a 100 minutes linear gradient from 3 to 40% ACN (in 0.1% FA) at 200

nl/min. The eluting peptides were on-line injected into a Q-Exactive Plus (Thermo Scientific) mass spectrometer using a nano-electrospray source. Ionisation (2.4 kV) was performed using a stainless steel emitter and a heated capillary temperature of 250˚C. Full MS scans were acquired over the m/z range 400-1500 with a mass resolution of 70 000 (at m/z 200). Full scan target was set $3x10^6$ with a maximum fill time of 50 ms. The 10 most intense peaks with charge state 2-4 were fragmented in the HCD collision cell with a normalized collision energy of 28%. The mass range was set to 140-2000, and a mass resolution of 17500 (at m/z 200). The target value for fragment scans was set at $5x10^4$, the intensity threshold was kept at $1x10^4$ and the maximum allowed accumulation times were 100 ms. Peptide match was set to preferred, isolation window at 1.6 m/z and isotope exclusion was on. The dynamic exclusion was set to 30 s.

Database search for protein identification and quantitative analysis

The raw Q_{EX} LC-MS/MS data were processed with the work flow described for plant proteomics using the Maxquant (Cox & Mann, 2008; Smaczniak et al., 2012) and the resulting masses file containing all the fragment information was matched against the an in house TAIR protein sequence database based on The Arabidopsis Information Resource (TAIR10, www.Arabidopsis.org). The proteinGroups.txt file generated from Maxquant was used for quantitative analysis. Intensity determination and normalization was performed by algorithm MaxLFQ (Cox et al., 2014). The LFQ value generated was used for further statistics analysis. A protein was identified as present if LFQ is above background value for all the three biological replicates and followed for further statistical analysis. Proteins identified in either monosome fraction or polyosome fractions were assigned as proteins with specific ribosome association. Proteins present in both fractions were first Log2 transformed followed by differential analysis with LIMMA package to identify proteins with different ribosome associations. Two fold difference with corrected *P*-value<0.05 (Benjamini-Hochberg) were used as criteria for differential presence. A monosome and polysome specifically associated protein was assigned if it was identified by

either one of above two methods.

Gene and protein function classification and over-representation analysis

DAVID Bioinformatics Resources 6.7 were used for gene functional annotation and GO enrichment analysis (Huang et al., 2009). Panther classification system was used for ribosome protein classification and statistical overrepresentation test (Mi et al., 2013). Panther protein class was used as reference list which contains 26,684 proteins in the database. Bonferroni corrected protein list was generated with P-value<0.05

Acknowledgements

This work was supported by the Netherlands Organization for Scientific Research (project number 821.02.010). Financial support came from the Marie Curie ITN project MERIT, Grant Agreement number 264474 (AP), ZonMW project, project number 435002002 and we thank Bio4Energy, a Strategic Research Environment appointed by the Swedish government, for supporting this work. We thank Service XS, Leiden, The Netherlands for performing the microarray hybridizations. We thank Berend Snel and Alessia Peviani of the Utrecht University for their suggestion in the setup of the motif analysis and Bas Dekkers for valuable suggestion on the manuscript

Supporting Information

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

Figure S1. Sequence features of transcripts associated with the monosome (M) or polysome (P) fractions in seeds and seedlings.

Figure S2. Total RNA and ribosome associated protein content.

Figure S3. Ribosome proteome analysis on seeds.

Figure. S4. Gene 1.1 ST Genechip quality assessment and reproducibility.

Supplemental Table S1. GO enrichment analysis of monosome specifc transcripts in seeds. GO-terms of polysome specifc transcripts in DS are listed according to their p-value in each list.

Supplemental Table S2. GO enrichment analysis of polysome specifc transcripts in seeds. GO-terms of polysome specifc transcripts in seeds are listed according to their p-value in each list.

Supplemental Table S3. GO enrichment analysis of monosome specifc transcripts in seedlings. GO-terms of polysome specifc transcripts in seeds are listed according to their *P*-value in each list.

Supplemental Table S4. GO enrichment analysis of polysome specifc transcripts in seedlings. GO-terms of polysome specifc transcripts in seeds are listed according to their p-value in each list.

Supplemental Table S5. mRNA half-life and decay rate of monosome and polysome specifc transcripts in seeds and seedlings.

Supplemental Table S6. Ribosome proteome identified for the seeds. In total, 1398 proteins are identified in either monosome or polysome fractions.

Supplemental Table S7. Functional categorization of total seed riboproteome and monosome and polysome associated riboproteome in seeds. Significant GO-terms of monosome and polysome enriched ribosome proteins in seeds (P-value<0.05).

Supplemental Table S8. Total ribosomal proteins (RPs) in seeds.

Supplemental Table S9. Total RNA binding proteins (RBPs) in seeds.

Supplemental Table S10. Total polysome and monosome specifc ribosome proteins in seeds.

Supplemental Table S11. Polysome and monosome specifc ribosome proteins in seeds. Top 3 significan protein classes in each ribosome association state are listed (P-value<0.05).

Supplemental Table S12. Total RNA binding proteins (RBPs) excluding ribosomal protein in seeds.

Supplemental Table S13. Differential localized RNA binding proteins (RBPs) excluding ribosomal protein in seeds.

Supplemental Table S14. Functional categorization of differential localized RNA binding proteins (RBPs) in seeds.

Chapter 4

Combined transcriptome and translatome analyses reveal a role for transcriptional inhibition of the tryptophan dependent auxin biosynthesis pathway in the control of seed dormancy

Bing Bai[1,2], Johannes Hanson[1,3], Leónie Bentsink[2]

[1] Department of Molecular Plant Physiology, Utrecht University, 3584 CH Utrecht, The Netherlands

[2] Wageningen Seed Laboratory, Laboratory of Plant Physiology, Wageningen University, 6708 PB Wageningen, The Netherlands

[3] Umeå Plant Science Center, Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden

Abstract

We investigated the role of transcriptional and translational regulation during seed imbibition in dormant and non-dormant (after-ripened) seeds of the very dormant near isogenic line *DELAY OF GERMINATION 1*. Ribosomal profiling was performed on dormant and non-dormant seeds imbibed for 6 and 24 hours. Transcriptome and translatome changes were investigated and the transcription inhibitor cordycepin was used to study the contribution of transcription. Ribosomal profiles of non-dormant seeds imbibed in cordycepin mimic those of dormant seeds. The translational efficiency of mRNA species is not affected by seed dormancy. Dormancy during imbibition is mainly transcriptionally regulated by the tryptophan-dependent auxin and indole glucosinolate biosynthesis pathways.

Keywords: transcription, translatome profiling, ribosome, seed dormancy, seed germination, *Arabidopsis thaliana*

Introduction

Seed dormancy, defined as the inability of a viable seed to germinate under optimal conditions, is an important adaptive trait for plants to survive in nature (Bewley, 1997). Seed dormancy is a complex trait for which substantial natural genetic variation is present in *Arabidopsis thaliana*. *DELAY OF GERMINATION 1* (*DOG1*), is one of the most dominant quantitative trait loci underlying this natural genetic variation (Alonso-Blanco et al., 2003; Bentsink et al., 2006; Bentsink et al., 2010; Huang et al., 2010). The gene encodes a protein with unknown but conserved function across plant species (Graeber et al., 2014). The strong dormancy phenotype of the Cape Verde Islands (Cvi) allele of *DOG1* provides an ideal model system to investigate molecular pathways regulating seed dormancy.

Seed dormancy can be relieved by after-ripening, which refers to a period of seed dry storage after seed harvest. During this period, gene transcription and metabolism are limited due to the low moisture content and small nuclear size due to high chromatin condensation (Fait et al., 2006; van Zanten et al., 2011; Gao et al., 2013; Meimoun et al., 2014). However, differences in the level of individual proteins have been reported between dormant and after-ripened dry seeds (Chibani et al., 2006). Post-transcriptional regulation is proposed to regulate seed dormancy release possibly through mRNA oxidation (Bazin et al., 2011; El-Maarouf-Bouteau et al., 2013), which inhibits protein translation in vitro (Bazin et al., 2011). In the dry stage, dormant and after-ripened seeds hardly show any difference in transcriptional patterns, however transcriptional changes become visible when the seeds are imbibed (Cadman et al., 2006; Finch-Savage et al., 2007). Recently, both a role for transcriptional and post-transcriptional regulation in the control of germination have been proposed for sunflower and Arabidopsis (Layat et al., 2014; Basbouss-Serhal et al., 2015; Dekkers et al., 2015). Here, we investigated the role of transcription and translation during the imbibition of dormant (D) and non-dormant, after-ripened (AR) seeds, making use of ribosome profiling and transcription inhibitors. Ribosome profiling uses a sucrose-gradient based fractionation method for separation of mRNAs based on their association to polysomes and thus identifies mRNAs that are being translated.

These mRNAs can be identified by high throughput profiling techniques such as microarray analysis and RNA-sequencing (Mustroph et al., 2009; Layat et al., 2014; Lin et al., 2014; Vragovic et al., 2015). The ratio between polysomal bound mRNA and total mRNA of a specific mRNA represents the translational efficiency (TE) of that mRNA. (Arava et al., 2003; Halbeisen & Gerber, 2009; Ingolia et al., 2009; Jiao & Meyerowitz, 2010; Liu et al., 2012; Liu, MJ et al., 2013; Juntawong et al., 2014; Krishnan et al., 2014; de Klerk et al., 2015). We show that, upon imbibition, after-ripened and dormant seeds can be discriminated by their transcriptional patterns. Moreover, we propose a role for the tryptophan dependent auxin and indole glucosinolate pathway in the regulation of germination

Materials and methods

Plant material and seed germination conditions

The *Arabidopsis thaliana* near isogenic line carrying the Cvi introgression of *DOG1* fragment in a Landsberg *erecta* genetic background (NIL*DOG17-1*; here referred to as NIL*DOG1*) was originally introduced by (Alonso-Blanco et al., 2003) and was used in the current study for its high dormancy behavior. NIL*DOG1* plants were grown in three biological blocks and seeds were harvested at maturity. Half of the freshly harvested dormant seeds were stored directly in -80°C to retain dormancy and the other half were after-ripened at ambient conditions (20-25°C and 40-60% RH) for dormancy release. Seed germination was scored following the after-ripening until 100% of germination was reached. Germination experiments were performed as described previously (Joosen et al., 2010). 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 (143 μmol m^{-2} s^{-1}). Pictures were taken once a day for a period of 6 days using a Nikon D80 camera fixed to a repro stand with a 60 mm macro objective. The camera was connected to a computer with Nikon Camera Control Pro software version 2.0. Clustering of seeds was prevented as much as possible. Germination was scored using the Germinator package (Joosen *et al.*, 2010). The effect of transcription inhibition on seed germination was tested by germinating the completely non-dormant seeds on different dosages (0.1, 1, 10, 100 and 1000μM) of the transcription inhibitors cordycepin and α-amanitin. (Sigma, USA) in petri dish. Seed germination and seedling establishment were evaluated every day. For the ribosome isolation, both dormant and non-dormant seeds were spread on wetted papers filled with either 1mM cordecepin or water to ensure homogeneous spacing. The petri dishes were wrapped with parafilm (M, Hach, USA) to prevent water loss during seed imbibition. The experiment was carried out in a 22°C incubator under continuous light (143 μmol m^{-2} s^{-1}). Seeds were sampled at 6 and 24 hours after imbibition (HAI). The harvested tissue was frozen in liquid nitrogen followed by freeze-drying. The dry material was stored at -80°C for further

analyses.

Isolation of total RNA and polysomal RNA and polysome analysis

For the isolation of polysomal RNA, about 400 mg of freeze-dried tissue was extracted with 8 ml of polysome extraction buffer, PEB (400 mM Tris, pH 9.0, 200 mM KCl, 35 mM $MgCl_2$, 5 mM EGTA 50 µg/mL Cycloheximide, 50 µg/mL Chloramphenicol) modified from (Subramanian, 1978). The extracts were loaded on top of a sucrose cushion (1.75 M sucrose in PEB) and centrifuged (18h, 90,000 g) using a Beckman Ti70 rotor for 18 h (Beckman Coulter, Brea, USA). The resulting pellet was resuspended in wash buffer (200 mM Tris, pH 9.0, 200 mM KCl, 0.025 M EGTA, 35 mM $MgCl_2$, 5 mM DTT, 50 µg/mL Cycloheximide, 50 µg/mL Chloramphenicol) and loaded on a 20-60% linear sucrose gradient, centrifuged at 190,000 g for 1.5 h at 4°C using Beckman SW55 rotor (Beckman Coulter). After ultracentrifugation, the gradients were fractionated into 20 fractions using a Teledyne Isco Density Gradient Fractionation System (Teledyne Isco. Lincoln, USA) with online spectrophotometric detection of (254 nm). The polysomal fractions were pooled for polysomal RNA isolation. The ribosome abundance is reflected by the area under the curve and was calculated after subtracting the baseline obtained by measuring a blank gradient and normalizing to total area under the curve to account for possible uneven loading of the gradients.

Microarray hybridization and data analysis

Affymetrix Arabidopsis Gene 1.1 ST Arrays (Affymetrix, Santa Clara, USA) were hybridized using the GeneChip® 3› IVT Express kit (cat. # 901229) according to instructions from the manufacturer. Hybridization data were analyzed and gene specific signal intensities were computed using the R statistical programming environment (www.R-project.com) and the BioConductor package Affy (Gautier et al., 2004) and the Brainarray cdf file ver. 17.1.0 (http://brainarray.mbni.med.umich.edu/). We display some familiar features of the perfect match and mismatch probe (PM and MM. DNA microarray data are available in the Gene Expression Omnibus

(GEO) repository (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE75368. The limma and affy package were used for RMA normalization (Irizarry et al., 2003) Probe set intensity signals that never exceeded the noise threshold (logExprs<4 in all samples) were removed. A linear model and Empirical Bayes methods were applied for assessing differential expression (Smyth, 2005; Diboun et al., 2006). DAVID Bioinformatics Resources 6.7 (https://david.ncifcrf.gov/home.jsp) were used for gene functional annotation and GO enrichment analysis.

Results and discussion

Translatome profiling reveals translational inhibition in dormant seeds.

Seeds of NIL*DOG1*-Cvi (Cape Verde Islands) were used to investigate the importance of translation during the imbibition of dormant seeds. The germination capacity of NIL*DOG1* seeds was followed during after-ripening. Freshly harvested seeds do not germinate, but dry storage (after-ripening) releases this dormancy and germination frequency gradually increases until it reaches 100% after 120 days (Figure 1A). The freshly harvested dormant (D) and fully after-ripened non-dormant seeds (ND), that present fully contrasting dormancy levels were used to determine the role of translation in dormant imbibed seed, using a translatomics approach. For that, ribosomal profiles were run on dormant and after-ripened, non-dormant, seeds at 6 and 24 hours after the start of imbibition (HAI). At these time points seeds do not show visible germination yet (Figure 1B). The ribosomal profiles reveal no differences between dormant and non-dormant seeds at 6 HAI, however at 24 HAI non-dormant seeds show a larger proportion of polysomes than dormant seeds, indicating non-dormant seeds are actively translating mRNAs (Figure 2C, D).

Inhibition of germination in dormant seeds is controlled by transcriptional regulation

We determined whether the difference in ribosomal profiles between dormant and non-dormant seeds at 6 and 24 HAI is regulated at the transcriptional or post-

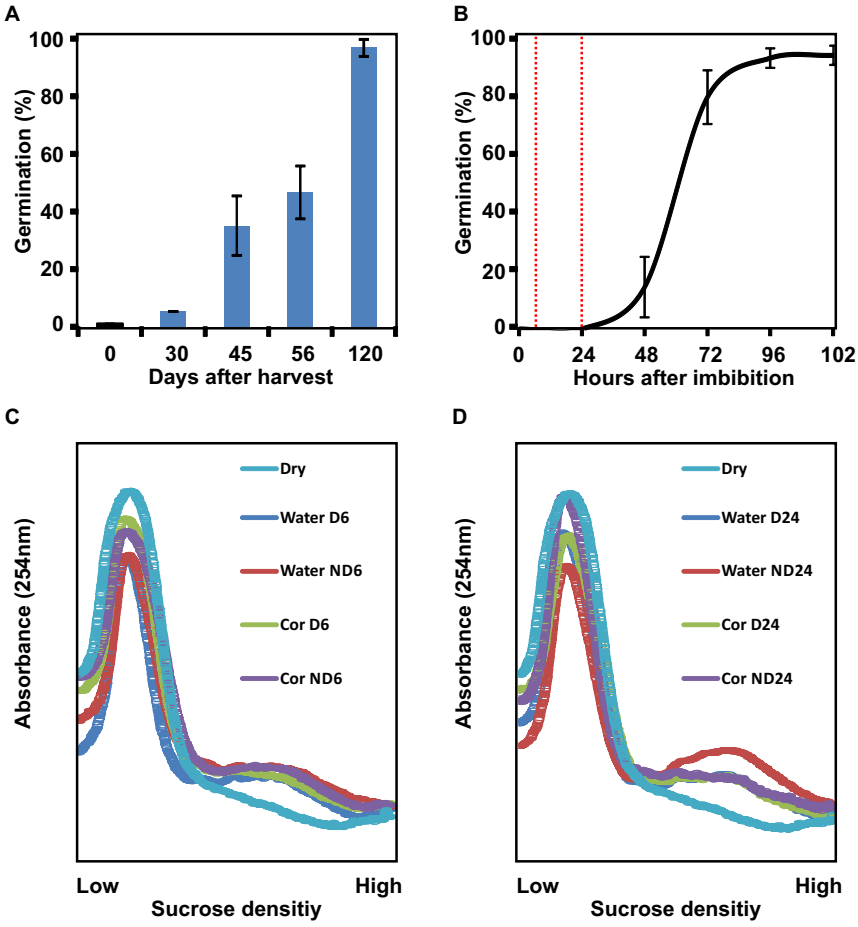


Figure 1. Germination behavior of NIL*DOG1* seeds following after-ripening. (A) Seed maximum germination determined after harvest at different times following after-ripening. (B) Germination dynamics of seeds after dormancy release. Red dashed lines represent the two imbibition stages at 6 and 24 hours imbibition (HAI) at which the transcriptome and translatome are compared. (C) Ribosome profiles of Dormant (D) and Non-dormant (ND) seeds after 6 hours imbibition in both water and 1 mM cordycepin (Cor) (based on its inhibitory effect on seed germination as shown in Figure 3). (D) Ribosome profiles of D and ND seeds after 24 HAI in both water and 1 mM cordycepin (Cor).

transcriptional level. Therefore, total and polysomal associated mRNAs of dormant and non-dormant seeds were compared at two imbibition time-points. At 6 HAI, 304 and 300 genes were up-regulated and 315 and 258 genes down-regulated in dormant compared to non-dormant seed in the total and polysomal mRNA fractions, respectively. At 24 HAI, numbers increased to 991 and 1136 up-regulated and

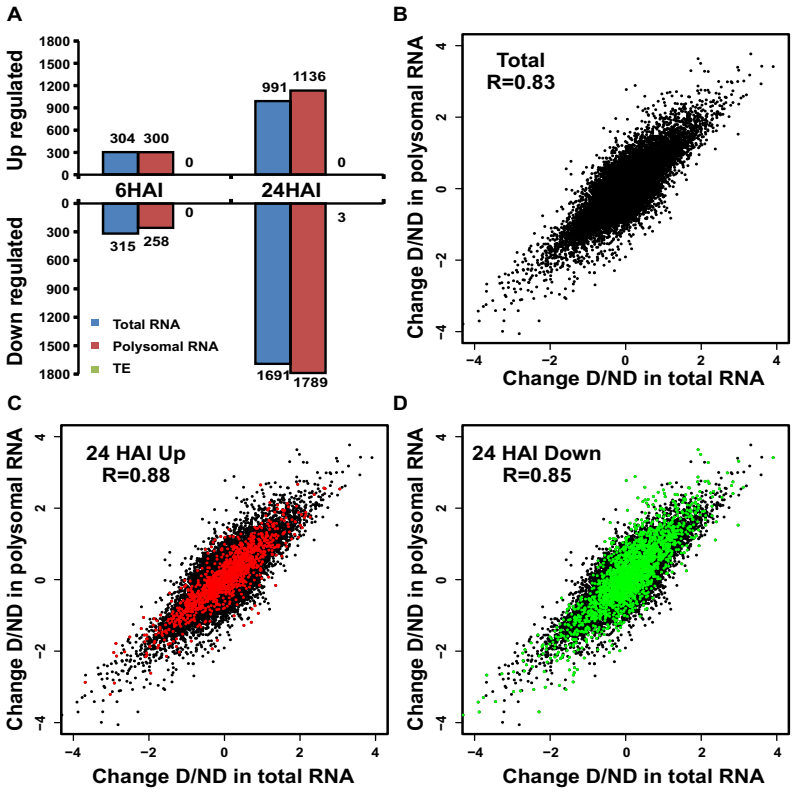


Figure 2. The influence of dormancy level on gene transcription and translation. (A) The effect of dormancy on total RNA abundance, polysomal RNA abundance and translational efficiency (TE=polysomal RNA abundance/total mRNA abundance) at two stages of imbibition. Up-regulated genes are genes that are higher in dormant versus non-dormant seed and down-regulated genes are higher in non-dormant seeds. (B) The correlation between changes of an individual RNA species in the total RNA in D compared to ND seeds (Log2D/ND) and changes of that same RNA species in the polysomal RNA in D compared to ND seeds (Log2D/ND) (C) The correlation between total RNA changes (Log2D/ND) and polysomal RNA changes (Log2D/ND) of genes up-regulated in D seeds compared with ND seeds at 24 HAI. (D) The correlation between total RNA changes (Log2D/ND) and polysomal RNA changes (Log2D/ND) of genes down-regulated in D seeds compared to ND seeds at 24 HAI. Correlation coefficients are shown for each correlation. All the correlations are highly significance (*P* -value< 2.2E-16). The red and green dots indicate the up- and down-regulated genes at 24 HAI in D seeds compared to ND seeds, respectively.

1691 and 1789 down-regulated genes in total and polysomal fractions, respectively (Figure 3A). To addressed whether translational regulation is important for the imbibition of dormant seeds, the TE was analyzed. Our analysis showed that the dormancy state hardly affected the TE of the individual mRNAs (Figure 2A). This

is confirmed by the highly significant correlation between the change of a specific mRNA species in the total mRNA in dormant versus non-dormant seeds and the change of the same mRNA in the polysomal mRNA (Figure 2B). This correlation is maintained if only the differentially up- or down-regulated genes at 24 HAI are considered (Figure 2C, D). Overall, these analyses show that the observed increased translation in non-dormant (Figure 1D) seeds depends on transcription and that the impact of translational regulation is relatively minor. This conclusion contrasts with recently published work (Basbouss-Serhal et al., 2015), where differences in translational regulation were found during imbibition of after-ripened and dormant Columbia seeds. The discrepancy observed may be caused by the different experimental system used, we define dormancy release based on after-ripening time and use fully discriminative stages (zero versus 100 percent germination), while Basbouss-Serhal et al (2015) studied temperature dependent germination that leads to a germination difference of 40% between the two dormancy stages. In addition, we imbibed the seeds in light while the seeds in the study of Basbouss-Serhal et al (2015) were imbibed in darkness. Light signaling has previous been shown to greatly impact translation in seedlings (Juntawong & Bailey-Serres, 2012; Gamm et al., 2014; Missra et al., 2015). This apparent contradiction is possibly related to the importance of light dependent mechanism in after-ripening of seeds (Leon & Owen, 2003; Hofmann, 2014) or other differences in experimental conditions and further illustrates the complex multifactorial regulation of seed germination (Bentsink & Koornneef, 2008; Rajjou et al., 2012).

Role of transcription during the imbibition of dormant seed

Transcriptional regulation in the control of seed dormancy is important and was further investigated using transcriptional inhibitors. For this, two transcriptional inhibitors were used, α-amanitin and cordycepin. α-amanitin was reported to inhibit seedling establishment but not germination, whereas cordycepin fully inhibited germination (Rajjou et al., 2004). In our study, both transcriptional inhibitors had a dosage dependent effect on seed germination and seedling growth (Figure 3A, B), however, 1

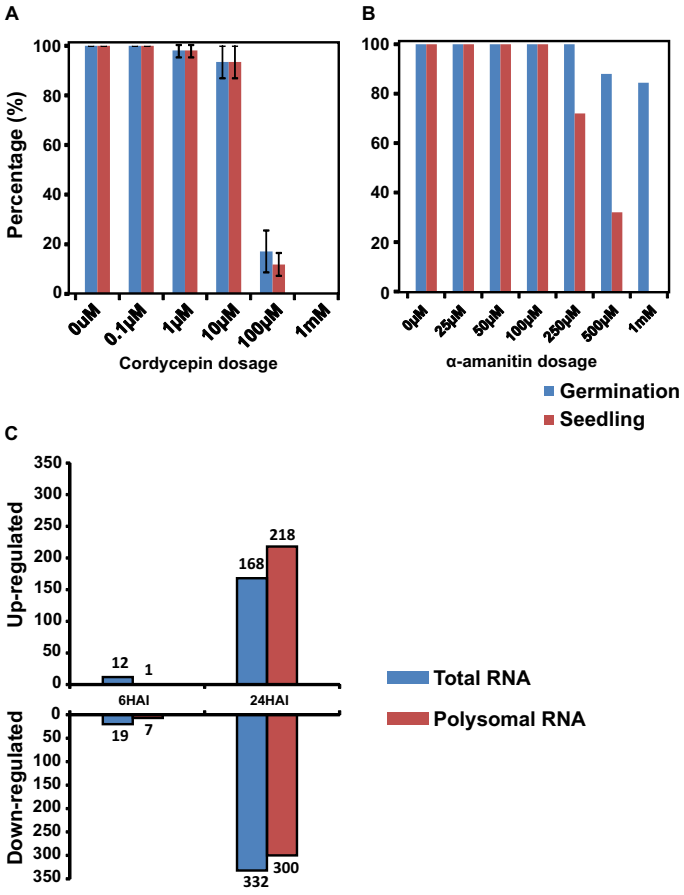


Figure 3. Cordycepin effect on transcription during two stages of seed imbibition. (A) Dosage dependent germination inhibition by cordycepin. (B) Dosage dependent germination inhibition by α-amanitin. (C) Bar graph represents the number of transcripts influenced by cordycepin in ND seeds at 6 and 24 HAI respectively. Transcripts differentially affected by cordycepin in total RNA and polysomal RNA.

mM cordycepin completely abolished seed germination and seedling establishment. The difference in inhibitory effect between α-amanitin and cordycepin might be due to differential uptake of these compounds by seeds. Cordycepin (1 mM) was used to inhibit transcription because it effectively blocked seed germination. Interestingly, the ribosomal profiles of the cordycepin treated seeds mimic those of the dormant seeds (Figure 1C, D). Next, the effect of cordycepin on the total and polysomal associated mRNAs was investigated. The number of differentially expressed genes

between 6 HAI imbibition in cordycepin and in water is relative low, 12 and 1 genes are up-regulated and 19 and 7 genes are down-regulated in cordycepin compared to water in the total and polysomal associated mRNAs, respectively (Figure 3). This indicates that cordycepin hardly affects transcription at 6 HAI, either because it has not (yet) penetrated the seed or that cordycepin is not very effective in blocking transcription at early seed imbibition. At 24 HAI these differences have increased to 168 and 218 up-regulated genes and 332 and 300 down-regulated genes in the total and polysomal associated mRNAs, respectively. Interestingly, up-regulationof genes is observed suggesting that there cordycepin does not completely block transcription. Overall, the minor effect of the inhibition of transcription at 6 HAI suggests that transcriptional activity at early time points does not determine to whether seeds are able to germinate or not.

Role of the tryptophan dependent auxin biosynthesis pathway in germination inhibition

In the presence of the transcriptional inhibitor cordycepin, dormant and non-dormant seeds do not germinate. Since germination inhibition was regulated at the transcriptional level for dormant and non-dormant seeds, the overlap in the respective transcriptomes was investigated. In this investigation, the changes over time were taken into account. Time dependent changes from 6 to 24 HAI in dormant, non-germinating seeds were compared with the same time dependent changes in non-dormant, germinating seeds. 899 and 1,266 genes were affected positively and negatively affected by dormancy level, respectively (Figure 4A). GO analysis revealed that these genes are related to abiotic stress responses such as light, heat, oxidative stimulus and a series of metabolic processes (Table S2). Similarly, time dependent changes from 6 to 24 HAI in cordecypin treated seeds were compared with the same time dependent changes in non-dormant, germinating seeds. 93 and 124 genes were affected positively and negatively by cordycepin, respectively (Figure 4A). The positively affected GO classes are redox balancing, hormone signaling (e.g. ethylene, ABA and cytokinin) and hormone metabolism (*NINE-CIS-*

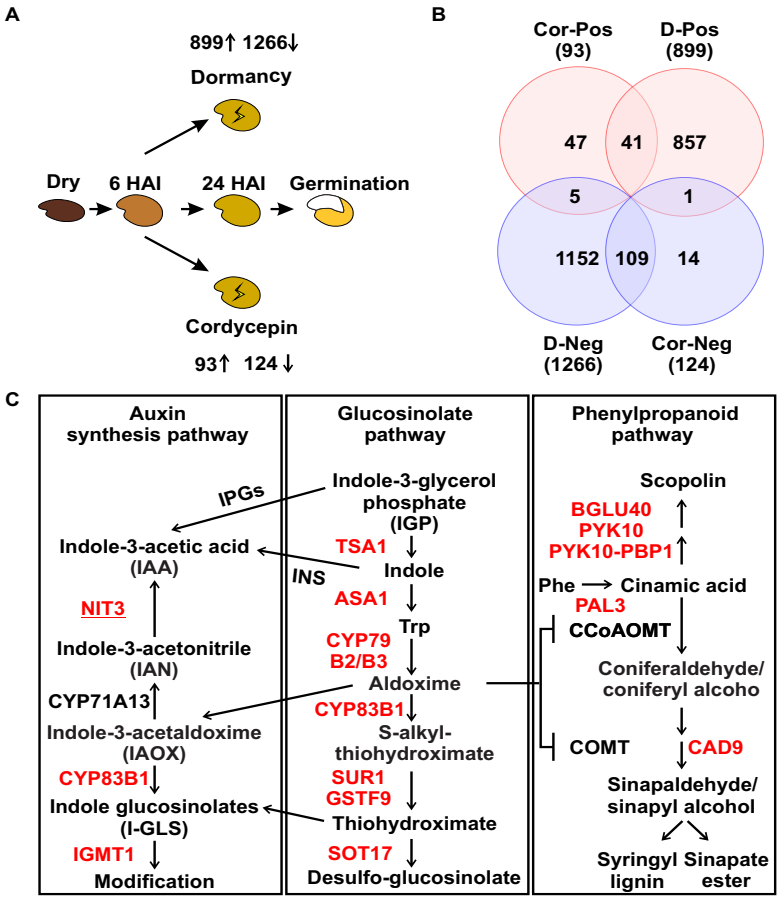


Figure 4. The effect of after ripening and cordycepin on the transcriptional changes from 6 to 24 HAI. (A) Dormancy and cordycepin are two independent factors which block seed germination. The upwards and downwards arrows indicate the genes positively and negatively influenced by dormancy or cordycepin in the 6-24 HAI comparison. (B) Venn-diagram comparing the genes influenced by dormancy and cordycepin in the 6-24 HAI comparison. Cor and D-pos and -neg indicate the genes positively and negatively influenced by cordycepin and dormancy respectively. (C) The enriched metabolic pathway for the genes influenced by both cordycepin and dormancy during seed imbibition. The genes presented in red are repressed by both cordycepin and dormancy and the underlined gene (NIT3) was repressed by cordycepin only. The genes in black encode enzymes catalyzing metabolic conversions that were not detected in the current study. GSTF9, glutathione S-transferase PHI9; SOT17, sulfotransferase17; SUR1, SUPERROOT1; IGMT1, indole glucosinolate o-methyltransferase 1; TSA1, tryptophan synthase alpha chain; ASA1, anthranilate synthase alpha subunit1; BGLU40, beta-glucosidase40; PYK10, beta-glucosidase23; CAD9, cinnamyl alcohol dehydrogenase9; PAL3, phenyl alanine ammonia-lyase 3; PBP1, PYK10-binding protein 1; NIT3, NITRILASE3; IPGs, indole-3-glycerol-phosphate synthases; INS, indole synthase; CCoAOMT, caffeoyl-CoA O-methyltransferase; COMT, caffeic acid O-methyltransferase; Trp, tryptophan; Phe, phenylalanine.

EPOXYCAROTENOID DIOXYGENASE5, 6 and *GIBBERELLIN 20-OXIDASE3*) while sulfur and indole derived biosynthetic process are negatively affected (Table S3). Approximately ten-fold more dormancy dependent differential gene expression was observed compared to the changes caused by transcriptionally inhibited seeds. However the overlap between the dormant and cordycepin-treated groups is significant (Figure 4B; 41 genes up and 109 genes down in non-germinating seeds). This large overlap suggests that seed dormancy is partly controlled by transcriptional regulation.

Interestingly, many of the overlapping genes are related to the tryptophan dependent auxin (indole-3-acetic acid, IAA) and the glucosinolate biosynthesis pathways (Figure 4C). In literature, this tryptophan dependent auxin biosynthesis pathway has been linked to the indole glucosinolate pathway (Zhao et al., 2002; Sugawara et al., 2009) and several genes related to both pathways are down regulated in the transcriptome of none germinating seeds (Figure 4; Table S5). This includes transcripts coding for enzymes involved in the conversion of indole-3-glycerol phosphate (IGP) to indole by TRYPTOPHAN SYNTHASE1 (TSA1), indole to tryptophan by ANTHRANILATE SYNTHASE ALPHA1 (ASA1), and the key glucosinolate biosynthesis related cytochrome P450 CYP79B1/B2, CYP83B1, SUPEROOT1 (SUR1), GLUTATHIONE S-TRANSFERASE9 (GSTF9), the glucosinolate synthesis enzyme SULFOTRANSFERASE17 (SOT17) and the glucosinolate modification enzyme INDOLE GLUCOSINOLATE O-METHYLTRANSFERASE1 (IGMT1). Further, the glucosinolate biosynthesis pathway likely interacts with the phenylpropanoid biosynthetic pathway through aldoxime, which has been reported to inhibit CAFFEIC ACID O-METHYLTRANSFERASE (COMT) and thus the production of phenylpropanoids such as sinapaldehyde and sinapyl alchohol (Hemm et al., 2003). Also enzymes in the phenylpropanoid pathway (PHENYL ALANINE AMMONIA-LYASE3 (PAL3), CINNAMYL ALCOHOL DEHYDROGENASE9 (CAD9), BETA GLUCOSIDASE40 (BGLU40), BGLU40 (PYK10) and PYK10-BINDING PROTEIN1 (PBP1) are down regulated in non-germinating seeds. These findings suggest that there is no or reduced auxin biosynthesis through the tryptophan dependent auxin

synthesis pathway in non-germinating (dormant and transcriptionally inhibited) seeds. A role for auxin in seed dormancy has earlier been proposed and auxin has shown to both positively and negatively affect seed germination often in cross-talk with the plant hormones ABA and GA (Brady et al., 2003; Liu et al., 2007; Liu, AH et al., 2013; Liu, X et al., 2013). Evidence for the dormancy increasing effect of auxin was mostly obtained by investigating tryptophan dependent auxin biosynthesis mutants such as *yuc1yuc6* and auxin signaling mutants, while the endogenous auxin level on seed dormancy was not investigated. Conclusive evidence on the effect of auxin on seed dormancy requires measurement of auxin levels, since auxin can also be synthesized by the tryptophan independent pathway, which, as recently revealed contributes to early embryogenesis (Jian et al., 2000; Wang et al., 2015). The tryptophan independent auxin synthesis pathway is conserved among different species (Wright et al., 1991; Normanly et al., 1993), but a role in regulating seed dormancy during seed imbibition has so far not been reported.

In addition, in none germinating seeds we find several cell cycle genes, including *CYCD1, CDCA2;3*, *DMC1*, *CMT3*, *MCM3*, *SEC14* (Table S5) to be negatively regulated. This might be a down-stream effect of the auxin pathway and would suggest reduced auxin levels since auxin is known to influence the cell division at both embryonic and vegetative stages (De Veylder et al., 2007). However, based on the current investigation, reduced expression of cell cycle genes could be the consequence instead of the cause of the germination inhibition.

Conclusion

Seed dormancy has been investigated for many years and major regulators of dormancy induction have been identified but limited attention has been paid to how germination is inhibited in dormant imbibed seeds. This manuscript reinforces the importance of transcriptional regulation in seed germination and shows that transcription is essential for successful seed germination and for polysome formation, and thus efficient translation. Auxin might be a key factor in differentiating dormant and non-dormant imbibed seeds, However, how the auxin pathway is activated

in non-dormant seeds is unknown. In-depth studies on the imbibition of seeds at different after-ripening stages might provide insight in this regulation.

Supporting information

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

Figure S1. Gene 1.1 ST Genechip quality assessment and reproducibility.

Figure S2. Cordycepin (a) and α-amanitin (b) effects on seed germination and seedling growth.

Supplemental Table S1. Gene ontology for the positive and negative interacted genes betwen cordycepin and dormancy level at 24 HAI. GO term of enriched biological process and KEGG pathways are shown.

Supplemental Table S2. Gene ontology for the trancriptional network affected by dormancy during imbibition from 6 to 24 HAI. GO term of enriched biological process and KEGG pathways are shown.

Supplemental Table S3. Gene ontology for the trancriptional network affected by cordycepin during imbibition from 6 to 24 HAI. GO term of enriched biological process and KEGG pathways are shown.

Supplemental Table S4. Gene ontology for the trancriptional network differentially affected by dormancy and cordycepin during imbibition from 6 to 24 HAI. GO term of enriched biological process and KEGG pathways are shown.

Supplemental Table S5. Gene ontology for the converged gene set affected by both dormancy and cordycepin during imbibition from 6 to 24 HAI. GO term of enriched biological process and KEGG pathways are shown.

Supplemental Table S6. Genes overlapped between dormancy and cordycepin effect on the transcriptional changes between 6 and 24 HAI. Gene postively and negatively affected by both factors are indicated together with gene descrptions

Chapter 5

Role for *ABCI20* in regulating seed maturation

Bing Bai[1, 2], Jan H.G. Cordewener[3], Antoine H.P. America[4, 5], Johannes Hanson[1, 6], Leónie Bentsink[2]

[1] Department of Molecular Plant Physiology, Utrecht University, 3584 CH Utrecht, The Netherlands

[2] Wageningen Seed Laboratory, Laboratory of Plant Physiology, Wageningen University, 6708 PB Wageningen, The Netherlands

[3] BU Bioscience, Plant Research International, P.O. Box 619, 6700 AP, Wageningen, The Netherlands

[4] Centre for BioSystems Genomics, P.O. Box 98, 6700 AB, Wageningen, The Netherlands

[5] Netherlands Proteomics Centre, P.O. Box 80082, 3508 TB, Utrecht, The Netherlands

[6] Umeå Plant Science Center, Department of Plant Physiology, Umeå University, SE-901 87 Umeå, Sweden

Abstract

Genetic variation for seed dormancy widely exists in *Arabidopsis thaliana* and previously the *Delay of Germination 6* (NIL*DOG6*) was identified in a quantitative trait loci (QTL) analysis. Here, we identified an ATP Binding Cassette family member gene *ABCI20* in a screen for NIL*DOG6* downstream targets. The *abci20* null mutant, has a strongly enhanced seed dormancy and is less sensitive to nitrate, GA and cold stratification for dormancy release. The fully after-ripened seed are more sensitive to ABA and show reduced seed longevity. Transcriptome analyses revealed reduced mRNA levels for genes associated with ABA- and ethylene signaling, autophagy, and oxidative stress reduction. The repression of key negative regulators for ABA signaling and positive regulators for ethylene signaling seem to explain the high ABA sensitivity and thus the high dormancy phenotype while the reduced expression of oxidative stress scavenger genes and four important autophagy related genes (*ATGs*) likely contribute to the low seed longevity. Ribosome profiling unveils an increased polysome in *abci20* dry seeds. Proteomic investigation of the *abci20* polysome fractions revealed the presence of non-ribosomal proteins involved in metabolic processes and proteasome related protein degradation. Possibly, during seed maturation proteolysis is affected in *abci20* resulting in denatured protein accumulation in the high density fractions. Interestingly, *abci20* is the first identified mutant that shows a negative correlation between seed dormancy and longevity and therefore, *abci20* provides a most useful tool for investigating the trade-off between these two important seed traits.

Keywords: *Arabidopsis thaliana*, ATP-binding cassette, translation, mRNA decay, seed dormancy, longevity, ribosome proteome

Introduction

Seed dormancy determines the timing of seed germination and is an important adaptive trait that prevents germination in suboptimal conditions. In *Arabidopsis thaliana*, seed dormancy is classified as physiologically non-deep, meaning that embryos released from surrounding structures grow normally, and that dormancy is lost through moist chilling (stratification) or after-ripening (Holdsworth et al., 2008). Seed dormancy is a complex trait affected by many different factors. In particular, abscisic acid (ABA) and gibberellins (GA) are important hormones in determining dormancy levels through their synthesis, catabolism and signaling (Bewley, 1997). Mutations affecting ABA synthesis, catabolism and signaling can significantly change seed dormancy levels, indicating the central role of ABA in regulating seed dormancy (Sondheimer et al., 1968; Koornneef et al., 1982; Karssen et al., 1983; Koornneef et al., 1984; Lin et al., 2007; Matakiadis et al., 2009; Liu et al., 2013; Shu et al., 2013). Nitrate and cold stratification are efficient factors to release seed dormancy in Arabidopsis by largely unknown mechanisms that are somehow linked to ABA and GA levels and signaling (Hilhorst et al., 1986; Jacobsen et al., 2002; Ali-Rachedi et al., 2004; Yamauchi et al., 2004; Alboresi et al., 2005).

Several studies have employed transcriptome analyses to study the release of seed dormancy both in the dry and imbibed state (Cadman et al., 2006, Finch-Savage et al., 2007, Carrera et al., 2008, Holdsworth et al., 2008 and Bentsink et al., 2010). These analyses revealed sets of differently expressed genes. Among these are key genes such as *DOG1* that, during seed maturation, directly determine the level of dormancy (Bentsink et al., 2006, Nakabayashi et al., 2012). Recently, it has been shown that regulation of seed dormancy also involves post-transcriptional processes. The REDUCED DORMANCY5 controls germination by suppressing the *PUMILIO9 (APUM9)* and *APUM11* encoding RNA binding proteins. It is proposed that the APUMs are involved in determining translation efficiency of specific stored mRNAs in seeds during imbibition (Xiang et al., 2014). Post-transcriptional regulation can affect seed dormancy levels by controlling mRNA stability, translation, as well as RNA and protein modifications such as oxidation (Li and Foley, 1996; Bazin et al.,

2011; El-Maarouf-Bouteau et al., 2013; Galland et al., 2014; Meimoun et al., 2014; Galland and Rajjou, 2015).

In this study, we investigated the expression profile of dormant and after-ripened NIL*DOG6* seeds during imbibition and identified an important downstream target gene regulating seed dormancy and seed longevity.

Results

abci20 mutant seeds are dormant, hyper sensitive to ABA and less storable

Delay of germination 6 (*DOG6*) is a major QTL affecting seed dormancy release. The near isogenic line (NIL) NIL*DOG6*-Cvi contains a Cvi introgression fragment at the *DOG6* position in the L*er* background and this line was used to study its down-stream targets (Alonso-Blanco et al., 2003). Transcriptome analyses on dormant and after-ripened imbibed NIL*DOG6* seeds revealed many translation related genes among the differentially expressed genes (Figure 1). Seed dormancy screening of T-DNA knock out mutants of these differentially expressed genes identified an extremely dormant phenotype for a T-DNA line with an insertion in the *ATP-BINDING CASSETTE I20* (*ABCI20*; AT5G02270) gene (Supplemental Figure 1). The T-DNA insertion in this gene (*abci20*) is located between the conserved Walk A and ABC sequence motifs (Figure 1A), resulting in undetectable *ABCI20* mRNA (Figure 1B). The *abci20* seeds required around 40 days to reach 50% germination (DSDS50; Figure 1C) while Col-0 only required 4-5 days to reach 50% germination. Cold stratification, gibberellin (GA) and nitrate treatments are routinely used procedures for dormancy breakage. Four days cold stratification for freshly harvested seeds, results in 94% germination for Col-0, while only 85% of the *abci20* seeds could germinate. When germinated in GA and nitrate, *abci20* still needed significantly longer after-ripening than Col-0 to release dormancy (Figure 1C). ABA levels in the *abci20* mutant did not differ from Col-0, in spite of the increased *abci20* dormancy level compared to Col-0 (Figure 1D). The ABA sensitivity of *abci20* was investigated and an increased ABA sensitivity in completely non-dormant seeds (after 1-year dry storage) was observed (Figure 1E). Interestingly, *abci20* also showed reduced germination after artificial ageing, which is a measure for seed storability (Figure 1E). The expression of *ABCI20* was quantified during the plant life cycle and two expression peaks were observed, one during seed maturation and one during seed germination (Supplemental Figure 2).

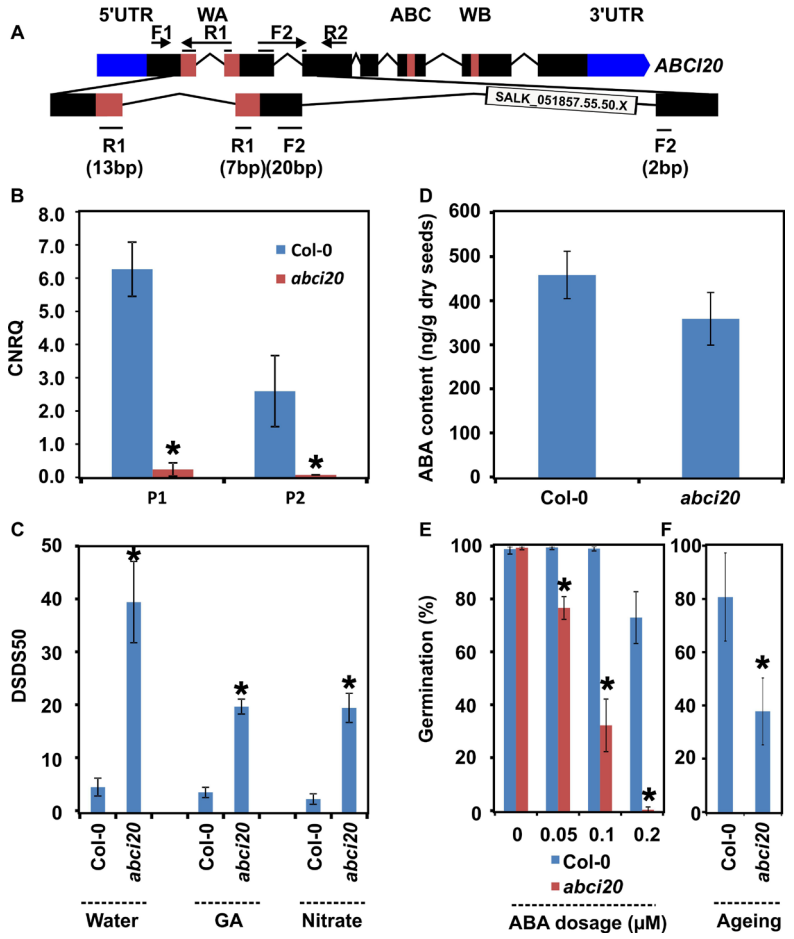
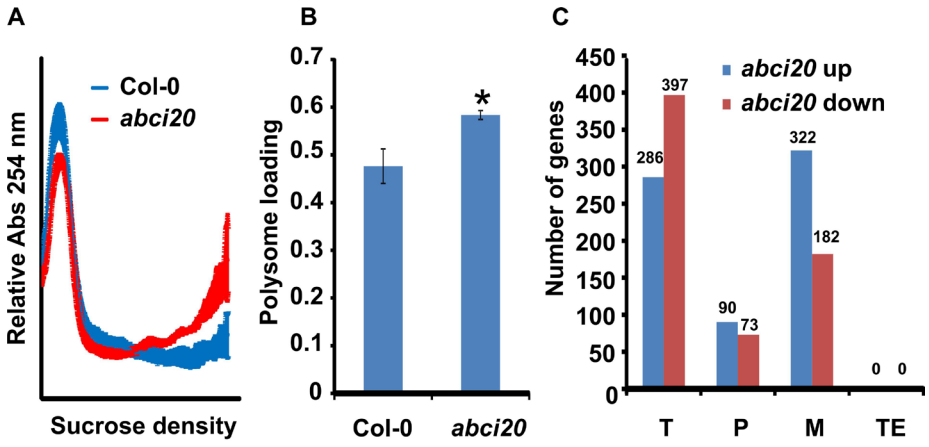


Figure 1. Characterization of the *abci20* mutant. (A) Gene model of *ABCI20*. The arrows indicated the region of two primer pairs used for transcript quantification. The box with SALK line ID indicates the T-DNA insertion site. (B) Quantification of *ABCI20* expression in wild type Columbia (Col-0) and T-DNA knockout mutant (*abci20*) with two primer pairs shown in A. (C) The dormancy levels of Col-0 and *abci20* as evaluated by the days of seed dry storage to reach 50% germination (DSDS50) in water, GA and nitrate. (D) Quantification of ABA content in Col-0 and *abci20*. (E) ABA sensitivity of *abci20* seeds and (F) the seed longevity of *abci20* seeds measured by after artificial ageing. Data represent means ± SD. *P*-value<0.05 as indicated by asterisks. *P*-values are calculated by a Student's t-test (n=3).

abci20 causes transcriptional reprograming

ABCI20 was among many genes involved in translation and differentially expressed in the NIL*DOG6* line. Translatome profiling was performed on dry seeds of the *abci20* mutant and Col-0 in order to investigate a potential role for *ABCI20*



D

Regulation	GO term	*P*-value	FC	Gene
abci20-up	Metabolic process	4.22E-05	5.3	*GAPC1; PGI1; UXS1*
	Stress response	5.13E-05	2.2	*HSP20/21/ 22/17.6/18.2*
abci20-down	Autophagy	2.73E-03	14.00	*ATG5/7/8H/8E*
	Ethylene response	7.54E-04	3.88	*EIN2; EBF2; EIL3/5*
	ABA response	5.61E-03	2.83	*HAI1/2; AFP1/3; PYL12; PP2C*
	Oxidation reduction	6.44E-04	1.92	*FDH; FER3; CAT1;GPX7*

Figure 2. Changed ribosome profile, transcriptome and translatome in *abci20*. (A) Mature dry seed ribosome profile of *abci20* and Col-0. (B) The polysome loading of *abci20* and Col-0 calculated based on the ratio between polysome area and total area under the curve. Data represent means ± SD. *P*-value<0.05 as indicated by asterisks. *P*-values are calculated by a Student's t-test (n=3). (C) The number of genes differentially expressed in total RNA (T), polysomal RNA (P) and monosomal RNA (M) pools in *abci20*. Translational efficiency (TE) is represented by the ratio between total mRNA and polysomal mRNA abundance. (D) The representative of the significant GO enrichment in the transcriptome of *abci20* with *P*-value for FDR, Fold Change (FC) and representative genes.

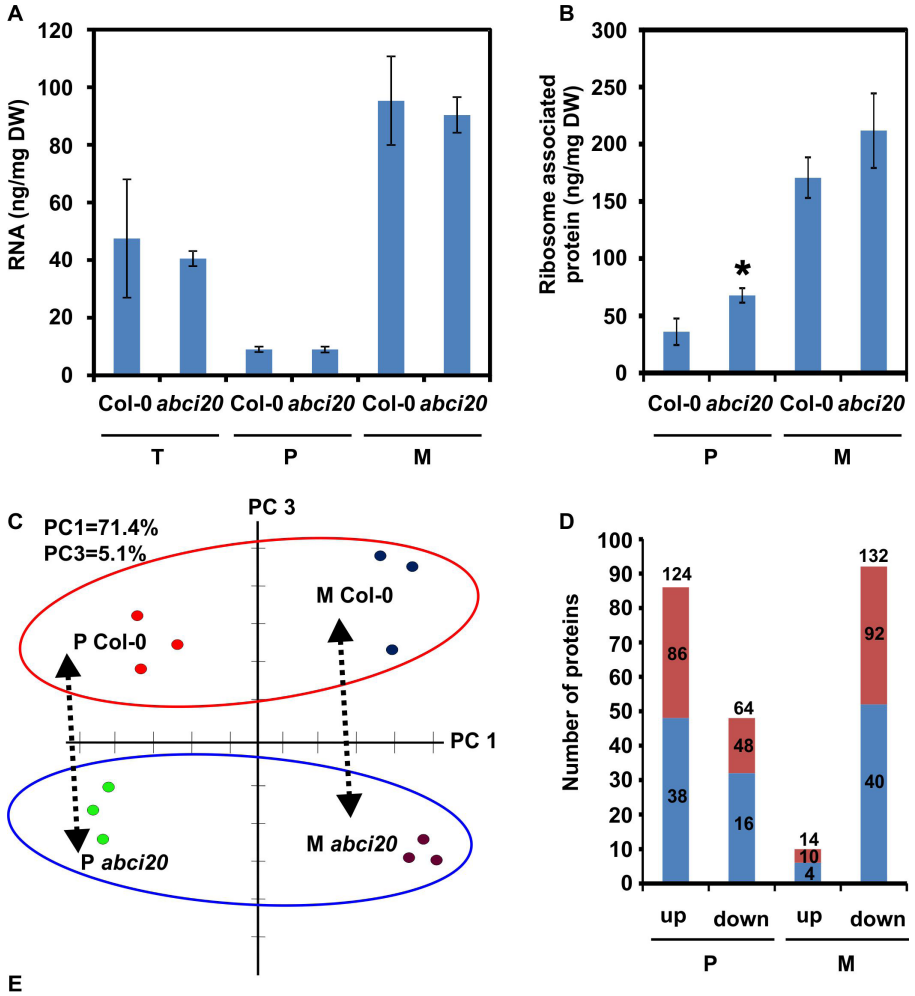
in protein translation. Surprisingly, in dry *abci20* seeds the polysome region is enhanced concurrent with a decrease in the monosome region (Figure 2A and B). This distinct ribosome profile of *abci20* compared to Col-0 was further investigated using microarray analysis of total mRNAs as well as monosomal and polysomal

mRNA pools. In this analysis, 286 and 397 genes were identified that were more than 2 fold up- and down-regulated, respectively, in the *abci20* compared to Col-0 (Figure 2C, Supplemental table 1). No differences in mRNA translational efficiency (TE; ratio between polysomal and total mRNA abundance) were observed in *abci20* compared to Col-0 (Figure 2C) and therefore the observed changes in the polysome profile resulted from transcriptional changes. The gene expression profile of *abci20* revealed that genes enriched in metabolic process and stress response (*HEAT SHOCK PROTEINS* (*HSPs*) including *HSP20*, *HSP21*, *HSP22.0*, *HSP17.6II* and *HSP18.2*) were up-regulated in *abci20*, as well as gibberellin (GA) metabolic genes (*GA2OX2*, *GA3OX3*, *GA20OX3*) and genes encoding key enzymes for gibberellins catabolism. Genes associated with ethylene and ABA signaling, autophagy and oxidative stress were down-regulated, including, *ETHYLENE INSENSITIVE2* (*EIN2*), down-stream positive regulators *ETHYLENE INSENSITIVE3-LIKE3* (*EIL3*), *ETHYLENE INSENSITIVE3-LIKE5* (*EIL5*) and *EIN3-BINDING F-BOX2* (*EBF2*) two ABA negative regulators: *HIGHLY ABA-INDUCED PP2C GENE1* and *2* (*HAI1* and *HAI2*), *SERINE/THREONINE PROTEIN PHOSPHATASE2A* (*PP2A*), *MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE19* (*MAPKKK19*), *REGULATORY COMPONENTS OF ABA RECEPTOR6* (*RCAR6*), *ABI5 binding protein* (*ABP1* and *ABP3*) and *ABA INSENSITIVE RING PROTEIN2* (*AIRP2*), oxidation reduction genes *CATALASE1* (*CAT1*), *ATYPICAL CYS HIS RICH THIOREDOXIN1* (*ACHT1*), *FERREDOXIN4* (*FD4*), *FERRIC REDUCTION OXIDASE2* (*FRO2*), *FERRITIN3* (*FER3*) and *GLUTATHIONE PEROXIDASE7* (GPX7), and autophagy genes *ATG5*, *ATG7*, *ATG8H* and *ATG8E* (Supplemental table 2). Also, ABA responsive element (ABRE)-like-binding motifs were highly enriched in ~25% of the promoters (*P*=8.62E-09) of down-regulated transcripts in *abci20*, and in the same set 17 and 22 transcripts were enriched in ABA binding factors (ABFs) binding site motif and ABRE binding site motif, respectively (Supplemental table 1). ABA is an important phytohormone mediating diverse responses during seed maturation, including the induction of seed dormancy and desiccation tolerance (Finkelstein et al., 2002; Verdier et al., 2013; Maia et al., 2014; Mao and Sun, 2015). We thus compared the genes regulated by *ABCI20* to the differentially regulated genes during

seed maturation from 14 days after flowering (DAF) until fully matured (Angelovici et al., 2009) and observed that ~37% of *abci20* down-regulated genes accumulated in wild type Wassilewskija (Ws) during seed maturation from 14 DAF until the end of maturation, while 26% of up-regulated genes in *abci20* decreased in Ws during the same maturation period (Supplemental Figure 3). This strongly indicates that *ABCI20* is involved in modulating seed maturation.

Large non-ribosomal complexes co-migrate with polysomes in *abci20*

To further investigate the aberrant polysomal profile in dry seeds of the *abci20* mutant (Figure 2A, B), the RNA content in the polysome and monosome fraction was quantified and this showed that there are no significant differences between the ribosome fractions (Figure 3A). However, analysis of the amounts of protein present in the corresponding fractions showed that there was a significantly higher amount of polysomal proteins in *abci20* (67.8 ng and 36.0 ng per mg dry seed, respectively, *P*-value<0.05) (Figure 3B) which likely contribute to the enhanced polysomal profile in *abci20*. High throughput liquid chromatography–mass spectrometry (LC-MS) was performed on both monosome and polysome fractions to identify proteins that contribute to the changes in the ribosome profile of *abci20*. In total 2,635 proteins were identified as present in at least one fraction of *abci20* and Col-0. PCA on the identified proteins clearly differentiated genotypes and the ribosome association form (Figure 3C). In each of the fractions unique proteins were identified (Figure 3D), however most interesting were the proteins that were specific for polysomes (Figure 3E, Supplemental table 4). These proteins were involved in diverse metabolic process including gluconeogenesis and glycolysis, pyruvate metabolism and myriads of stress responses related to oxidation and cadmium ions. Interestingly, components of the ubiquitin-proteasome system (UPS) were differentially enriched in *abci20*. Specifically enriched in the *abci20* polysome fraction are REGULATORY PARTICLE TRIPLE-A ATPASE3, 5B and 6A, SIGNALOSOME SUBUNIT4, PROTEASOME BETA SUBUNIT C1 and PROTEASOME BETA SUBUNIT PAA



E

Polysome	Regulation	GO term	*P*-value	FC	Protein
Polysome	*abci20*-up	Metabolic process	2.13E-04	8.03	ADG1; ENOC; ASD1
		Abiotic stimulus	3.05E-04	2.83	LEA2; HSP23.6; HD2C;ACO3
		Proteosome	1.77E-03	8.83	RPT6A; PBC1; RPT3;PAA2
	abci20-down	Ribosome biogenesis	4.22E-05	3.68	RLS26; PES; WD40; RPL34;
		Seed development	3.05E-04	4.77	LEA2; HSP23.6; HD2C;ACO3
Monosome	*abci20*-up	Metabolic process	0.07	24.13	G6PD6; TGG1
	abci20-down	Protein transport	3.47E-06	5.40	VSR4;TOM20; ERMO2;RAB7b/1c

Figure 3. Ribosome proteome of *abci20*. (A) RNA content in total seed extract (T) and polysome and monosome fraction in *abci20* and Col-0. (B) Protein content in total seed extract (T) and polysome and monosome fraction in *abci20* and Col-0. Data represent means ± SD. *P*-value<0.05 as indicated by asterisks. (C) PCA of the proteins affect in *abci20* and Col-0 in polysomal RNA (P) and monosomal RNA (M) pools. (D) The number of proteins differentially present in *abci20* in association with polysome and monosome. The blue bars indicate proteins identified in both genotypes in corresponding fractions but differentially enriched in each genotype and red bars indicate proteins specifically detected as present or absent in *abci20* in corresponding ribosome fractions. Total number of differentially enriched proteins in *abci20* is indicated. (E) The representative of the significant GO enrichment in the proteins differentially enriched in *abci20* in each ribosome association state with *P*-value for FDR, Fold Change (FC) and representative proteins.

2. In contrast, reduced abundance in *abci20* monosomes was observed of HAPLESS 15, UBIQUITIN-PROTEIN LIGASE 2, RPN3a, SEC31B and CUL4 RING ubiquitin ligase complex localized transducin WD-40 repeat family protein. Additionally, proteins reduced in *abci20* were associated with GO terms ribosome biogenesis and embryo development (Figure 3E, Supplemental table 4).

Discussion

The *DOG6* locus controls seed dormancy (Alonso-Blanco et al., 2003; Bentsink et al., 2010) and here target genes and processes were identified and studied. Transcriptional analyses on the NIL*DOG6* near isogenic line revealed that many genes related to translation are down regulated in NIL*DOG6* imbibed dormant compared to non-dormant seeds, suggestion that *DOG6* inhibits germination by inhibiting translation. In Col-0, *ABCI20* expression peaks during seed maturation and seed germination. In imbibed dormant NIL*DOG6* seeds *ABCI20* expression is reduced compared to that of non-dormant seeds. In *abci20* mutant seeds no *ABCI20* expression is observed and these seeds are highly dormant, ABA hypersensitive and badly storable, properties that are induced during seed maturation and maintained during seed germination. Therefore, a role for *ABCI20* in the control of late seed maturation is proposed.

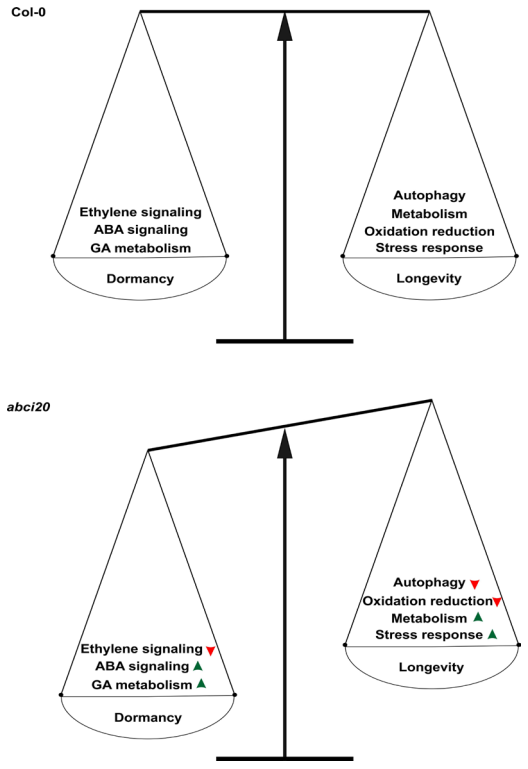


Figure 4. The trade-off between seed dormancy and longevity affected in *abci20*. Molecular processes associated with this trade-off changes in *abci20* are indicated as enhanced level of changes with a green arrow upwards and reduced level of changes with a red arrow downwards.

5

A role for *ABCI20* in the regulation of late seed maturation

The *abci20* mutant is severely affected in processes that are induced during seed maturation, particularly the high level of seed dormancy and the reduced seed longevity (Figure 4). Seed maturation has been studied in detail and it is know that seed dormancy in Arabidopsis is induced from 14 DAP onwards. Before 14 DAP seeds are capable to germinate, after that seeds require a dormancy breaking treatment such as a period of seed dry storage, cold stratification, GA addition or other dormancy breaking treatments (Jarvis et al., 1968; Arc et al., 2012; Hauvermale et al., 2015). Recently, it was shown for seeds of Medicago that seed longevity is only induced during the late seed maturation stage (Righetti et al., 2015). Seed dormancy and longevity are most likely linked processes but induced at different seed maturation stages. Apparently, there is a trade-off between seed dormancy and seed longevity, and QTLs controlling these traits are negatively correlated (Nguyen et al., 2012).

Dormant genotypes have reduced seed longevity and less dormant genotypes are better storable, and also environmental factors affect dormancy and longevity in a contrasting manner. Short days and low temperature increase seed dormancy and reduce seed longevity and in the other way around long days and higher temperatures decrease dormancy and increase longevity (He et al., 2014). How these processes are linked and controlled is unknown, but it was found that the *DOG1* gene regulates both processes (Nguyen et al., 2012). The *DOG*1-Cvi allele increases seed dormancy and decreases seed longevity as was shown by complementation analyses. It appears that the induction of high dormancy levels negatively affects seed longevity. This strong negative correlation is also observed in the *abci20* mutant. This makes *abci20* an interesting gene for studying the regulation of seed dormancy and longevity, in addition to *DOG1*. Detailed investigations of the *abci20* mutant lead to the remarkable observation of a different ribosomal profile compared to Col-0. Normally the dry seed ribosome profile shows a very clear monosome peak without any detectable signal in the polysome area, while the *abci20* mutant shows an increased signal in the polysome area (Figure 2A). This increase is not explained by an increase in RNA levels (Figure 3A) but by increased protein levels (Figure 3B). It is presently unclear, whether these dense protein complexes co-migrate in association with polysomes or whether that are of the same density as polysomes. Detailed proteomic analysis revealed that GO terms related to metabolism, abiotic stimuli and the proteasome are overrepresented among the proteins present in the polysome fraction of *abci20*. Normally, at the end of seed maturation, metabolic processes are inactivated, and the degradation of seed protein is initiated at late seed maturation (Li et al., 2007). Possibly, these processes are not completed in the *abci20* mutant as suggested by the large transcriptional changes in the *abci20* mutant. In the *abci20* several negative regulators of the ABA signaling cascade are down regulated, including *RCAR6*, *HAI1, HAI2*, *AFP1*, *AFP2*, *AIRP2* and *MAPKKK19* (Gosti et al., 1999; Garcia et al., 2008; Miura et al., 2009; Nishimura et al., 2010; Raghavendra et al., 2010; Cho et al., 2011; Bhaskara et al., 2012; Hu et al., 2014). The transcriptional changes strongly point to increased ABA signaling in *abci20*, which would explain the ABA hypersensitivity

and the high dormancy phenotype. Mutants of several of these genes show enhanced ABA sensitivity but these genes were so far not associated with increased dormancy. In addition to the effect of *ABCI20* on ABA signaling, also ethylene signaling and GA metabolism are affected and cross-talk between these hormones and ABA might explain the dormancy phenotype. The important ethylene signaling factors *EIN2* and *EILs* are down-regulated, likely leading to reduced ethylene signaling (Guo and Ecker, 2003). The key GA biosynthetic enzymes *GA20ox3* and *GA3ox3* and the GA catabolic gene *GA2ox2* (Zhu et al., 2006) are up-regulated in *abci20*. The suggested hormonal interplay might explain the aberrant seed dormancy and ABA sensitivity but it remains unclear how seed longevity is affected.

Transcriptional changes that are over-represented in *abci20* seeds and that might explain the longevity phenotype are repressed genes related to oxidation reduction and autophagy. Genes related to oxidation reduction are normally up-regulated during seed maturation but these levels are reduced in *abci20* and include *CAT1*, *ACHT1*, *FD4*, *FRO2*, *FER3* and *GPX7* (Figure 2D, Supplemental table 2). Reducing oxidative damage can extend seed lifespan, while increased oxidative damage generally shortens the lifespan, which is consistent with the free radical theory of ageing (FRTA) (Harman, 1956; Perez et al., 2009; Fontana et al., 2010). Other genes shown to be related to seed longevity and that are down regulated in *abci20* are autophagy genes (*ATG5*, *7*, *8H* and *8E*) (Figure 2D, Supplemental table 2). Autophagy recycles damaged or unwanted cell materials upon encountered stress conditions or during specific developmental processes (Liu and Bassham, 2012). Mutants of these *ATGs* give phenotypes related to accelerated senescence, hypersensitivity to starvation conditions and slower growth as reviewed by (Liu and Bassham, 2012). However, seed phenotypes have not been described for these mutants. In Arabidopsis *ABCI20* has two homologs, *ABCI19* and *ABCI21* but mutants in these genes do not have a distinguishable seed phenotype (data not shown). These *ABCI* genes are homologous to the yeast *Caf16* gene, which shares 27% similarity with *ABCI20*, and 32% and 35% similarity with *ABCI19* and *ABCI21,* respectively (Supplemental Figure 4A). In yeast Caf16 plays a role in mRNA stability by interacting with the

Ccr4-Not deadenylation complex. The Ccr4-Not core complex of 1×10^6 Dalton (1 MDa) consists of nine conserved subunits: Not 1, Not2, Not3, Not4/Mot2, Not 5, Caf130, Caf40, POP2/Caf1 and Ccr4 (Nasertorabi et al., 2011; Miller and Reese, 2012). The larger, less defined 1.9-MDa Ccr4-Not complex contains in addition Caf4 and Caf16 (Liu et al., 2001). We show here that *ABCI20* regulates the expression of genes that affect seed dormancy, longevity and ABA sensitivity. Whether *ABCI20* plays a role in mRNA stability like the yeast homologue requires further studies.

Materials and methods

Plant material and seed germination conditions

The expression analysis of NIL*DOG6* was based on microarray experiments (GEO GSE74331, Supplemental table 6). The *Arabidopsis thaliana* T-DNA mutants of genes differentially expressed between dormant and non-dormant imbibed NIL*DOG6* seeds were ordered from NASC and were used in the current study. Plants were grown in three biological blocks and seeds were harvested at maturity. Half of the freshly harvested dormant seeds were stored directly in -80°C to keep the dormancy and the rest were after-ripened at ambient conditions (20-25°C and 40-60% RH) dormancy release. Seed germination was scored following the after-ripening until 100% of germination was reached. Germination experiments were performed as described previously (Joosen et al., 2010). In brief, two layers of blue blotter germination paper (Anchorpaper company, St Paul) 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 (143 μmol m^{-2} s^{-1}). Pictures were taken once a day for a period of 6 days using a Nikon D80 camera fixed to a repro stand with a 60 mm macro objective. The camera was connected to a computer with Nikon Camera Control Pro software version 2.0. Clustering of seeds was prevented as much as possible. Germination was scored using the Germinator package (Joosen et al., 2010). The calculation of DSDS50 (days of seed dry storage required to reach

50% germination) and the germination under stress treatment is according to (He et al., 2014).

Isolation of total RNA and polysomal RNA and polysome analysis

For the isolation of polysomal RNA, about 400 mg of freshly harvested seeds was extracted with 8 ml of polysome extraction buffer, PEB (400 mM Tris, pH 9.0, 200 mM KCl, 35 mM $MgCl_2$, 5 mM EGTA 50 µg/mL Cycloheximide, 50 µg/mL Chloramphenicol) modified from (Subramanian, 1978). Aliquot of the raw extract was used for total RNA isolation. The extracts were loaded on top of a sucrose cushion (1.75 M sucrose in PEB) and centrifuged (18h, 90,000 g) using a Beckman Ti70 rotor for 18 h (Beckman Coulter, Brea, USA). The resulting pellet was resuspended in wash buffer (200 mM Tris, pH 9.0, 200 mM KCl, 0.025 M EGTA, 35 mM $MgCl_2$, 5 mM DTT, 50 µg/mL Cycloheximide, 50 µg/mL Chloramphenicol) and loaded on a 20-60% linear sucrose gradient, centrifuged at 190,000 g for 1.5 h at 4°C using Beckman SW55 rotor (Beckman Coulter). After ultracentrifugation, the gradients were fractionated into 20 fractions using a Teledyne Isco Density Gradient Fractionation System (Teledyne Isco. Lincoln, USA) with online spectrophotometric detection of (254 nm). The polysomal fractions were pooled for polysomal RNA isolation. The ribosome abundance is reflected by the area under the curve and was calculated after subtracting the baseline obtained by measuring a blank gradient and normalizing to total area under the curve to account for possible uneven loading of the gradients.

Microarray hybridization and data analysis

Affymetrix Arabidopsis Gene 1.1 ST Arrays (Affymetrix, Santa Clara, USA) were hybridized using the GeneChip® 3› IVT Express kit (cat. # 901229) according to instructions from the manufacturer. Hybridization data were analyzed and gene specific signal intensities were computed using the R statistical programming environment (www.R-project.com) and the BioConductor package Affy (Gautier et al., 2004) and the Brainarray cdf file ver. 17.1.0 (http://brainarray.mbni.med.umich.

edu/). Our analyses make use of three data sets: a small experimental study consisting of five MGU74A mouse GeneChip arrays, part of the data from an extensive spike-in study conducted by Gene Logic and Wyeth's Genetics Institute involving 95 HG-U95A human GeneChip arrays; and part of a dilution study conducted by Gene Logic involving 75 HG-U95A GeneChip arrays. We display some familiar features of the perfect match and mismatch probe (PM and MM. DNA microarray data are available in the Gene Expression Omnibus (GEO) repository (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE76904. The LIMMA package was used for normalization and to obtain differential gene expression data (Smyth, 2005; Diboun et al., 2006). The limma and affy package were used for RMA normalization (Irizarry et al., 2003). Probe set intensity signals that never exceeded the noise threshold (logExprs<4 in all samples) were removed. A linear model and Empirical Bayes methods were applied for assessing differential expression (Smyth, 2004). Principle component analysis (PCA) was performed using TM4 (Saeed et al., 2003).

Ribosomal protein extraction, tryptic digestion and peptide purification

The ribosome fractionation was performed as described above. The fractions corresponding to monosome and polysome regions were separately pooled and 1.1 ml of methanol was added to and 0.6 ml of each pooled ribosome extract and mixed thoroughly. After adding 0.3 ml of chloroform, vortexing and centrifugation for 15 minutes at 16,000 g, the supernatant was carefully removed with a capillary pipette. The air-dried protein pellet was dissolved in 100 µl of 8 M urea and an aliquot was taken for protein estimation using the Qubit® protein HS Assay Kit (ThermoFisher, Waltham, USA). Between 4-15 ug of protein dissolved in 100 ul of 8 M urea was used for digestion after reduction (5 mM DTT; 30 min at 37°C) and alkylation (15 mM iodoacetamide; 30 min at RT in the dark) of the cysteines. After adding 75 ul of 0.1 M ammonium bicarbonate buffer (pH8.3), 25 µl of Trypsin/Lys-C Mix (0.02 ug/µl, Promega, Madison, USA) was added for digestion (3 hours at 37°C). After dilution with 200 ul of 0.1 M ammonium bicarbonate, 4 ul of modified trypsin

(0.05 µg/µl, Promega, Madison, USA) was added for overnight incubation at 37°C. Digestion was terminated by adding 0.1% trifluoroacetic acid (TFA). The tryptic digests were cleaned by reverse phase SPE (solid phase extraction) on an Oasis HLB Elution Plate (Water). Peptides were eluted with 0.1 ml of 50% acetonitrile (ACN), 0.1% TFA and dried by vacuum centrifugation. For 2D-LC-MS analysis the peptides were dissolved in 20 µl of 0.1 M ammonium formate (pH 10).

Mass spectrometry measurement

For high resolution separation of the ribosomal protein digests, a nanoAcquity 2-D UPLC system (Waters Corporation, Manchester, UK) was used employing orthogonal reverse phase separation at high and low pH, respectively. With this 2-D set up, the pool of peptides was eluted from the first dimension XBridge C18 trap column (in 20 mM amononium formate pH10) in two steps of 20% and 65% ACN. For the second dimension we used a BEH C18 column (75 µm ×25 cm, water UK) eluting with a 100 minutes linear gradient from 3 to 40% ACN (in 0.1% FA) at 200 nl/min. The eluting peptides were on-line injected into a Q-Exactive Plus (Thermo Scientific) mass spectrometer using a nano-electrospray source. Ionisation (2.4 kV) was performed using a stainless steel emitter and a heated capillary temperature of 250˚C. Full MS scans were acquired over the m/z range 400-1500 with a mass resolution of 70 000 (at m/z 200). Full scan target was set 3x106 with a maximum fill time of 50 ms. The 10 most intense peaks with charge state 2-4 were fragmented in the HCD collision cell with a normalized collision energy of 28%. The mass range was set to 140-2000, and a mass resolution of 17500 (at m/z 200). The target value for fragment scans was set at 5x104, the intensity threshold was kept at 1x104 and the maximum allowed accumulation times were 100 ms. Peptide match was set to preferred, isolation window at 1.6 m/z and isotope exclusion was on. The dynamic exclusion was set to 30 s.

Database search for protein identification and quantitative analysis

The raw Q_{EX} LC-MS/MS data were processed with the work flow described for plant proteomics using the Maxquant (Cox and Mann, 2008; Smaczniak et al., 2012) and the resulting masses file containing all the fragment information was matched against the an in house TAIR protein sequence database based on The Arabidopsis Information Resource (TAIR10, www.Arabidopsis.org). The proteinGroups.txt file generated from Maxquant was used for quantitative analysis. Intensity determination and normalization was performed by algorithm MaxLFQ (Cox et al., 2014). The LFQ value generated was used for further statistics analysis. A protein was identified as present if LFQ is above background value for all the three biological replicates and followed for further statistical analysis. Proteins identified in either monosome fraction or polyosome fractions were assigned as proteins with specific ribosome association. Proteins present in both fractions were first Log2 transformed followed by differential analysis with LIMMA package to identify proteins with different ribosome associations. Two fold difference with corrected P-value<0.05 (Benjamini-Hochberg) were used as criteria for differential presence. A monosome and polysome specifically associated protein was assigned if it was identified by either one of above two methods.

Gene and protein function classification and over-representation analysis

DAVID Bioinformatics Resources 6.7 were used for gene functional annotation and GO enrichment analysis (Huang et al., 2009). Panther classification system was used for ribosome protein classification and statistical overrepresentation test (Mi et al., 2013). Panther protein class was used as reference list which contains 26,684 proteins in the database. Bonferroni corrected protein list was generated with *P*-value<0.05

Supporting Information

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

Figure S1. The identification of *abci20* as a high dormant mutant line among the down-stream targets of DOG6. *P*-value<0.0001 as indicated by the asterisks

Figure S2. Expression profile of *ABCI20*. (A) The expression of *ABCI20* across plant development. (B) The expression of *ABCI20* during seed germination in both fresh harvested (Fresh) and after-ripened (AR) seeds in both micropylar endosperm (ME) and radicle (RAD) (Dekkers et al., 2015).

Figure S3. The comparison of gene differentially expressed in *abci20* to the genes regulated during seed maturation (Angelovici et al., 2009).

Figure S4. Sequence alignment of ABCI20 with other homologs from different species. (A) Sequence similarity matrix. Data from up-right panel indicates the percentage of similarity between each compared sequence pairs and down-left panel indicates the number of aligned amino acids in each pair (B) Phylogenetic tree for all the evaluated sequences.

Supplementary Table S1. Differentially regulated genes in *abci20*

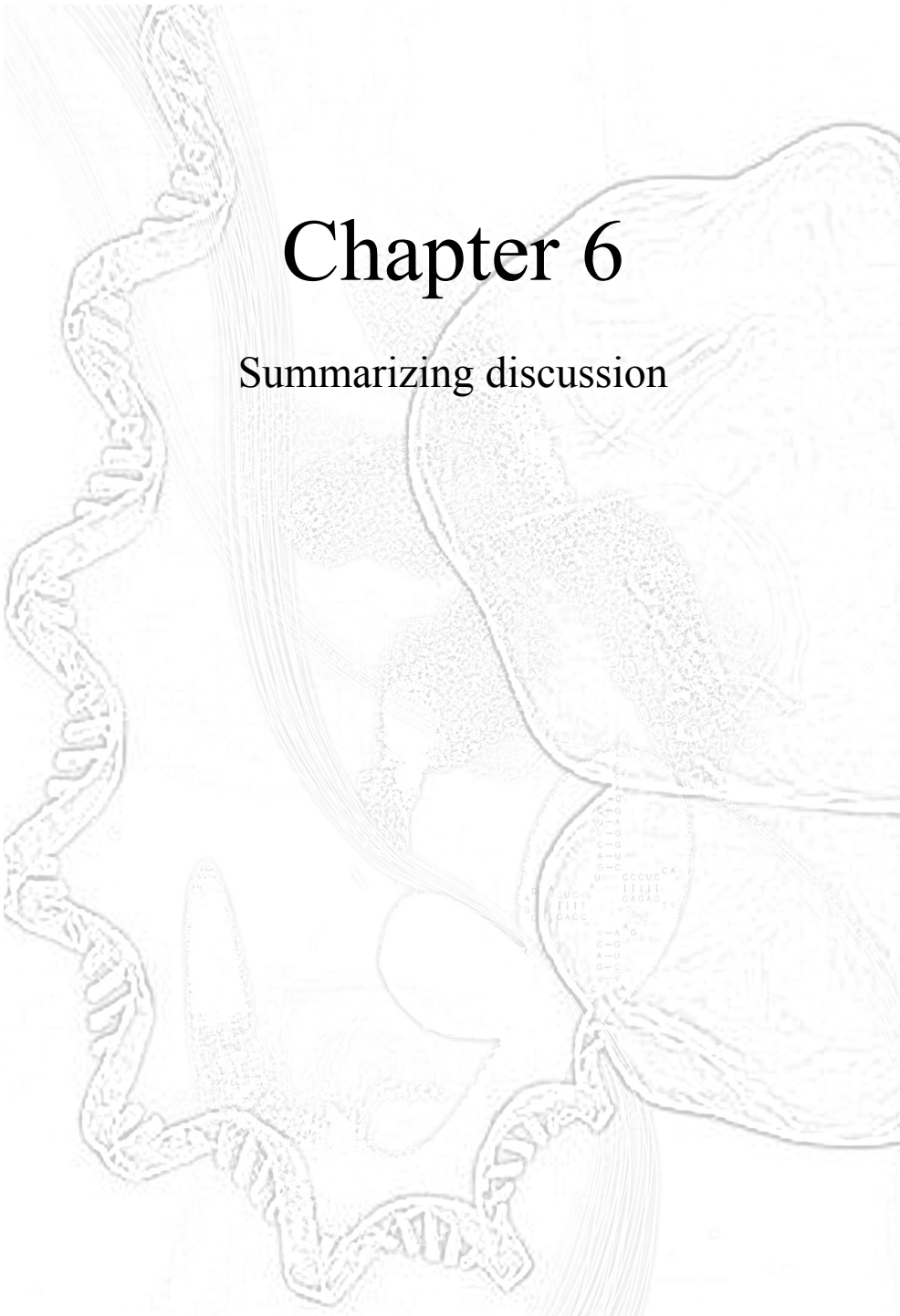
Supplementary Table S2. GO analysis of gene differentially expressed in *abci20*

Supplementary Table S3. Ribosome proteins differentially present in polysome and monosome fraction of *abci20*. Specific proteins are the protein only detected or complete absent in that ribosome fraction and enriched peroteins are protein detected in all fractions but signicantly high or low in that fraction for *abci20* mutant

Supplementary Table S4. GO analysis of protein differentially expressed in *abci20* in both monosome and polysome association

Supplementary Table S5. Sequence feature associated with *abci20* differentially regulated genes

Supplementary Table S6. Down-stream targets of DOG6. The dormant and non-dormant of NIL*DOG6* seeds after 24 hours imbibition were compared and the differentially regulated genes are arranged in the columns with Log2(Fold Change), adjusted *P*-value and gene descriptions.



Chapter 6

Summarizing discussion

Seeds provide an ideal model for studying translational regulation

Seeds represent one of the most important evolutionary innovations. They not only provide food reserves but also enable plants to keep dormant until the arrival of optimal conditions for germination and thus endow plants the ability to cope with harsh environmental conditions after seed dispersal. Maternally derived reserves including mRNA are essential for seed germination. During development and germination seeds cycle through phases of dehydration and rehydration. Dehydration (or desiccation) occurs during seed maturation and rehydration takes place when the seed imbibes water for germination. This makes the seed an interesting system to investigate translational activity associated with developmentally controlled dehydration and rehydration processes. Knowledge on translational regulation might allow the manipulation of gene expression to improve seed quality related traits such as seed germination, dormancy and longevity.

Translatome profiling provides a powerful tool to investigate gene translation processes

Sucrose gradient centrifugation is used to purify molecules such as virus, ribosomes and membranes. It is widely used to investigate cellular ribosome dynamics that can give insight into the protein translation process (Woese, 1961; Marcus et al., 1966; Rinaldi & Monroy, 1969; App et al., 1971; Fountain & Bewley, 1973; Speigel & Marcus, 1975). With the development of high through-put techniques for RNA profiling such as microarray analyses and RNA-sequencing, also the identification of mRNAs associated with the ribosomes became feasible. Numerous reports have employed this technique to study environmental stimuli such as oxidation (Halbeisen & Gerber, 2009), heat (Yanguez et al., 2013), hypoxia (Branco-Price et al., 2005), darkness (Liu et al., 2013), sugar feeding (Gamm et al., 2014) and endoplasmic reticulum (ER) stress (Krishnan et al., 2014) in variable species. Ribosomes can also be purified using the TRAP technique (translating ribosome affinity purification) (Zanetti et al., 2005; Reynoso et al., 2015; Sorenson & Bailey-Serres, 2015) that makes use of a tagged ribosome fusion protein driven by different promoters to

immuno-precipitate ribosomes in different tissue types followed by mRNA profiling. This technique has also been successfully used in translatome studies, using the His-FLAG tagged ribosome protein RPL18 (Mustroph et al., 2009; Lin et al., 2014; Vragovic et al., 2015). In my thesis, I choose sucrose gradient translatome profiling since I am interested in the seed related developmental stages. Translatome profiling in these different seed stages has an advantage over tagging ribosomes since it is challenging to identify promoters that are expressed evenly in all seed stages and tissues at high enough levels for immunoprecipitation. In addition, I have studied protein translation in different genotypes, i.e. Columbia wild type, NIL*DOG1* and the *abci20* mutant. Thus the use of sucrose gradient fractionation avoids the transformation and time-consuming selection of transformants that is required for TRAP. In this thesis, translatome profiling techniques are widely applied for studying the translational control during seed germination and post-germinative growth (**Chapter 2**), mRNAs associated to the ribosome in seeds and seedlings (**Chapter 3**), translational regulation of seed dormancy (**Chapter 4**) and gene (*ABCI20*) specific translational control (**Chapter 5**).

Translational regulation is highly selective

Translational regulation refers to the control of the levels of protein synthesized from a mRNA. A highly significant correlation exists between total transcripts and polysomal associated transcripts and is observed during all the studied seed stages (**Chapter 2-5**). Interestingly, during seed germination, two stages with extensive changes in translational efficiency were identified, that were named the Hydration Shift and Germination Shift stages, based on the time windows at which they were identified and that correspond to seed hydration and germination *sensu stricto*, respectively (**Chapter 2**). In contrast, hardly any translational regulation was identified when investigating the imbibition of dormant seeds (**Chapter 4**) or the *abci20* dry seeds (**Chapter 5**). My combined results indicate the importance of translational regulation during germination, while being less important during dormancy regulation. The high correlation of transcription and translation for

dormancy suggests that here transcription is the key regulator **Chapters 4** and **5**).

Monosomes are the main storage form for mRNAs in dry seeds

Up to now, research on the translatome has mainly focussed on the polysomes since ribosomes tend to be present as clusters on mRNAs that are actively translated. However, when translation is limiting such as in dry seeds, the polysome is hardly visible in the ribosome profile. In **Chapter 3**, I focussed on mRNA storage in dry seeds. I showed that seeds and seedlings have a largely different translational behaviour seen by the single monosome peak in the seeds versus the dramatically enhanced polysome level in seedlings. Given the limited polysomes in the dry seed (**Chapter 2**, Figure 1), it is intriguing to know where/how mRNAs are stored in these seeds. A genome-wide study of the ribosome associated mRNAs and protein shows that monosomes and polysomes are recruited on different mRNA species in association with different messenger ribonucleoproteins (mRNPs) especially those involved in mRNA storage and processing such as stress granule (SG) and processing body (P-bodies) components. These mRNPs are emerging as important cytosolic domains for non-translating mRNA storage and processing. SGs and P-bodies provide targets for investigating mRNA storage, processing and reinitiating for translation. Future studies should be focussed on the function of these monosome and polysome associated proteins in relation to mRNA storage, possibly in relation to seed dormancy, longevity and desiccation tolerance.

The identification of the first mutant with negative correlation between seed dormancy and longevity

Seed dormancy and longevity are negatively correlated as has been recently revealed based on natural genetic variation studies in near isogenic lines (Nguyen *et al.*, 2012; He et al., 2014). In these studies high dormancy levels have been correlated with low seed longevity and vice versa, however no mutant had been identified so far with the same negative correlation. In **Chapter 5** I report on the identification of a dormancy gene, *ABCI20*, identified in the transcriptome of the *Delay of Germination 6* (*DOG6*)

near isogenic line. The *abci20* mutant seeds are highly dormant, hypersensitive to ABA and have a low longevity. This dormancy and longevity trade-off phenotype provides the potential to manipulate seed dormancy and longevity in one go and generate transgenic seeds with low dormancy and high longevity, a highly desirable seed trait in the seed industry. With the approach established in the earlier chapters, the polysome profile of *abci20* seeds was studied and it displayed a strongly enhanced polysome region compared to wild type seeds. This observation stimulated my interest to identify the downstream targets of *ABCI20* in relation to seed dormancy and longevity and to identify factors that contribute to the enhanced polysome level in *abci20*. With this purpose, total and polysomal RNAs were quantified and used for transcriptome and translatome studies. A transcriptional interplay between ABA signalling, ethylene signalling and GA metabolism was identified in relation to the seed dormancy phenotype while oxidation reduction, stress response, autophagy and primary metabolic process are potentially contributing to the observed longevity phenotype. No genes were identified that are under translational control in mature dry *abci20* seeds and thus the enhanced polysome region in *abci20* cannot be explained by translational changes. A further investigation on the polysome associated proteins reveals that the proteins that co-migrate with polysomes contribute to the increased polysome level in the *abci20*. These proteins are enriched for metabolic process and proteasome components, suggesting that the maturation program is interrupted in *abci20*, consistent with the observation that the transcripts that differentially accumulated in *abci20* for a large part overlap with maturation genes (**Chapter 5**, Supplemental Figure 3). The changed expression in association with high dormancy and low longevity in *abci20* seeds (**Chapter 5**, Figure 2, 3) could be explained by the antioxidant and energy theory recently proposed to explain the trade-off mechanism between seed dormancy and longevity (Nguyen & Bentsink, 2015). Based on these results, a model for ABCI20 regulation is proposed while a deeper understanding on how ABCI20 modulates these seed traits remains to be identified.

The significance of translational research in seeds

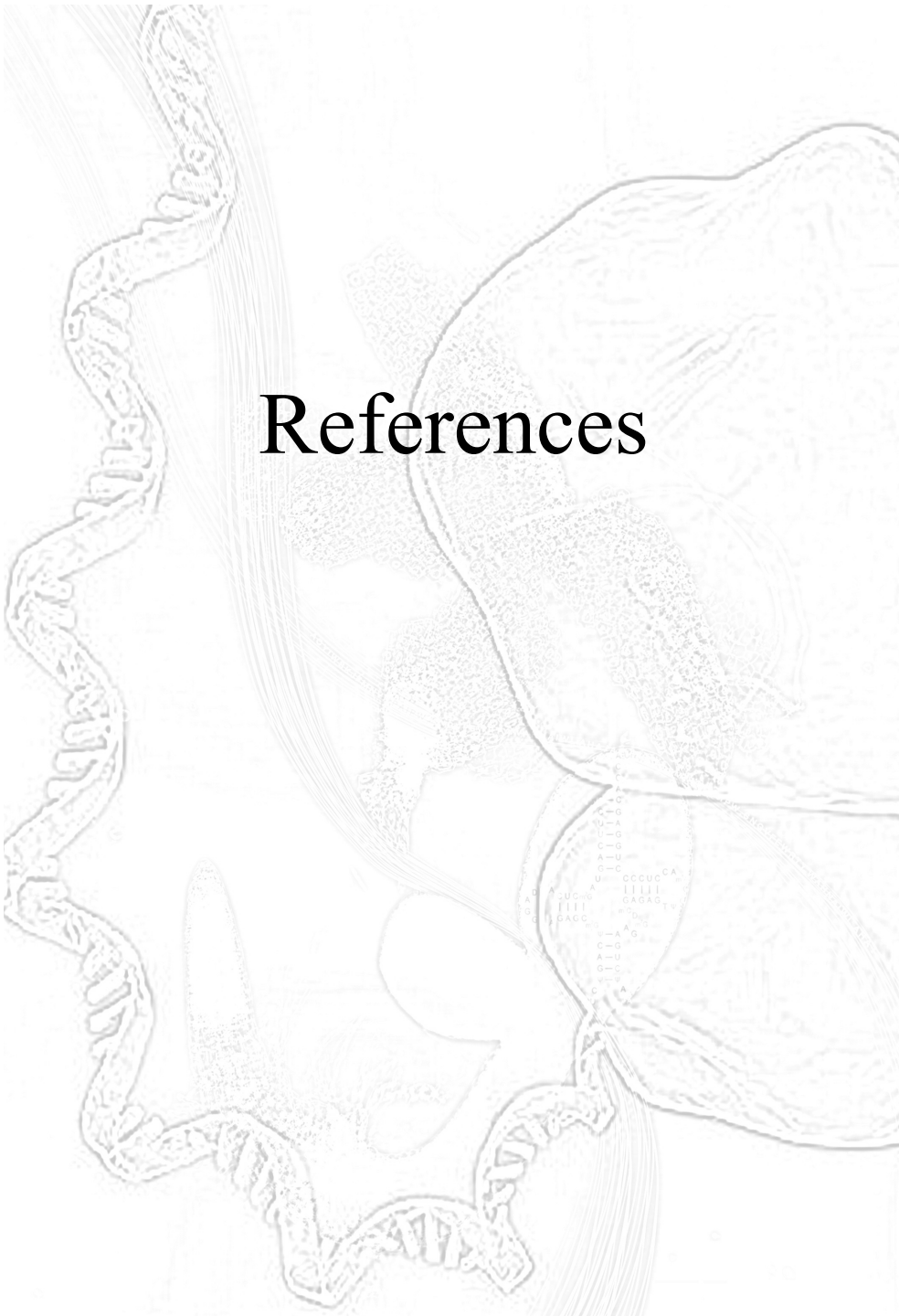
Translational regulation is the control of the levels of protein synthesized from its mRNA. It regulates the last step before the protein is eventually synthesized and thus directly determines the protein synthesis rate. Translatome profiling revealed important translational regulators such as IER3, which in mouse protects the macrophage from cell death during activation (Schott et al., 2014) and Arabidopsis *EIN3-BINDING F-BOX1* (*EBF1*) and *EBF2*, which are translationally repressed through its 3'UTR by *Ethylene Insensitive2* (*EIN2*) (Li et al., 2015; Merchante et al., 2015). Seeds do not have photosynthesis during heterotrophic growth and thus early seed germination is based on reserves stored in the seeds including proteins and mRNAs. The translation of these stored mRNAs determines their functionality during seed germination. In this thesis, mRNAs identified under translational regulation during this process display distinct sequence features (**Chapter 2**). The further understanding of these translationally regulated genes regarding their functional importance for seed germination is important. Another intriguing aspect is the identification of translational regulators, especially mRNA binding proteins (mRNPs) that modulate gene translation during seed germination. These factors likely regulate mRNA translation by recognizing specific mRNA features such as the identified mRNA structure and specific motifs. The identification and characterization of these mRNPs could provide insight in how the interaction of translational regulators and mRNAs is achieved and on the structural basis for the interaction and further translational regulation. Further experimentation will focus on identifying which translational regulatory factors or mRNPs that are responsible for the extensive translational regulation during germination.

Frontier in transcriptional and translational regulation

New technologies, especially cost effective sequencing techniques, dramatically advance the landscape of biological research from a single gene, transcript, protein and metabolite to the global evaluation of whole genome, transcriptome, proteome and metabolome. Microarray and RNA-sequencing (RNA-seq) can evaluate thousands

of transcripts simultaneously in small amounts of mRNA. RNA-seq is becoming more popular since it does not require prior knowledge about the genome and thus can be applied to any species. The complexity of gene regulation makes it difficult to directly convert expression data to mathematical models predicting the final protein production, functionality, or metabolic changes. This inconsistency requires a more detailed readout about gene expression, in an organ and tissue specific way. Methods for this include dissection-based methodologies such as laser capture microdissection (LCM) and sorting based methodologies such as fluorescence-activated cell sorting (FACS) and fluorescence-activated nuclei sorting (FANS) Recently, the evaluation of transcripts recruited by polysomes engaged in translation as a fraction of total transcripts has provided a way to tackle translational regulation by employing TRAP or sucrose-gradient based translatome and ribosome profiling. These techniques help in understanding gene expression regulation at the mRNA level (Arava et al., 2003; Ingolia et al., 2009; Mustroph et al., 2009; Liu et al., 2012; Yanguez et al., 2013; Gamm et al., 2014; Layat et al., 2014; Lin et al., 2014). In contrast, fluorescence-activated nuclei sorting (FANS) and isolation of nuclei tagged in specific cell types (INTACT) provide access to nuclear processes (Deal & Henikoff, 2011; Steiner et al., 2012; Amin et al., 2014; Steiner & Henikoff, 2015). More recently, ribosome profiling has been applied on a subcellular level by investigating translational processes associated with endoplasmic reticulum (ER) and mitochondria (Jan et al., 2014; Williams et al., 2014). These techniques greatly improve the resolution of gene expression regulation analysis and thus provide promising applications in seed biology for gene regulation studies.

In all, this thesis provides new perspectives for research in the seed biology field. Cereal seeds such as rice, wheat and maize are the main source of nourishment of billions of people in the world. Seeds also act as the most important propagation material. Translational regulation in seeds, as an emerging research area, may provide novel insights into the regulatory processes that facilitate the targeted engineering of crops to improved propagation quality traits such as seed dormancy, longevity, germination and post-germination growth.



References

Ajtkhozhin, M.A., Doschanov, K.J., and Akhanov, A.U. (1976). Informosomes as a stored form of mRNA in wheat embryos. FEBS letters **66,** 124-126.

Akashi, H. (1994). Synonymous codon usage in Drosophila melanogaster: natural selection and translational accuracy. Genetics **136,** 927-935.

Alboresi, A., Gestin, C., Leydecker, M.T., Bedu, M., Meyer, C., and Truong, H.N. (2005). Nitrate, a signal relieving seed dormancy in Arabidopsis. Plant Cell Environmentment **28,** 500-512.

Alexa, A., and Rahnenfuhrer, J. (2010). topGO: topGO: Enrichment analysis for Gene Ontology. R package version 2.18.0.

Ali-Rachedi, S., Bouinot, D., Wagner, M.H., Bonnet, M., Sotta, B., Grappin, P., and Jullien, M. (2004). Changes in endogenous abscisic acid levels during dormancy release and maintenance of mature seeds: studies with the Cape Verde Islands ecotype, the dormant model of *Arabidopsis thaliana*. Planta **219,** 479-488.

Almoguera, C., and Jordano, J. (1992). Developmental and environmental concurrent expression of sunflower dry-seed-stored low-molecular-weight heat-shock protein and Lea mRNAs. Plant molecular biology **19,** 781-792.

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

Amin, N.M., Greco, T.M., Kuchenbrod, L.M., Rigney, M.M., Chung, M.I., Wallingford, J.B., Cristea, I.M., and Conlon, F.L. (2014). Proteomic profiling of cardiac tissue by isolation of nuclei tagged in specific cell types (INTACT). Development **141,** 962-973.

Andrei, M.A., Ingelfinger, D., Heintzmann, R., Achsel, T., Rivera-Pomar, R., and Luhrmann, R. (2005). A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. RNA-a Publication of the RNA Society **11,** 717-727.

Angelovici, R., Fait, A., Zhu, X., Szymanski, J., Feldmesser, E., Fernie, A.R., and Galili, G. (2009). Deciphering transcriptional and metabolic networks associated with lysine metabolism during Arabidopsis seed development. Plant Physiology **151,** 2058-2072.

App, A.A., Bulis, M.G., and McCarthy, W.J. (1971). Dissociation of ribosomes and seed germination. Plant Physiology **47,** 81-86.

Aranda, A., and Proudfoot, N. (2001). Transcriptional termination factors for RNA polymerase II in yeast. Molecular cell **7,** 1003-1011.

Arava, Y., Wang, Y., Storey, J.D., Liu, C.L., Brown, P.O., and Herschlag, D. (2003). Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A **100,** 3889-3894.

Arc, E., Chibani, K., Grappin, P., Jullien, M., Godin, B., Cueff, G., Valot, B., Balliau, T., Job, D., and Rajjou, L. (2012). Cold stratification and exogenous nitrates entail similar functional proteome adjustments during Arabidopsis seed dormancy release. Journal of proteome research **11,** 5418-5432.

Arciga-Reyes, L., Wootton, L., Kieffer, M., and Davies, B. (2006). UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in Arabidopsis. Plant Journal **47,**

480-489.

Arimoto, K., Fukuda, H., Imajoh-Ohmi, S., Saito, H., and Takekawa, M. (2008). Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. Nature cell biology **10,** 1324-1332.

Aspart, L., Meyer, Y., Laroche, M., and Penon, P. (1984). Developmental regulation of the synthesis of proteins encoded by stored mRNA in radish embryos. Plant Physiology **76,** 664-673.

Azzouz, N., Panasenko, O.O., Deluen, C., Hsieh, J., Theiler, G., and Collart, M.A. (2009). Specific roles for the Ccr4-Not complex subunits in expression of the genome. RNA-a Publication of the RNA Society **15,** 377-383.

Baena-Gonzalez, E., Rolland, F., Thevelein, J.M., and Sheen, J. (2007). A central integrator of transcription networks in plant stress and energy signalling. Nature **448,** 938-942.

Baerenfaller, K., Grossmann, J., Grobei, M.A., Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., Zimmermann, P., Grossniklaus, U., Gruissem, W., and Baginsky, S. (2008). Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. Science **320,** 938-941.

Bailey-Serres, J. (1999). Selective translation of cytoplasmic mRNAs in plants. Trends Plant Science **4,** 142-148.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Research **37,** W202-208.

Bailly, C., and Kranner, I. (2011). Analyses of reactive oxygen species and antioxidants in relation to seed longevity and germination. Methods Molecular Biology **773,** 343-367.

Balagopal, V., and Parker, R. (2009). Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. Current opinion in cell biology **21,** 403-408.

Balzer, E., and Moss, E.G. (2007). Localization of the developmental timing regulator Lin28 to mRNP complexes, P-bodies and stress granules. RNA biology **4,** 16-25.

Barak, S., Singh Yadav, N., and Khan, A. (2014). DEAD-box RNA helicases and epigenetic control of abiotic stress-responsive gene expression. Plant signaling & behavior **9,** e977729.

Basbouss-Serhal, I., Soubigou-Taconnat, L., Bailly, C., and Leymarie, J. (2015). Germination potential of dormant and nondormant Arabidopsis seeds is driven by distinct recruitment of messenger RNAs to polysomes. Plant Physiology **168,** 1049-1065.

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

Beaudoin, N., Serizet, C., Gosti, F., and Giraudat, J. (2000). Interactions between abscisic acid and ethylene signaling cascades. Plant Cell **12,** 1103-1115.

Bechade, C., Rostaing, P., Cisterni, C., Kalisch, R., La Bella, V., Pettmann, B., and Triller, A. (1999). Subcellular distribution of survival motor neuron (SMN) protein:

possible involvement in nucleocytoplasmic and dendritic transport. European Journal of Neuroscience **11,** 293-304.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). MRNA degradation by miRNAs and GW182 requires both CCR4 : NOT deadenylase and DCP1 : DCP2 decapping complexes. Gene Development **20,** 1885-1898.

Behura, S.K., and Severson, D.W. (2011). Coadaptation of isoacceptor tRNA genes and codon usage bias for translation efficiency in *Aedes aegypti* and *Anopheles gambiae*. Insect Molecular Biology **20,** 177-187.

Beltaifa, S., Webster, M.J., Ligons, D.L., Fatula, R.J., Herman, M.M., Kleinman, J.E., and Weickert, C.S. (2005). Discordant changes in cortical TrkC mRNA and protein during the human lifespan. The European journal of neuroscience **21,** 2433-2444.

Beltran-Pena, E., Ortiz-Lopez, A., and Sanchez de Jimenez, E. (1995). Synthesis of ribosomal proteins from stored mRNAs early in seed germination. Plant molecular biology **28,** 327-336.

Bentsink, L., and Koornneef, M. (2008). Seed dormancy and germination. The Arabidopsis book / American Society of Plant Biologists **6,** e0119.

Bentsink, L., Jowett, J., Hanhart, C.J., and Koornneef, M. (2006). Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis. Proc Natl Acad Sci U S A **103,** 17042-17047.

Bentsink, L., Hanson, J., Hanhart, C.J., Blankestijn-de Vries, H., Coltrane, C., Keizer, P., El-Lithy, M., Alonso-Blanco, C., de Andres, M.T., Reymond, M., van Eeuwijk, F., Smeekens, S., and Koornneef, M. (2010). Natural variation for seed dormancy in Arabidopsis is regulated by additive genetic and molecular pathways. Proc Natl Acad Sci U S A **107,** 4264-4269.

Bewley, J.D. (1997). Seed Germination and Dormancy. Plant Cell **9,** 1055-1066.

Bewley, J.D., Bradford, K.J., Hilhorst, H.W.M., and Nonogaki, H. (2013). Seeds Physiology of Development, Germination and Dormancy, . (Springer).

Bhaskara, G.B., Nguyen, T.T., and Verslues, P.E. (2012). Unique drought resistance functions of the highly ABA-induced clade A protein phosphatase 2Cs. Plant Physiology **160,** 379-395.

Bleecker, A.B., Estelle, M.A., Somerville, C., and Kende, H. (1988). Insensitivity to ethylene conferred by a dominant mutation in *Arabidopsis thaliana*. Science **241,** 1086-1089.

Bogamuwa, S., and Jang, J.C. (2013). The Arabidopsis tandem CCCH zinc finger proteins AtTZF4, 5 and 6 are involved in light-, abscisic acid- and gibberellic acid- mediated regulation of seed germination. Plant Cell Environment **36,** 1507-1519.

Borukhov, S., Polyakov, A., Nikiforov, V., and Goldfarb, A. (1992). Grea Protein - a Transcription Elongation-Factor from Escherichia-Coli. Proc Natl Acad Sci U S A **89,** 8899-8902.

Brady, S.M., Sarkar, S.F., Bonetta, D., and 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. Plant Journal **34,** 67-75.

Branco-Price, C., Kawaguchi, R., Ferreira, R.B., and Bailey-Serres, J. (2005). Genome-wide analysis of transcript abundance and translation in Arabidopsis seedlings subjected to oxygen deprivation. Ann Bot **96,** 647-660.

Branco-Price, C., Kaiser, K.A., Jang, C.J., Larive, C.K., and Bailey-Serres, J. (2008). Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*. Plant Journal **56,** 743-755.

Browning, K.S. (2004). Plant translation initiation factors: it is not easy to be green. Biochemical Society transactions **32,** 589-591.

Buchan, J.R., and Parker, R. (2009). Eukaryotic stress granules: the ins and outs of translation. Molecular cell **36,** 932-941.

Buchan, J.R., Muhlrad, D., and Parker, R. (2008). P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. Journal of Cell Biology **183,** 441-455.

Buitink, J., Vu, B.L., Satour, P., and Leprince, O. (2003). The re-establishment of desiccation tolerance in germinated radicles of *Medicago truncatula* Gaertn. seeds. Seed Science Research **13,** 273-286.

Bulmer, M. (1987). Coevolution of codon usage and transfer RNA abundance. Nature **325,** 728-730.

Burrill, C.P., Westesson, O., Schulte, M.B., Strings, V.R., Segal, M., and Andino, R. (2013). Global RNA structure analysis of poliovirus identifies a conserved RNA structure involved in viral replication and infectivity. Journal of virology **87,** 11670-11683.

Cadman, C.S., Toorop, P.E., Hilhorst, H.W., and Finch-Savage, W.E. (2006). Gene expression profiles of Arabidopsis Cvi seeds during dormancy cycling indicate a common underlying dormancy control mechanism. The Plant Journal **46,** 805-822.

Cech, T.R., and Steitz, J.A. (2014). The noncoding RNA revolution-trashing old rules to forge new ones. Cell **157,** 77-94.

Chen, D., Sarid, S., and Katchalski, E. (1968). Studies on the nature of messenger RNA in germinating wheat embryos. Proc Natl Acad Sci U S A **60,** 902-909.

Chen, G., Gharib, T.G., Huang, C.C., Taylor, J.M., Misek, D.E., Kardia, S.L., Giordano, T.J., Iannettoni, M.D., Orringer, M.B., Hanash, S.M., and Beer, D.G. (2002). Discordant protein and mRNA expression in lung adenocarcinomas. Molecular & cellular proteomics : MCP **1,** 304-313.

Chen, M., MacGregor, D.R., Dave, A., Florance, H., Moore, K., Paszkiewicz, K., Smirnoff, N., Graham, I.A., and Penfield, S. (2014). Maternal temperature history activates Flowering Locus T in fruits to control progeny dormancy according to time of year. Proc Natl Acad Sci U S A **111,** 18787-18792.

Chiba, Y., Johnson, M.A., Lidder, P., Vogel, J.T., van Erp, H., and Green, P.J. (2004). AtPARN is an essential poly(A) ribonuclease in Arabidopsis. Gene **328,** 95-102.

Chibani, K., Ali-Rachedi, S., Job, C., Job, D., Jullien, M., and Grappin, P. (2006). Proteomic analysis of seed dormancy in Arabidopsis. Plant Physiology **142,** 1493-1510.

Chitnis, V.R., Gao, F., Yao, Z., Jordan, M.C., Park, S., and Ayele, B.T. (2014). After-ripening induced transcriptional changes of hormonal genes in wheat seeds: the cases of brassinosteroids, ethylene, cytokinin and salicylic acid. Plos One **9,** e87543.

Chiwocha, S.D.S., Cutler, A.J., Abrams, S.R., Ambrose, S.J., Yang, J., Ross, A.R.S., and Kermode, A.R. (2005). The etr1-2 mutation in *Arabidopsis thaliana* affects the abscisic acid, auxin, cytokinin and gibberellin metabolic pathways during maintenance of seed dormancy, moist-chilling and germination. Plant Journal **42,** 35-48.

Cho, S.K., Ryu, M.Y., Seo, D.H., Kang, B.G., and Kim, W.T. (2011). The Arabidopsis RING E3 Ubiquitin Ligase AtAIRP2 Plays Combinatory Roles with AtAIRP1 in Abscisic Acid-Mediated Drought Stress Responses. Plant Physiology **157,** 2240-2257.

Chu, C.Y., and Rana, T.M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. PLoS biology **4,** 1122-1136.

Comai, L., and Harada, J.J. (1990). Transcriptional activities in dry seed nuclei indicate the timing of the transition from embryogeny to germination. Proc Natl Acad Sci U S A **87,** 2671-2674.

Cook, C.E., Coggon, P., Mcphail, A.T., Wall, M.E., Whichard, L.P., Egley, G.H., and Luhan, P.A. (1972). Germination Stimulants .2. Structure of Strigol - Potent Seed-Germination Stimulant for Witchweed (Striga-Lutea Lour). Journal of the American Chemical Society **94,** 6198-&.

Copeland, P.R., and Wormington, M. (2001). The mechanism and regulation of deadenylation: Identification and characterization of Xenopus PARN. RNA **7,** 875-886.

Corbineau, F., Xia, Q., Bailly, C., and El-Maarouf-Bouteau, H. (2014). Ethylene, a key factor in the regulation of seed dormancy. Frontiers in plant science **5,** 539.

Costa, M.C.D., Righetti, K., Nijveen, H., Yazdanpanah, F., Ligterink, W., Buitink, J., and Hilhorst, H.W.M. (2015). A gene co-expression network predicts functional genes controlling the re-establishment of desiccation tolerance in germinated *Arabidopsis thaliana* seeds. Planta **242,** 435-449.

Cougot, N., Babajko, S., and Seraphin, B. (2004). Cytoplasmic foci are sites of mRNA decay in human cells. Journal of Cell Biology **165,** 31-40.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology **26,** 1367-1372.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Molecular & Cellular Proteomics **13,** 2513-2526.

Crick, F.H. (1966a). Codon--anticodon pairing: the wobble hypothesis. Journal of molecular biology **19,** 548-555.

Crick, F.H.C. (1966b). Codon-Anticodon Pairing : The Wobble Hypothesis Journal of

molecular biology **19,** 548-555.

Cui, Y., Ramnarain, D.B., Chiang, Y.C., Ding, L.H., McMahon, J.S., and Denis, C.L. (2008). Genome wide expression analysis of the CCR4-NOT complex indicates that it consists of three modules with the NOT module controlling SAGA-responsive genes. Molecular Genetics and Genomics **279,** 323-337.

Cutler, S., Ghassemian, M., Bonetta, D., Cooney, S., and McCourt, P. (1996). A protein farnesyl transferase involved in abscisic acid signal transduction in Arabidopsis. Science **273,** 1239-1241.

Davies, M.J. (2005). The oxidative environment and protein damage. BBA-Proteins Proteom **1703,** 93-109.

De Gara, L., de Pinto, M.C., Moliterni, V.M., and D'Egidio, M.G. (2003). Redox regulation and storage processes during maturation in kernels of *Triticum durum*. Journal of Experimental Botany **54,** 249-258.

de Klerk, E., Fokkema, I.F., Thiadens, K.A., Goeman, J.J., Palmblad, M., den Dunnen, J.T., von Lindern, M., and t Hoen, P.A. (2015). Assessing the translational landscape of myogenic differentiation by ribosome profiling. Nucleic Acids Research.

De Veylder, L., Beeckman, T., and Inze, D. (2007). The ins and outs of the plant cell cycle. Nature Reviews Molecular Cell Biology **8,** 655-665.

Deal, R.B., and Henikoff, S. (2011). The INTACT method for cell type-specific gene expression and chromatin profiling in *Arabidopsis thaliana*. Nat Protocol **6,** 56-68.

Decca, M.B., Carpio, M.A., Bosc, C., Galiano, M.R., Job, D., Andrieux, A., and Hallak, M.E. (2007). Post-translational arginylation of calreticulin - A new isospecies of calreticulin component of stress granules. Journal Biology Chemistry **282,** 8237-8245.

Decker, C.J., and Parker, R. (2012). P-Bodies and Stress Granules: Possible roles in the control of translation and mRNA degradation. Cold Spring Harbor Perspectives in Biology **4**.

Dehlin, E., Wormington, M., Korner, C.G., and Wahle, E. (2000). Cap-dependent deadenylation of mRNA. Embo Journal **19,** 1079-1086.

Dekkers, B.J., Pearce, S., van Bolderen-Veldkamp, R.P., Hanson, J., Holdsworth, M.J., and Bentsink, L. (2015). Dormant and after-ripened seeds are distinguished by early transcriptional differences in the imbibed state (Under review).

Dekkers, B.J., Costa, M.C., Maia, J., Bentsink, L., Ligterink, W., and Hilhorst, H.W. (2015). Acquisition and loss of desiccation tolerance in seeds: from experimental model to biological relevance. Planta **241,** 563-577.

Dekkers, B.J., Pearce, S., van Bolderen-Veldkamp, R.P., Marshall, A., Widera, P., Gilbert, J., Drost, H.G., Bassel, G.W., Muller, K., King, J.R., Wood, A.T., Grosse, I., Quint, M., Krasnogor, N., Leubner-Metzger, G., Holdsworth, M.J., and Bentsink, L. (2013). Transcriptional dynamics of two seed compartments with opposing roles in Arabidopsis seed germination. Plant Physiology **163,** 205-215.

Delseny, M., Aspart, L., and Guitton, Y. (1977). Disappearance of stored polyadenylic acid and mRNA during early germination of radish (*Raphanus sativus* L.) embryo axes.

Planta **135,** 125-128.

Denis, C.L., and Chen, J.J. (2003). The CCR4-NOT complex plays diverse roles in mRNA metabolism. Progress in Nucleic Acid Research and Molecular Biology **73,** 221-250.

Dennis, M.D., and Browning, K.S. (2009). Differential phosphorylation of plant translation initiation factors by *Arabidopsis thaliana* CK2 holoenzymes. Journal Biology Chemistry **284,** 20602-20614.

Dennis, M.D., Person, M.D., and Browning, K.S. (2009). Phosphorylation of plant translation initiation factors by CK2 enhances the in vitro interaction of multifactor complex components. Journal Biology Chemistry **284,** 20615-20628.

DePaula, M., PerezOtaola, M., Darder, M., Torres, M., Frutos, G., and MartinezHonduvilla, C.J. (1996). Function of the ascorbate-glutathione cycle in aged sunflower seeds. Physiologia Plantarum **96,** 543-550.

Deprost, D., Yao, L., Sormani, R., Moreau, M., Leterreux, G., Nicolai, M., Bedu, M., Robaglia, C., and Meyer, C. (2007). The Arabidopsis TOR kinase links plant growth, yield, stress resistance and mRNA translation. EMBO reports **8,** 864-870.

Devos, C.H.R., Kraak, H.L., and Bino, R.J. (1994). Aging of Tomato Seeds Involves Glutathione Oxidation. Physiologia Plantarum **92,** 131-139.

Diboun, I., Wernisch, L., Orengo, C.A., and Koltzenburg, M. (2006). Microarray analysis after RNA amplification can detect pronounced differences in gene expression using limma. BMC genomics **7,** 252.

Ding, Y., Tang, Y., Kwok, C.K., Zhang, Y., Bevilacqua, P.C., and Assmann, S.M. (2014a). In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. Nature **505,** 696-700.

Ding, Z.J., Yan, J.Y., Li, G.X., Wu, Z.C., Zhang, S.Q., and Zheng, S.J. (2014b). WRKY41 controls Arabidopsis seed dormancy via direct regulation of *ABI3* transcript levels not downstream of ABA. The Plant Journal **79,** 810-823.

Divi, U.K., and Krishna, P. (2010). Overexpression of the Brassinosteroid Biosynthetic Gene AtDWF4 in Arabidopsis Seeds Overcomes Abscisic Acid-induced Inhibition of Germination and Increases Cold Tolerance in Transgenic Seedlings. Journal of Plant Growth Regulation **29,** 385-393.

Doroshenk, K.A., Crofts, A.J., Morris, R.T., Wyrick, J.J., and Okita, T.W. (2009). Proteomic analysis of cytoskeleton-associated RNA binding proteins in developing rice seed. Journal of Proteome Research **8,** 4641-4653.

Dos Santos, C.V., Laugier, E., Tarrago, L., Massot, V., Issakidis-Bourguet, E., Rouhier, N., and Rey, P. (2007). Specificity of thioredoxins and glutaredoxins as electron donors to two distinct classes of Arabidopsis plastidial methionine sulfoxide reductases B. FEBS letters **581,** 4371-4376.

Drummond, D.A., and Wilke, C.O. (2008). Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. Cell **134,** 341-352.

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

Duret, L. (2000). tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. Trends in genetics : TIG **16,** 287-289.

El-Maarouf-Bouteau, H., Meimoun, P., Job, C., Job, D., and Bailly, C. (2013). Role of protein and mRNA oxidation in seed dormancy and germination. Frontiers in Plant Science **4,** 77.

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

Falcone Ferreyra, M.L., Pezza, A., Biarc, J., Burlingame, A.L., and Casati, P. (2010). Plant L10 ribosomal proteins have different roles during development and translation under ultraviolet-B stress. Plant Physiology **153,** 1878-1894.

Fasken, M.B., and Corbett, A.H. (2005). Process or perish: quality control in mRNA biogenesis. Nature Structural & Molecular Biology **12,** 482-488.

Feng, H., Chen, Q., Feng, J., Zhang, J., Yang, X., and Zuo, J. (2007). Functional characterization of the Arabidopsis eukaryotic translation initiation factor 5A-2 that plays a crucial role in plant growth and development by regulating cell division, cell growth, and cell death. Plant Physiology **144,** 1531-1545.

Fernie, A.R., and Stitt, M. (2012). On the discordance of metabolomics with proteomics and transcriptomics: coping with increasing complexity in logic, chemistry, and network interactions scientific correspondence. Plant Physiology **158,** 1139-1145.

Finch-Savage, W.E., Cadman, C.S., Toorop, P.E., Lynn, J.R., and Hilhorst, H.W. (2007). Seed dormancy release in Arabidopsis Cvi by dry after-ripening, low temperature, nitrate and light shows common quantitative patterns of gene expression directed by environmentally specific sensing. The Plant Journal **51,** 60-78.

Finkelstein, R.R. (1994). Mutations at two new Arabidopsis ABA response loci are similar to the abi3 mutations. The Plant Journal **5,** 765-771.

Finkelstein, R.R., Gampala, S.S., and Rock, C.D. (2002). Abscisic acid signaling in seeds and seedlings. Plant Cell **14 Suppl,** S15-45.

Fontana, L., Partridge, L., and Longo, V.D. (2010). Extending Healthy Life Span-From Yeast to Humans. Science **328,** 321-326.

Fountain, D.W., and Bewley, J.D. (1973). Polyribosome Formation and Protein-Synthesis in Imbibed but Dormant Lettuce Seeds. Plant Physiology **52,** 604-607.

Fu, Q., Wang, B.C., Jin, X., Li, H.B., Han, P., Wei, K.H., Zhang, X.M., and Zhu, Y.X. (2005). Proteomic analysis and extensive protein identification from dry, germinating Arabidopsis seeds and young seedlings. Journal of Biochemistry and Molecular Biology **38,** 650-660.

Fujimura, K., Kano, F., and Murata, M. (2008). Identification of PCBP2, a facilitator of IRES-mediated translation, as a novel constituent of stress granules and processing bodies. Rna-a Publication of the RNA Society **14,** 425-431.

Galau, G.A., Jakobsen, K.S., and Hughes, D.W. (1991). The Controls of Late Dicot

embryogenesis and early germination. Physiologia Plantarum **81,** 280-288.

Galland, M., and Rajjou, L. (2015). Regulation of mRNA translation controls seed germination and is critical for seedling vigor. Frontiers in Plant Science **6,** 284.

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

Gallardo, K., Job, C., Groot, S.P., Puype, M., Demol, H., Vandekerckhove, J., and Job, D. (2001). Proteomic analysis of arabidopsis seed germination and priming. Plant Physiology **126,** 835-848.

Gallie, D.R., Le, H., Caldwell, C., Tanguay, R.L., Hoang, N.X., and Browning, K.S. (1997). The phosphorylation state of translation initiation factors is regulated developmentally and following heat shock in wheat. Journal Biology Chemistry **272,** 1046-1053.

Gamm, M., Peviani, A., Honsel, A., Snel, B., Smeekens, S., and Hanson, J. (2014). Increased sucrose levels mediate selective mRNA translation in Arabidopsis. BMC Plant Biology **14,** 306.

Gao, F., Rampitsch, C., Chitnis, V.R., Humphreys, G.D., Jordan, M.C., and Ayele, B.T. (2013). Integrated analysis of seed proteome and mRNA oxidation reveals distinct post-transcriptional features regulating dormancy in wheat (*Triticum aestivum* L.). Plant Biotechnology Journal **11,** 921-932.

Garcia, M.E., Lynch, T., Peeters, J., Snowden, C., and Finkelstein, R. (2008). A small plant-specific protein family of ABI five binding proteins (AFPs) regulates stress response in germinating Arabidopsis seeds and seedlings. Plant Molecular Biology **67,** 643-658.

Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics **20,** 307-315.

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nature reviews. Genetics **15,** 829-845.

Gibon, Y., Usadel, B., Blaesing, O.E., Kamlage, B., Hoehne, M., Trethewey, R., and Stitt, M. (2006). Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in Arabidopsis rosettes. Genome Biology **7**.

Goeres, D.C., Van Norman, J.M., Zhang, W.P., Fauver, N.A., Spencer, M.L., and Sieburth, L.E. (2007). Components of the Arabidopsis mRNA decapping complex are required for early seedling development. Plant Cell **19,** 1549-1564.

Gong, Z.Z., Lee, H., Xiong, L.M., Jagendorf, A., Stevenson, B., and Zhu, J.K. (2002). RNA helicase-like protein as an early regulator of transcription factors for plant chilling and freezing tolerance. Proc Natl Acad Sci U S A **99,** 11507-11512.

Gong, Z.Z., Dong, C.H., Lee, H., Zhu, J.H., Xiong, L.M., Gong, D.M., Stevenson, B., and Zhu, J.K. (2005). A DEAD box RNA helicase is essential for mRNA export and important for development and stress responses in Arabidopsis. Plant Cell **17,** 256-267.

Gosai, S.J., Foley, S.W., Wang, D., Silverman, I.M., Selamoglu, N., Nelson, A.D.,

Beilstein, M.A., Daldal, F., Deal, R.B., and Gregory, B.D. (2015). Global Analysis of the RNA-protein interaction and RNA secondary structure landscapes of the Arabidopsis nucleus. Molecular Cell **57,** 376-388.

Gosti, F., Beaudoin, N., Serizet, C., Webb, A.A.R., Vartanian, N., and Giraudat, J. (1999). ABI1 protein phosphatase 2C is a negative regulator of abscisic acid signaling. Plant Cell **11,** 1897-1909.

Goulet, I., Boisvenue, S., Mokas, S., Mazroui, R., and Cote, J. (2008). TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules. Human Molecular Genetics **17,** 3055-3074.

Graeber, K., Linkies, A., Steinbrecher, T., Mummenhoff, K., Tarkowska, D., Tureckova, V., Ignatz, M., Sperber, K., Voegele, A., de Jong, H., Urbanova, T., Strnad, M., and Leubner-Metzger, G. (2014). DELAY OF GERMINATION 1 mediates a conserved coat-dormancy mechanism for the temperature- and gibberellin-dependent control of seed germination. Proc Natl Acad Sci U S A **111,** E3571-3580.

Guo, H.W., and Ecker, J.R. (2003). Plant responses to ethylene gas are mediated by SCF (EBF1/EBF2)-dependent proteolysis of EIN3 transcription factor. Cell **115,** 667-677.

Guo, L., Li, X., and Tang, Q.Q. (2015). Transcriptional regulation of adipocyte differentiation: a central role for CCAAT/enhancer-binding protein (C/EBP) beta. Journal of Biology Chemistry **290,** 755-761.

Halbeisen, R.E., and Gerber, A.P. (2009). Stress-dependent coordination of transcriptome and translatome in yeast. PLoS biology **7,** e1000105.

Hammett, J.R., and Katterman, F.R. (1975). Storage and metabolism of poly(adenylic acid)-mRNA in germinating cotton seeds. Biochemistry **14,** 4375-4379.

Harman, D. (1956). Aging: a theory based on free radical and radiation chemistry. Journal of Gerontology **11,** 298-300.

Harris, B., and Dure, L., 3rd. (1976). Polyadenylylation of stored mRNA in cotton seed germination. Progress in Nucleic Acid Research and Molecular Biology **19,** 113-118.

Hauvermale, A.L., Tuttle, K.M., Takebayashi, Y., Seo, M., and Steber, C.M. (2015). Loss of *Arabidopsis thaliana* seed dormancy is associated with increased accumulation of the GID1 GA hormone receptors. Plant and Cell Physiology **56,** 1773-1785.

He, H.Z., Vidigal, D.D., Snoek, L.B., Schnabel, S., Nijveen, H., Hilhorst, H., and Bentsink, L. (2014). Interaction between parental environment and genotype affects plant and seed performance in Arabidopsis. Journal of Experimental Botany **65,** 6603-6615.

Hemm, M.R., Ruegger, M.O., and Chapple, C. (2003). The Arabidopsis *ref2* mutant is defective in the gene encoding CYP83A1 and shows both phenylpropanoid and glucosinolate phenotypes. Plant Cell **15,** 179-194.

Hershberg, R., and Petrov, D.A. (2008). Selection on codon bias. Annual review of genetics **42,** 287-299.

Hershberg, R., and Petrov, D.A. (2009). General rules for optimal codon choice. PLoS Genetics **5,** e1000556.

Hilhorst, H.W.M., Smitt, A.I., and Karssen, C.M. (1986). Gibberellin-biosynthesis and gibberellin-sensitivity mediated stimulation of seed-germination of *Sisymbrium officinale* by red-light and nitrate. Physiologia Plantarum **67,** 285-290.

Hofmann, N. (2014). Cryptochromes and seed dormancy: the molecular mechanism of blue light inhibition of grain germination. Plant Cell **26,** 846.

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

Hossain, Z., Amyot, L., McGarvey, B., Gruber, M., Jung, J., and Hannoufa, A. (2012). The translation elongation factor eEF-1B beta 1 is involved in cell wall biosynthesis and plant development in *Arabidopsis thaliana*. Plos One **7**.

Howell, K.A., Narsai, R., Carroll, A., Ivanova, A., Lohse, M., Usadel, B., Millar, A.H., and Whelan, J. (2009). Mapping metabolic and transcript temporal switches during germination in rice highlights specific transcription factors and the role of RNA instability in the germination process. Plant Physiology **149,** 961-980.

Hsu, Y.F., Chen, Y.C., Hsiao, Y.C., Wang, B.J., Lin, S.Y., Cheng, W.H., Jauh, G.Y., Harada, J.J., and Wang, C.S. (2014). AtRH57, a DEAD-box RNA helicase, is involved in feedback inhibition of glucose-mediated abscisic acid accumulation during seedling development and additively affects pre-ribosomal RNA processing with high glucose. Plant Journal **77,** 119-135.

Hu, R.B., Zhu, Y.F., Shen, G.X., and Zhang, H. (2014). TAP46 plays a positive role in the ABSCISIC ACID INSENSITIVE5-regulated gene expression in Arabidopsis. Plant Physiology **164,** 721-734.

Hua, Y.M., and Zhou, J.H. (2004). Survival motor neuron protein facilitates assembly of stress granules. FEBS letters **572,** 69-74.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protocol **4,** 44-57.

Huang, X., Schmitt, J., Dorn, L., Griffith, C., Effgen, S., Takao, S., Koornneef, M., and Donohue, K. (2010). The earliest stages of adaptation in an experimental plant population: strong selection on QTLS for seed dormancy. Molecular Ecology **19,** 1335-1351.

Hughes, D.W., and Galau, G.A. (1989). Temporally modular gene expression during cotyledon development. Genes Development **3,** 358-369.

Ihle, j.N., and Dure, l.S. (1970). The Temporal Separation of Transcription and Translation and its Control in Cotton Embryogenesis and Germination. Plant Growth Substances**,** 216-221.

Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. Molecular Biology and Evolution **2,** 13-34.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling.

Science **324,** 218-223.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics **4,** 249-264.

Iskakov, B.K., and Aitkhozhin, M.A. (1979). Proteins of polyribosome-bound informosome of germinating wheat embryos. Molekuliarnaia Biologiia **13,** 1124-1129.

Ito, K., Takahashi, A., Morita, M., Suzuki, T., and Yamamoto, T. (2011). The role of the CNOT1 subunit of the CCR4-NOT complex in mRNA deadenylation and cell viability. Protein Cell **2,** 755-763.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nature reviews. Molecular cell biology **11,** 113-127.

Jacobsen, J.V., Pearce, D.W., Poole, A.T., Pharis, R.P., and Mander, L.N. (2002). Abscisic acid, phaseic acid and gibberellin contents associated with dormancy and germination in barley. Physiol Plant **115,** 428-441.

Jan, C.H., Williams, C.C., and Weissman, J.S. (2014). Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. Science **346,** 1257521.

Jarvis, B.C., Franklan.B, and Cherry, J.H. (1968). Increased template activity of chromatin associated with breaking of seed dormancy by gibberellic Acid. Plant Physiology **S 43,** S18-&.

Jian, O.Y., Shao, X., and Li, J.Y. (2000). Indole-3-glycerol phosphate, a branchpoint of indole-3-acetic acid biosynthesis from the tryptophan biosynthetic pathway in *Arabidopsis thaliana*. Plant Journal **24,** 327-333.

Jiao, Y., and Meyerowitz, E.M. (2010). Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. Molecular Systems Biology **6,** 419.

Jimenez-Lopez, S., Mancera-Martinez, E., Donayre-Torres, A., Rangel, C., Uribe, L., March, S., Jimenez-Sanchez, G., and Sanchez de Jimenez, E. (2011). Expression profile of maize (*Zea mays* L.) embryonic axes during germination: translational regulation of ribosomal protein mRNAs. Plant & Cell Physiology **52,** 1719-1733.

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

Juntawong, P., and Bailey-Serres, J. (2012). Dynamic Light Regulation of Translation Status in *Arabidopsis thaliana*. Frontiers in Plant Science **3,** 66.

Juntawong, P., Girke, T., Bazin, J., and Bailey-Serres, J. (2014). Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis. Proc Natl Acad Sci U S A **111,** E203-212.

Kanai, M., Hayashi, M., Kondo, M., and Nishimura, M. (2013). The plastidic DEAD-box RNA helicase 22, HS3, is essential for plastid functions both in seed development and in seedling growth. Plant & Cell Physiology **54,** 1431-1440.

Karssen, C.M., Brinkhorst-van der Swan, D.L., Breekland, A.E., and Koornneef, M.

(1983). Induction of dormancy during seed development by endogenous abscisic acid: studies on abscisic acid deficient genotypes of *Arabidopsis thaliana* (L.) Heynh. Planta **157,** 158-165.

Kawaguchi, R., Girke, T., Bray, E.A., and Bailey-Serres, J. (2004). Differential mRNA translation contributes to gene regulation under non-stress and dehydration stress conditions in *Arabidopsis thaliana*. The Plant Journal **38,** 823-839.

Kawahara, H., Imai, T., Imataka, H., Tsujimoto, M., Matsumoto, K., and Okano, H. (2008). Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. The Journal of cell biology **181,** 639-653.

Kedersha, N., Cho, M.R., Li, W., Yacono, P.W., Chen, S., Gilks, N., Golan, D.E., and Anderson, P. (2000). Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. The Journal of Cell Biology **151,** 1257-1268.

Kedersha, N., Stoecklin, G., Ayodele, M., Yacono, P., Lykke-Andersen, J., Fitzler, M.J., Scheuner, D., Kaufman, R.J., Golan, D.E., and Anderson, P. (2005). Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. Journal of Cell Biology **169,** 871-884.

Kedersha, N.L., Gupta, M., Li, W., Miller, I., and Anderson, P. (1999). RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. Journal of Cell Biology **147,** 1431-1441.

Kertesz, M., Wan, Y., Mazor, E., Rinn, J.L., Nutter, R.C., Chang, H.Y., and Segal, E. (2010). Genome-wide measurement of RNA secondary structure in yeast. Nature **467,** 103-107.

Khan, A., Garbelli, A., Grossi, S., Florentin, A., Batelli, G., Acuna, T., Zolla, G., Kaye, Y., Paul, L.K., Zhu, J.K., Maga, G., Grafi, G., and Barak, S. (2014). The Arabidopsis STRESS RESPONSE SUPPRESSOR DEAD-box RNA helicases are nucleolar- and chromocenter-localized proteins that undergo stress-mediated relocalization and are involved in epigenetic gene silencing. The Plant Journal **79,** 28-43.

Kim, B., Nesvizhskii, A.I., Rani, P.G., Hahn, S., Aebersold, R., and Ranish, J.A. (2007). The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes. Proc Natl Acad Sci U S A **104,** 16068-16073.

Kimball, S.R., Horetsky, R.L., Ron, D., Jefferson, L.S., and Harding, H.P. (2003). Mammalian stress granules represent sites of accumulation of stalled translation initiation complexes. American Journal of Physiology-Cell Physiology **284,** C273-C284.

Klasens, B.I., Das, A.T., and Berkhout, B. (1998). Inhibition of polyadenylation by stable RNA secondary structure. Nucleic Acids Research **26,** 1870-1876.

Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Development **14,** 2452-2460.

Kong, Q., and Lin, C.L. (2010). Oxidative damage to RNA: mechanisms, consequences, and diseases. Cellular and Molecular Life Sciences : CMLS **67,** 1817-1829.

Koornneef, M., Reuling, G., and Karssen, C.M. (1984). The Isolation and Characterization

of Abscisic-Acid Insensitive Mutants of *Arabidopsis thaliana*. Physiologia Plantarum **61,** 377-383.

Koornneef, M., Jorna, M.L., Derswan, D.L.C.B., and Karssen, C.M. (1982). The isolation of abscisic-acid (ABA) deficient mutants by selection of induced revertants in non-germinating gibberellin sensitive lines of *Arabidopsis thaliana* (L) Heynh.. Theoretical and Applied Genetics **61,** 385-393.

Kozak, M. (2005). Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene **361,** 13-37.

Krishnan, K., Ren, Z., Losada, L., Nierman, W.C., Lu, L.J., and Askew, D.S. (2014). Polysome profiling reveals broad translatome remodeling during endoplasmic reticulum (ER) stress in the pathogenic fungus *Aspergillus fumigatus*. BMC Genomics **15,** 159.

Lageix, S., Lanet, E., Pouch-Pelissier, M.N., Espagnol, M.C., Robaglia, C., Deragon, J.M., and Pelissier, T. (2008). Arabidopsis eIF2alpha kinase GCN2 is essential for growth in stress conditions and is activated by wounding. BMC Plant Biology **8,** 134.

Lai, W.S., Kennington, E.A., and Blackshear, P.J. (2003). Tristetraprolin and its family members can promote the cell-free deadenylation of AU-rich element-containing mRNAs by poly(A) ribonuclease. Molecular and Cellular Biology **23,** 3798-3812.

Lastdrager, J., Hanson, J., and Smeekens, S. (2014). Sugar signals and the control of plant growth and development. Journal of Experimental Botany **65,** 799-807.

Lawford, G.R. (1969). The effect of incubation with puromycin on the dissociation of rat liver ribosomes into active subunits. Biochemical and Biophysical Research Communications **37,** 143-150.

Layat, E., Leymarie, J., El-Maarouf-Bouteau, H., Caius, J., Langlade, N., and Bailly, C. (2014). Translatome profiling in dormant and nondormant sunflower (*Helianthus annuus*) seeds highlights post-transcriptional regulation of germination. The New Phytologist **204,** 864-872.

Le, H., Browning, K.S., and Gallie, D.R. (1998). The phosphorylation state of the wheat translation initiation factors eIF4B, eIF4A, and eIF2 is differentially regulated during seed development and germination. Journal Biology Chemistry **273,** 20084-20089.

Le, H., Browning, K.S., and Gallie, D.R. (2000). The phosphorylation state of poly(A)-binding protein specifies its binding to poly(A) RNA and its interaction with eukaryotic initiation factor (eIF) 4F, eIFiso4F, and eIF4B. Journal Biology Chemistry **275,** 17452-17462.

Le, H., Tanguay, R.L., Balasta, M.L., Wei, C.C., Browning, K.S., Metz, A.M., Goss, D.J., and Gallie, D.R. (1997). Translation initiation factors eIF-iso4G and eIF-4B interact with the poly(A)-binding protein and increase its RNA binding activity. Journal Biology Chemistry **272,** 16247-16255.

Lejeune, F., and Maquat, L.E. (2005). Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Current Opinion in Cell Biology **17,** 309-315.

Leon-Kloosterziel, K.M., van de Bunt, G.A., Zeevaart, J.A., and Koornneef, M. (1996).

Arabidopsis mutants with a reduced seed dormancy. Plant Physiology **110,** 233-240.

Leon, R.G., and Owen, M.D.K. (2003). Regulation of weed seed dormancy through light and temperature interactions. Weed Science **51,** 752-758.

Leung, A.K.L., Calabrese, J.M., and Sharp, P.A. (2006). Quantitative analysis of argonaute protein reveals microRNA-dependent localization to stress granules. Proc Natl Acad Sci U S A **103,** 18125-18130.

Leung, J., Merlot, S., and Giraudat, J. (1997). The Arabidopsis *ABSCISIC ACID-INSENSITIVE2* (*ABI2*) and *ABI1* genes encode homologous protein phosphatases 2C involved in abscisic acid signal transduction. Plant Cell **9,** 759-771.

Li, B., and Foley, M.E. (1996). Transcriptional and Posttranscriptional Regulation of Dormancy-Associated Gene Expression by Afterripening in Wild Oat. Plant Physiology **110,** 1267-1273.

Li, F., Zheng, Q., Vandivier, L.E., Willmann, M.R., Chen, Y., and Gregory, B.D. (2012). Regulatory impact of RNA secondary structure across the Arabidopsis transcriptome. Plant Cell **24,** 4346-4359.

Li, Q., Wang, B.C., Xu, Y., and Zhu, Y.X. (2007). Systematic studies of 12S seed storage protein accumulation and degradation patterns during Arabidopsis seed maturation and early seedling germination stages. Journal of Biochemistry and Molecular Biology **40,** 373-381.

Li, W., Ma, M., Feng, Y., Li, H., Wang, Y., Ma, Y., Li, M., An, F., and Guo, H. (2015). EIN2-directed translational regulation of ethylene signaling in Arabidopsis. Cell **163,** 670-683.

Lin, P.C., Hwang, S.G., Endo, A., Okamoto, M., Koshiba, T., and Cheng, W.H. (2007). Ectopic expression of *ABSCISIC ACID 2*/*GLUCOSE INSENSITIVE 1* in Arabidopsis promotes seed dormancy and stress tolerance. Plant Physiology **143,** 745-758.

Lin, S.Y., Chen, P.W., Chuang, M.H., Juntawong, P., Bailey-Serres, J., and Jauh, G.Y. (2014). Profiling of translatomes of in vivo-grown pollen tubes reveals genes with roles in micropylar guidance during pollination in Arabidopsis. Plant Cell **26,** 602-618.

Linder, P., and Jankowsky, E. (2011). From unwinding to clamping - the DEAD box RNA helicase family. Nature reviews. Molecular Cell Biology **12,** 505-516.

Liu, A.H., Gao, F., Kanno, Y., Jordan, M.C., Kamiya, Y., Seo, M., and Ayele, B.T. (2013). Regulation of Wheat Seed Dormancy by After-Ripening Is Mediated by Specific Transcriptional Switches That Induce Changes in Seed Hormone Metabolism and Signaling. Plos One **8**.

Liu, H.Y., Chiang, Y.C., Pan, J., Chen, J., Salvadore, C., Audino, D.C., Badarinarayana, V., Palaniswamy, V., Anderson, B., and Denis, C.L. (2001). Characterization of CAF4 and CAF16 reveals a functional connection between the CCR4-NOT complex and a subset of SRB proteins of the RNA polymerase II holoenzyme. Journal Biology Chemistry **276,** 7541-7548.

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., Arenas-Huertero, C., and Chua, N.H. (2012). Genome-wide analysis uncovers regulation of long intergenic noncoding

RNAs in Arabidopsis. Plant Cell **24,** 4333-4345.

Liu, J.D., Rivas, F.V., Wohlschlegel, J., Yates, J.R., Parker, R., and Hannon, G.J. (2005). A role for the P-body component GW182 in microRNA function. Nature Cell Biology **7,** 1261-1266.

Liu, J.L., and Gall, J.G. (2007). U bodies are cytoplasmic structures that contain uridine-rich small nuclear ribonucleoproteins and associate with P bodies. Proc Natl Acad Sci U S A **104,** 11655-11659.

Liu, M.J., Wu, S.H., and Chen, H.M. (2012). Widespread translational control contributes to the regulation of Arabidopsis photomorphogenesis. Molecular Systems Biology **8,** 566.

Liu, M.J., Wu, S.H., Wu, J.F., Lin, W.D., Wu, Y.C., Tsai, T.Y., Tsai, H.L., and Wu, S.H. (2013). Translational landscape of photomorphogenic Arabidopsis. Plant Cell **25,** 3699-3710.

Liu, P.P., Montgomery, T.A., Fahlgren, N., Kasschau, K.D., Nonogaki, H., and Carrington, J.C. (2007). Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. Plant Journal **52,** 133-146.

Liu, X., Zhang, H., Zhao, Y., Feng, Z., Li, Q., Yang, H.Q., Luan, S., Li, J., and He, Z.H. (2013). Auxin controls seed dormancy through stimulation of abscisic acid signaling by inducing ARF-mediated ABI3 activation in Arabidopsis. Proc Natl Acad Sci U S A **110,** 15485-15490.

Liu, Y., and Bassham, D.C. (2012). Autophagy: pathways for self-eating in plant cells. Annual Review of Plant Biology **63,** 215-237.

Lorkovic, Z.J. (2009). Role of plant RNA-binding proteins in development, stress response and genome organization. Trends Plant Science **14,** 229-236.

Luo, R.X., and Dean, D.C. (1999). Chromatin remodeling and transcriptional regulation. Journal of the National Cancer Institute **91,** 1288-1294.

Maia, J., Dekkers, B.J., Provart, N.J., Ligterink, W., and Hilhorst, H.W. (2011). The re-establishment of desiccation tolerance in germinated *Arabidopsis thaliana* seeds and its associated transcriptome. Plos One **6,** e29123.

Maia, J., Dekkers, B.J., Dolle, M.J., Ligterink, W., and Hilhorst, H.W. (2014). Abscisic acid (ABA) sensitivity regulates desiccation tolerance in germinated Arabidopsis seeds. The New Phytologist **203,** 81-93.

Mao, Z., and Sun, W. (2015). Arabidopsis seed-specific vacuolar aquaporins are involved in maintaining seed longevity under the control of ABSCISIC ACID INSENSITIVE 3. Journal of Experimental Botany **66,** 4781-4794.

Marcus, A., and Feeley, J. (1964). Activation of protein synthesis in the imbibition phase of seed germination. Proc Natl Acad Sci U S A **51,** 1075-1079.

Marcus, A., and Feeley, J. (1965). Protein synthesis in imbibed seeds .2. polysome formation during imbibition. Journal Biology Chemistry **240,** 1675-&.

Marcus, A., Feeley, J., and Volcani, T. (1966). Protein synthesis in imbibed seeds .3. kinetics of amino acid incorporation ribosome activation and polysome formation.

Plant Physiology **41,** 1167-&.

Marre, E. (1967). Ribosome and enzyme changes during maturation and germination of the castor bean seed. Current Topics in Developmental Biology **2,** 75-105.

Matakiadis, T., Alboresi, A., Jikumaru, Y., Tatematsu, K., Pichon, O., Renou, J.P., Kamiya, Y., Nambara, E., and Truong, H.N. (2009). The Arabidopsis abscisic acid catabolic gene *CYP707A2* plays a key role in nitrate control of seed dormancy. Plant Physiology **149,** 949-960.

Matharu, N., and Ahituv, N. (2015). Minor loops in major folds: enhancer-promoter looping, chromatin restructuring, and their association with transcriptional regulation and disease. PLoS Genetics **11,** e1005640.

Mathews, M., Sonenberg, N., and Hershey, J. (2007). Origins and principles of translational control. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Mauger, D.M., and Weeks, K.M. (2010). Toward global RNA structure analysis. Nature Biotechnology **28,** 1178-1179.

Meimoun, P., Mordret, E., Langlade, N.B., Balzergue, S., Arribat, S., Bailly, C., and El-Maarouf-Bouteau, H. (2014). Is gene transcription involved in seed dry after-ripening? Plos One **9**.

Meinke, D.W. (1995). Molecular-Genetics of Plant Embryogenesis. Annual Review of Plant Biology **46,** 369-394.

Merchante, C., Brumos, J., Yun, J., Hu, Q., Spencer, K.R., Enriquez, P., Binder, B.M., Heber, S., Stepanova, A.N., and Alonso, J.M. (2015). Gene-specific translation regulation mediated by the hormone-signaling molecule EIN2. Cell **163,** 684-697.

Mi, H., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2013). Large-scale gene function analysis with the PANTHER classification system. Nature Protocol **8,** 1551-1566.

Miller, J.E., and Reese, J.C. (2012). Ccr4-Not complex: the control freak of eukaryotic cells. Critical Reviews in Biochemistry and Molecular Biology **47,** 315-333.

Missra, A., Ernest, B., Lohoff, T., Jia, Q., Satterlee, J., Ke, K., and von Arnim, A.G. (2015). The circadian clock modulates global daily cycles of mRNA ribosome loading. Plant Cell **27,** 2582-2599.

Mitchell, R.C., and Villiers, T.A. (1972). Polysome formation in light-controlled dormancy. Plant Physiology **50,** 671-674.

Mittler, R., Vanderauwera, S., Suzuki, N., Miller, G., Tognetti, V.B., Vandepoele, K., Gollery, M., Shulaev, V., and Van Breusegem, F. (2011). ROS signaling: the new wave? Trends in Plant Science **16,** 300-309.

Miura, K., Lee, J., Jin, J.B., Yoo, C.Y., Miura, T., and Hasegawa, P.M. (2009). Sumoylation of ABI5 by the Arabidopsis SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling. Proc Natl Acad Sci U S A **106,** 5418-5423.

Morel, J.B., Godon, C., Mourrain, P., Beclin, C., Boutet, S., Feuerbach, F., Proux, F., and Vaucheret, H. (2002). Fertile hypomorphic ARGONAUTE (AGO1) mutants impaired in post-transcriptional gene silencing and virus resistance. Plant Cell **14,** 629-639.

Murrell, A.M., Bockamp, E.O., Gottgens, B., Chan, Y.S., Cross, M.A., Heyworth, C.M., and Green, A.R. (1995). Discordant regulation of SCL/TAL-1 mRNA and protein during erythroid differentiation. Oncogene **11,** 131-139.

Mustroph, A., Zanetti, M.E., Jang, C.J., Holtan, H.E., Repetti, P.P., Galbraith, D.W., Girke, T., and Bailey-Serres, J. (2009). Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. Proc Natl Acad Sci U S A **106,** 18843-18848.

Nair, L.A., and Koundal, K.R. (1993). Changes in Polysomes and RNA in Developing Seeds of Chickpea (*Cicer-arietinum* L). Journal of Plant Biochemistry and Biotechnology **2,** 29-33.

Nakabayashi, K., Okamoto, M., Koshiba, T., Kamiya, Y., and 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,** 697-709.

Narsai, R., Law, S.R., Carrie, C., Xu, L., and Whelan, J. (2011). In-depth temporal transcriptome profiling reveals a crucial developmental switch with roles for RNA processing and organelle metabolism that are essential for germination in Arabidopsis. Plant Physiology **157,** 1342-1362.

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

Nasertorabi, F., Batisse, C., Diepholz, M., Suck, D., and Bottcher, B. (2011). Insights into the structure of the CCR4-NOT complex by electron microscopy. FEBS letters **585,** 2182-2186.

Nayak, N.R., Putnam, A.A., Addepalli, B., Lowenson, J.D., Chen, T., Jankowsky, E., Perry, S.E., Dinkins, R.D., Limbach, P.A., Clarke, S.G., and Downie, A.B. (2013). An Arabidopsis ATP-dependent, DEAD-box RNA helicase loses activity upon IsoAsp formation but is restored by PROTEIN ISOASPARTYL METHYLTRANSFERASE. Plant Cell **25,** 2573-2586.

Nelson, D.C., Scaffidi, A., Dun, E.A., Waters, M.T., Flematti, G.R., Dixon, K.W., Beveridge, C.A., Ghisalberti, E.L., and Smith, S.M. (2011). F-box protein MAX2 has dual roles in karrikin and strigolactone signaling in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A **108,** 8897-8902.

Nguyen, T.P., and Bentsink, L. (2015). Seed dormancy, longevity and their adaptation. In molecular mechanisms in plant adaptation, R. Laitinen, ed (Wiley-Blackwell), pp. 61-74.

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

Nicolai, M., Roncato, M.A., Canoy, A.S., Rouquie, D., Sarda, X., Freyssinet, G., and Robaglia, C. (2006). Large-scale analysis of mRNA translation states during sucrose

starvation in arabidopsis cells identifies cell proliferation and chromatin structure as targets of translational control. Plant Physiology **141,** 663-673.

Nishimura, N., Kitahata, N., Seki, M., Narusaka, Y., Narusaka, M., Kuromori, T., Asami, T., Shinozaki, K., and Hirayama, T. (2005). Analysis of ABA Hypersensitive Germination2 revealed the pivotal functions of PARN in stress response in Arabidopsis. Plant Journal **44,** 972-984.

Nishimura, N., Sarkeshik, A., Nito, K., Park, S.Y., Wang, A., Carvalho, P.C., Lee, S., Caddell, D.F., Cutler, S.R., Chory, J., Yates, J.R., and Schroeder, J.I. (2010). PYR/PYL/RCAR family members are major in-vivo ABI1 protein phosphatase 2C-interacting proteins in Arabidopsis. The Plant Journal **61,** 290-299.

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

Normanly, J., Cohen, J.D., and Fink, G.R. (1993). *Arabidopsis thaliana* auxotrophs reveal a tryptophan-independent biosynthetic-pathway for indole-3-acetic-acid. Proc Natl Acad Sci U S A **90,** 10355-10359.

Oberstrass, F.C., Auweter, S.D., Erat, M., Hargous, Y., Henning, A., Wenter, P., Reymond, L., Amir-Ahmady, B., Pitsch, S., Black, D.L., and Allain, F.H. (2005). Structure of PTB bound to RNA: specific binding and implications for splicing regulation. Science **309,** 2054-2057.

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

Ottoleng.S, Comi, P., Giglioni, B., Gianni, A.M., and Guidotti, G.G. (1973). Dissociation of Single Ribosomes as a Preliminary Step for Their Participation in Protein-Synthesis. European Journal of Biochemistry **33,** 227-232.

Palade, G.E. (1955). A small particulate component of the cytoplasm. The Journal of Biophysical and Biochemical Cytology **1,** 59-&.

Panasenko, O., Landrieux, E., Feuermann, M., Finka, A., Paquet, N., and Collart, M. (2006a). The yeast Ccr4-Not complex controls ubiquitination of the nascent-associated polypeptide complex. FEBS Journal **273,** 77-77.

Panasenko, O., Landrieux, E., Feuermann, M., Finka, A., Paquet, N., and Collart, M.A. (2006b). The yeast Ccr4-not complex controls ubiquitination of the nascent-associated polypeptide (NAC-EGD) complex. Journal Biology Chemistry **281,** 31389-31398.

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

Parker, R., and Song, H.W. (2004). The enzymes and control of eukaryotic mRNA turnover. Nature Structural & Molecular Biology **11,** 121-127.

Perez, I., Lin, C.H., McAfee, J.G., and Patton, J.G. (1997). Mutation of PTB binding sites causes misregulation of alternative 3' splice site selection in vivo. RNA **3,** 764-778.

Perez, V.I., Bokov, A., Van Remmen, H., Mele, J., Ran, Q.T., Ikeno, Y., and Richardson,

A. (2009). Is the oxidative stress theory of aging dead? BBA General Subjects **1790,** 1005-1014.

Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. Genes Development **20,** 2922-2936.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 microRNA in human cells. Science **309,** 1573-1576.

Piques, M., Schulze, W.X., Hohne, M., Usadel, B., Gibon, Y., Rohwer, J., and Stitt, M. (2009). Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in Arabidopsis. Molecular Systems Biology **5,** 314.

Polesskaya, A., Cuvellier, S., Naguibneva, I., Duquet, A., Moss, E.G., and Harel-Bellan, A. (2007). Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency. Genes Development **21,** 1125-1138.

Popescu, S.C., and Tumer, N.E. (2004). Silencing of ribosomal protein L3 genes in N. tabacum reveals coordinate expression and significant alterations in plant growth, development and ribosome biogenesis. The Plant Journal **39,** 29-44.

Poulson, R., and Beevers, L. (1973). RNA metabolism during the development of cotyledons of *Pisum staivum* L. Biochimica et biophysica acta **308,** 381-389.

Qu, X., Wen, J.D., Lancaster, L., Noller, H.F., Bustamante, C., and Tinoco, I., Jr. (2011). The ribosome uses two active mechanisms to unwind messenger RNA during translation. Nature **475,** 118-121.

Quint, M., Drost, H.G., Gabel, A., Ullrich, K.K., Bonn, M., and Grosse, I. (2012). A transcriptomic hourglass in plant embryogenesis. Nature **490,** 98-101.

Raghavendra, A.S., Gonugunta, V.K., Christmann, A., and Grill, E. (2010). ABA perception and signalling. Trends Plant Science **15,** 395-401.

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

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

Rajjou, L., Duval, M., Gallardo, K., Catusse, J., Bally, J., Job, C., and Job, D. (2012). Seed germination and vigor. Annual Review of Plant Biology **63,** 507-533.

Raker, V.A., Mironov, A.A., Gelfand, M.S., and Pervouchine, D.D. (2009). Modulation of alternative splicing by long-range RNA structures in Drosophila. Nucleic Acids Research **37,** 4533-4544.

Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussov, S., Albu, M., Zheng, H., Yang, A., Na, H., Irimia, M., Matzat, L.H., Dale, R.K., Smith, S.A., Yarosh, C.A., Kelly, S.M., Nabet, B., Mecenas, D., Li, W., Laishram, R.S., Qiao, M., Lipshitz, H.D., Piano, F., Corbett, A.H., Carstens, R.P., Frey, B.J., Anderson, R.A., Lynch, K.W., Penalva, L.O., Lei, E.P., Fraser, A.G., Blencowe,

B.J., Morris, Q.D., and Hughes, T.R. (2013). A compendium of RNA-binding motifs for decoding gene regulation. Nature **499,** 172-177.

Ren, M., Venglat, P., Qiu, S., Feng, L., Cao, Y., Wang, E., Xiang, D., Wang, J., Alexander, D., Chalivendra, S., Logan, D., Mattoo, A., Selvaraj, G., and Datla, R. (2012). Target of rapamycin signaling regulates metabolism, growth, and life span in Arabidopsis. Plant Cell **24,** 4850-4874.

Reverdatto, S.V., Dutko, J.A., Chekanova, J.A., Hamilton, D.A., and Belostotsky, D.A. (2004). mRNA deadenylation by PARN is essential for embryogenesis in higher plants. RNA **10,** 1200-1214.

Reynoso, M.A., Juntawong, P., Lancia, M., Blanco, F.A., Bailey-Serres, J., and Zanetti, M.E. (2015). Translating Ribosome Affinity Purification (TRAP) followed by RNA sequencing technology (TRAP-SEQ) for quantitative assessment of plant translatomes. Methods of Molecular Biology **1284,** 185-207.

Rhazi, L., Cazalis, R., Lemelin, E., and Aussenac, T. (2003). Changes in the glutathione thiol-disulfide status during wheat grain development. Plant Physiology Biochemistry **41,** 895-902.

Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases. Gene Development **23,** 1247-1269.

Righetti, K., Vu, J.L., Pelletier, S., Vu, B.L., Glaab, E., Lalanne, D., Pasha, A., Patel, R.V., Provart, N.J., Verdier, J., Leprince, O., and Buitink, J. (2015). Inference of Longevity-Related Genes from a Robust Coexpression Network of Seed Maturation Identifies Regulators Linking Seed Storability to Biotic Defense-Related Pathways. Plant Cell **27,** 2692-2708.

Rinaldi, A.M., and Monroy, A. (1969). Polyribosome formation and RNA synthesis in the early post-fertilization stages of the sea urchin egg. Developmental Biology **19,** 73-86.

Robaglia, C., Thomas, M., and Meyer, C. (2012). Sensing nutrient and energy status by SnRK1 and TOR kinases. Current Opinion in Plant Biology **15,** 301-307.

Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J.S. (2014). Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. Nature **505,** 701-705.

Ruijter, J.M., Ramakers, C., Hoogaars, W.M., Karlen, Y., Bakker, O., van den Hoff, M.J., and Moorman, A.F. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Research **37,** e45.

Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V., and Quackenbush, J. (2003). TM4: A free, open-source system for microarray data management and analysis. Biotechniques **34,** 374.

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

Sami-Subbu, R., Choi, S.B., Wu, Y., Wang, C., and Okita, T.W. (2001). Identification of a cytoskeleton-associated 120 kDa RNA-binding protein in developing rice seeds. Plant Molecular Biology **46,** 79-88.

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

Sano, N., Masaki, S., Tanabata, T., Yamada, T., Hirasawa, T., Kashiwagi, M., and Kanekatsu, M. (2013). RNA-binding proteins associated with desiccation during seed development in rice. Biotechnology Letters **35,** 1945-1952.

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

Sassone-Corsi, P. (2002). Unique chromatin remodeling and transcriptional regulation in spermatogenesis. Science **296,** 2176-2178.

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

Schott, J., Reitter, S., Philipp, J., Haneke, K., Schafer, H., and Stoecklin, G. (2014). Translational regulation of Specific mRNAs controls feedback inhibition and survival during macrophage activation. PLoS Genetics **10**.

Sen, G.L., and Blau, H.M. (2005). Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. Nature Cell Biology **7,** 633-U628.

Shabalina, S.A., Spiridonov, N.A., and Kashina, A. (2013). Sounds of silence: synonymous nucleotides as a key to biological regulation and complexity. Nucleic Acids Research **41,** 2073-2094.

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

Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science **300,** 805-808.

Shu, K., Zhang, H., Wang, S., Chen, M., Wu, Y., Tang, S., Liu, C., Feng, Y., Cao, X., and Xie, Q. (2013). ABI4 regulates primary seed dormancy by regulating the biogenesis of abscisic acid and gibberellins in arabidopsis. PLoS Genetics **9,** e1003577.

Singh, R., Valcarcel, J., and Green, M.R. (1995). Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science **268,** 1173-1176.

Siwecka, M.A., Golaszew.T, and Szarkows.Jw. (1973). Decrease of ribosome bound ribonucleolytic activity during rye embryo germination. Bulletin of the Polish Academy of Sciences **21,** 171-175.

Smaczniak, C., Li, N., Boeren, S., America, T., van Dongen, W., Goerdayal, S.S., de Vries, S., Angenent, G.C., and Kaufmann, K. (2012). Proteomics-based identification

of low-abundance signaling and regulatory protein complexes in native plant tissues. Nature Protocol **7,** 2144-2158.

Smeekens, S., Ma, J., Hanson, J., and Rolland, F. (2010). Sugar signals and molecular networks controlling plant growth. Current Opinion in Plant Biology **13,** 274-279.

Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology **3,** Article3.

Smyth, G.K. (2005). Limma: linear models for microarray data. (New York: Springer).

Sondheimer, E., Tzou, D.S., and Galson, E.C. (1968). Abscisic acid levels and seed dormancy. Plant Physiology **43,** 1443-1447.

Sorenson, R., and Bailey-Serres, J. (2014). Selective mRNA sequestration by OLIGOURIDYLATE-BINDING PROTEIN 1 contributes to translational control during hypoxia in Arabidopsis. Proc Natl Acad Sci U S A **111,** 2373-2378.

Sorenson, R., and Bailey-Serres, J. (2015). Rapid immunopurification of ribonucleoprotein complexes of plants. Methods in Molecular Biology **1284,** 209-219.

Spassov, D.S., and Jurecic, R. (2003). The PUF family of RNA-binding proteins: Does evolutionarily conserved structure equal conserved function? IUBMB Life **55,** 359-366.

Speigel, S., and Marcus, A. (1975). Polyribosome formation in early wheat embryo germination independent of either transcription or polyadenylation. Nature **256,** 228-230.

Sreenivasulu, N., Usadel, B., Winter, A., Radchuk, V., Scholz, U., Stein, N., Weschke, W., Strickert, M., Close, T.J., Stitt, M., Graner, A., and Wobus, U. (2008). Barley grain maturation and germination: Metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. Plant Physiology **146,** 1738-1758.

Staszak, A.M., and Pawlowski, T.A. (2014). Proteomic analysis of embryogenesis and the acquisition of seed dormancy in Norway maple (*Acer platanoides* L.). International Journal of Molecular Sciences **15,** 10868-10891.

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

Steiner, F.A., and Henikoff, S. (2015). Cell type-specific affinity purification of nuclei for chromatin profiling in whole animals. Methods in Molecular Biology **1228,** 3-14.

Steiner, F.A., Talbert, P.B., Kasinathan, S., Deal, R.B., and Henikoff, S. (2012). Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling. Genome Research **22,** 766-777.

Stoecklin, G., Stubbs, T., Kedersha, N., Wax, S., Rigby, W.F., Blackwell, T.K., and Anderson, P. (2004). MK2-induced tristetraprolin:14-3-3 complexes prevent stress granule association and ARE-mRNA decay. The EMBO Journal **23,** 1313-1324.

Subbiah, V., and Reddy, K.J. (2010). Interactions between ethylene, abscisic acid and cytokinin during germination and seedling establishment in Arabidopsis. Journal

Biosciences **35,** 451-458.

Subramanian, V. (1978). Polyribosome formation during early germination of bean embryonic axes. Indian Journal of Biochemistry & Biophysics **15,** 235-237.

Sugawara, S., Hishiyama, S., Jikumaru, Y., Hanada, A., Nishimura, T., Koshiba, T., Zhao, Y., Kamiya, Y., and Kasahara, H. (2009). Biochemical analyses of indole-3-acetaldoxime-dependent auxin biosynthesis in Arabidopsis. Proc Natl Acad Sci U S A **106,** 5430-5435.

Sun, X., Yang, Q., and Xia, X. (2013). An improved implementation of effective number of codons (nc). Molecular Biology and Evolution **30,** 191-196.

Suzuki, A., Igarashi, K., Aisaki, K., Kanno, J., and Saga, Y. (2010). NANOS2 interacts with the CCR4-NOT deadenylation complex and leads to suppression of specific RNAs. Proc Natl Acad Sci U S A **107,** 3594-3599.

Suzuki, A., Saba, R., Miyoshi, K., Morita, Y., and Saga, Y. (2012). Interaction between NANOS2 and the CCR4-NOT deadenylation complex Is essential for male germ cell development in mouse. Plos One **7**.

Suzuki, Y., and Minamikawa, T. (1985). On the role of stored mRNA in protein synthesis in embryonic axes of germinating vigna unguiculata seeds. Plant Physiology **79,** 327-331.

Swindell, W.R., Remmer, H.A., Sarkar, M.K., Xing, X., Barnes, D.H., Wolterink, L., Voorhees, J.J., Nair, R.P., Johnston, A., Elder, J.T., and Gudjonsson, J.E. (2015). Proteogenomic analysis of psoriasis reveals discordant and concordant changes in mRNA and protein abundance. Genome Medicine **7,** 86.

Tanaka, M., Chock, P.B., and Stadtman, E.R. (2007). Oxidized messenger RNA induces translation errors. Proc Natl Acad Sci U S A **104,** 66-71.

Thomas, J., Palusa, S.G., Prasad, K.V., Ali, G.S., Surabhi, G.K., Ben-Hur, A., Abdel-Ghany, S.E., and Reddy, A.S. (2012). Identification of an intronic splicing regulatory element involved in auto-regulation of alternative splicing of SCL33 pre-mRNA. The Plant Journal **72,** 935-946.

Thompson, J.E., Hopkins, M.T., Taylor, C., and Wang, T.W. (2004). Regulation of senescence by eukaryotic translation initiation factor 5A: implications for plant growth and development. Trends in Plant Science **9,** 174-179.

Toh, S., Kamiya, Y., Kawakami, N., Nambara, E., McCourt, P., and Tsuchiya, Y. (2012). Thermoinhibition uncovers a role for strigolactones in Arabidopsis seed germination. Plant and Cell Physiology **53,** 107-117.

Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J.J., Denis, C.L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. Cell **104,** 377-386.

Ui, A., Nagaura, Y., and Yasui, A. (2015). Transcriptional elongation factor ENL phosphorylated by ATM recruits polycomb and switches off transcription for DSB repair. Molecular Cell **58,** 468-482.

Valleriani, A., Zhang, G., Nagar, A., Ignatova, Z., and Lipowsky, R. (2011). Length-

dependent translation of messenger RNA by ribosomes. Physical Review. E, Statistical, nonlinear, and soft matter physics **83,** 042903.

Van Der Kelen, K., Beyaert, R., Inze, D., and De Veylder, L. (2009). Translational control of eukaryotic gene expression. Critical Reviews in Biochemistry and Molecular Biology **44,** 143-168.

Van Lijsebettens, M., Vanderhaeghen, R., De Block, M., Bauw, G., Villarroel, R., and Van Montagu, M. (1994). An S18 ribosomal protein gene copy at the Arabidopsis PFL locus affects plant development by its specific expression in meristems. The EMBO Journal **13,** 3378-3388.

van Zanten, M., Koini, M.A., Geyer, R., Liu, Y.X., Brambilla, V., Bartels, D., Koornneef, M., Fransz, P., and Soppe, W.J.J. (2011). Seed maturation in *Arabidopsis thaliana* is characterized by nuclear size reduction and increased chromatin condensation. Proc Natl Acad Sci U S A **108,** 20219-20224.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biology **3**.

Velez-Bermudez, I.C., and Schmidt, W. (2014). The conundrum of discordant protein and mRNA expression. Are plants special? Frontiers in Plant Science **5,** 619.

Verdier, J., Lalanne, D., Pelletier, S., Torres-Jerez, I., Righetti, K., Bandyopadhyay, K., Leprince, O., Chatelain, E., Vu, B.L., Gouzy, J., Gamas, P., Udvardi, M.K., and Buitink, J. (2013). A regulatory network-based approach dissects late maturation processes related to the acquisition of desiccation tolerance and longevity of *Medicago truncatula* seeds. Plant Physiology **163,** 757-774.

Vessey, J.P., Vaccani, A., Xie, Y., Dahm, R., Karra, D., Kiebler, M.A., and Macchi, P. (2006). Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules. The Journal of Neuroscience **26,** 6496-6508.

Voegele, A., Linkies, A., Muller, K., and Leubner-Metzger, G. (2011). Members of the gibberellin receptor gene family *GID1* (*GIBBERELLIN INSENSITIVE DWARF1*) play distinct roles during *Lepidium sativum* and *Arabidopsis thaliana* seed germination. Journal of Experimental Botany **62,** 5131-5147.

Vragovic, K., Sela, A., Friedlander-Shani, L., Fridman, Y., Hacham, Y., Holland, N., Bartom, E., Mockler, T.C., and Savaldi-Goldstein, S. (2015). Translatome analyses capture of opposing tissue-specific brassinosteroid signals orchestrating root meristem differentiation. Proc Natl Acad Sci U S A **112,** 923-928.

Wang, B., Chu, J.F., Yu, T.Y., Xu, Q., Sun, X.H., Yuan, J., Xiong, G.S., Wang, G.D., Wang, Y.H., and Li, J.Y. (2015). Tryptophan-independent auxin biosynthesis contributes to early embryogenesis in Arabidopsis. Proc Natl Acad Sci U S A **112,** 4821-4826.

Warf, M.B., and Berglund, J.A. (2010). Role of RNA structure in regulating pre-mRNA splicing. Trends in Biochemical Sciences **35,** 169-178.

Waterworth, W.M., Masnavi, G., Bhardwaj, R.M., Jiang, Q., Bray, C.M., and West, C.E. (2010). A plant DNA ligase is an important determinant of seed longevity. Plant

Journal **63,** 848-860.

Webster, C., Gaut, R.L., Browning, K.S., Ravel, J.M., and Roberts, J.K.M. (1991). Hypoxia enhances phosphorylation of eukaryotic initiation factor-4a in maize root-tips. Journal Biology Chemistry **266,** 23341-23346.

Wilczynska, A., Aigueperse, C., Kress, M., Dautry, F., and Weil, D. (2005). The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. Journal of Cell Science **118,** 981-992.

Williams, C.C., Jan, C.H., and Weissman, J.S. (2014). Targeting and plasticity of mitochondrial proteins revealed by proximity-specific ribosome profiling. Science **346,** 748-751.

Wilusz, C.J., Wormington, M., and Peltz, S.W. (2001). The cap-to-tail guide to mRNA turnover. Nature Reviews Molecular Cell Biology **2,** 237-246.

Woese, C.R. (1961). Unusual Ribosome Particles Occurring during Spore Germination. Journal of Bacteriology **82,** 695-&.

Wood, Z.A., Schroder, E., Harris, J.R., and Poole, L.B. (2003). Structure, mechanism and regulation of peroxiredoxins. Trends in Biochemical Sciences **28,** 32-40.

Wright, A.D., Sampson, M.B., Neuffer, M.G., Michalczuk, L., Slovin, J.P., and Cohen, J.D. (1991). Indole-3-acetic-acid biosynthesis in the mutant maize orange pericarp, a Tryptophan Auxotroph. Science **254,** 998-1000.

Wuitschick, J.D., and Karrer, K.M. (1999). Analysis of genomic G + C content, codon usage, initiator codon context and translation termination sites in *Tetrahymena thermophila*. The Journal of Eukaryotic Microbiology **46,** 239-247.

Xia, X., MacKay, V., Yao, X., Wu, J., Miura, F., Ito, T., and Morris, D.R. (2011). Translation initiation: a regulatory role for poly(A) tracts in front of the AUG codon in *Saccharomyces cerevisiae*. Genetics **189,** 469-478.

Xiang, Y., Nakabayashi, K., Ding, J., He, F., Bentsink, L., and Soppe, W.J. (2014). *REDUCED DORMANCY5* encodes a protein phosphatase 2C that is required for seed dormancy in Arabidopsis. Plant Cell **26,** 4362-4375.

Xu, J., and Chua, N.H. (2009). Arabidopsis Decapping 5 is required for mRNA decapping, P-body formation, and translational repression during postembryonic development. Plant Cell **21,** 3270-3279.

Yamauchi, Y., Ogawa, M., Kuwahara, A., Hanada, A., Kamiya, Y., and Yamaguchi, S. (2004). Activation of gibberllin biosynthesis and response pathways by low temperature during imbibition of *Arabidopsis thaliana* seeds. Plant Cell **16,** 367-378.

Yang, C., Zhang, C., Dittman, J.D., and Whitham, S.A. (2009). Differential requirement of ribosomal protein S6 by plant RNA viruses with different translation initiation strategies. Virology **390,** 163-173.

Yanguez, E., Castro-Sanz, A.B., Fernandez-Bautista, N., Oliveros, J.C., and Castellano, M.M. (2013). Analysis of genome-wide changes in the translatome of Arabidopsis seedlings subjected to heat stress. Plos One **8,** e71425.

Ye, Y., Gong, Z., Lu, X., Miao, D., Shi, J., Lu, J., and Zhao, Y. (2015). *GermoStatin*

Resistance locus 1 encodes a PHD domain protein involved in auxin-mediated seed dormancy and germination. The Plant Journal **85**, 3-15.

Yeakley, J.M., Morfin, J.P., Rosenfeld, M.G., and Fu, X.D. (1996). A complex of nuclear proteins mediates SR protein binding to a purine-rich splicing enhancer. Proc Natl Acad Sci U S A **93,** 7582-7587.

Yu, Y., Guo, G., Lv, D., Hu, Y., Li, J., Li, X., and Yan, Y. (2014). Transcriptome analysis during seed germination of elite Chinese bread wheat cultivar Jimai 20. BMC Plant Biology **14,** 20.

Zanetti, M.E., Chang, I.F., Gong, F., Galbraith, D.W., and Bailey-Serres, J. (2005). Immunopurification of polyribosomal complexes of Arabidopsis for global analysis of gene expression. Plant Physiology **138,** 624-635.

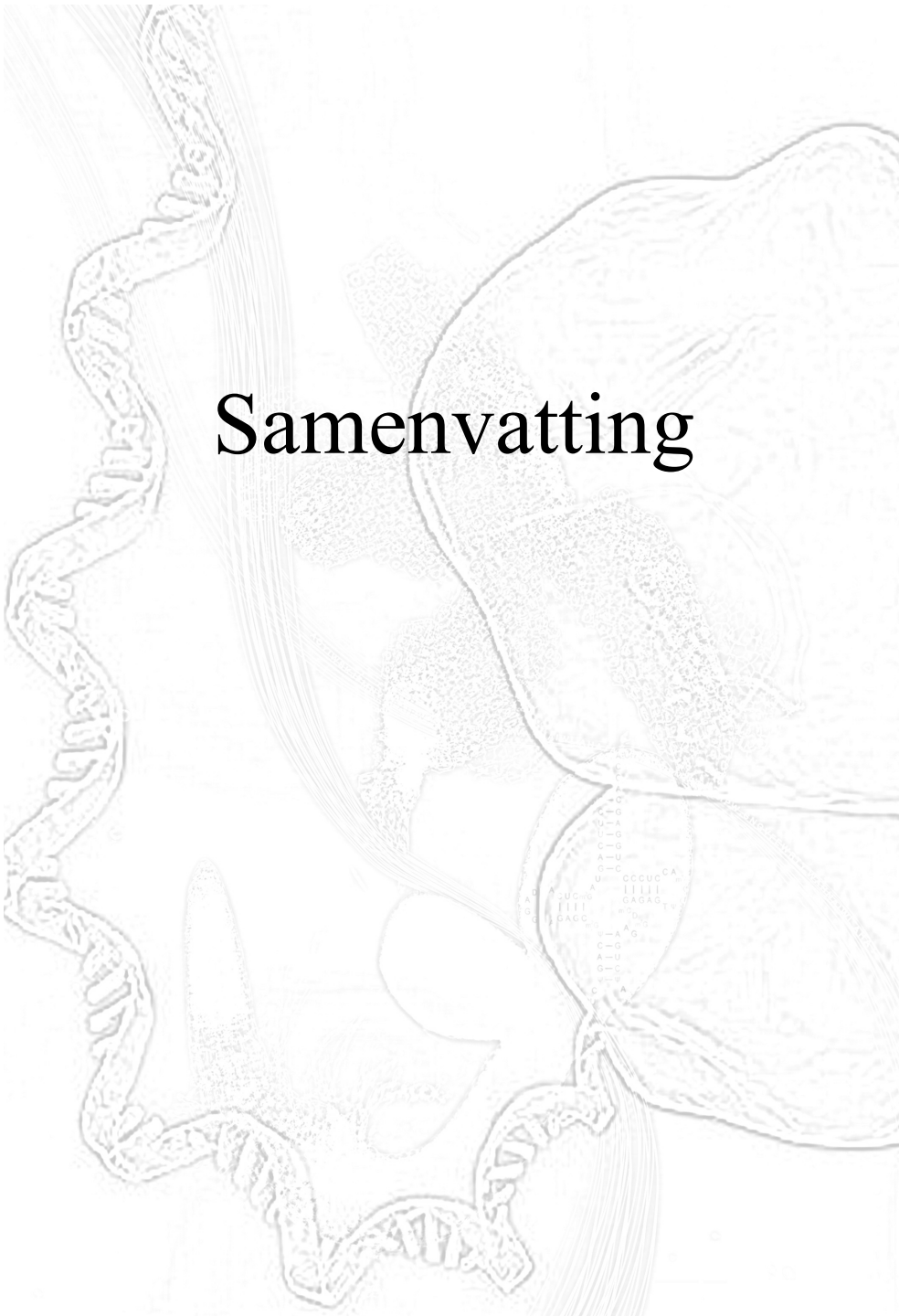
Zarudnaya, M.I., Kolomiets, I.M., Potyahaylo, A.L., and Hovorun, D.M. (2003). Downstream elements of mammalian pre-mRNA polyadenylation signals: primary, secondary and higher-order structures. Nucleic Acids Research **31,** 1375-1386.

Zhang, J.Y., Harnpicharnchai, P., Jakovljevic, J., Tang, L., Guo, Y.R., Oeffinger, M., Rout, M.P., Hiley, S.L., Hughes, T., and Woolford, J.L. (2007). Assembly factors Rpf2 and Rrs1 recruit 5S rRNA and ribosomal proteins RPL5 and RPL11 into nascent ribosomes. Gene Development **21,** 2580-2592.

Zhao, Y., Hull, A.K., Gupta, N.R., Goss, K.A., Alonso, J., Ecker, J.R., Normanly, J., Chory, J., and Celenza, J.L. (2002). Trp-dependent auxin biosynthesis in Arabidopsis: involvement of cytochrome P450s CYP79B2 and CYP79B3. Genes Development **16,** 3100-3112.

Zheng, Q., Ryvkin, P., Li, F., Dragomir, I., Valladares, O., Yang, J., Cao, K., Wang, L.S., and Gregory, B.D. (2010). Genome-wide double-stranded RNA sequencing reveals the functional significance of base-paired RNAs in Arabidopsis. PLoS Genetics **6,** e1001141.

Zhu, Y.Y., Nomura, T., Xu, Y.H., Zhang, Y.Y., Peng, Y., Mao, B.Z., Hanada, A., Zhou, H.C., Wang, R.X., Li, P.J., Zhu, X.D., Mander, L.N., Kamiya, Y., Yamaguchi, S., and He, Z.H. (2006). *ELONGATED UPPERMOST INTERNODE* encodes a cytochrome P450 monooxygenase that epoxidizes gibberellins in a novel deactivation reaction in rice. Plant Cell **18,** 442-456.



Samenvatting

Zaden vertegenwoordigen één van de belangrijkste evolutionaire innovaties. Ze verzorgen niet alleen reservevoedsel voor de zaailing, maar zorgen er ook voor dat een plant kan overleven in ongunstige omstandigheden. Onderdeel van de reserves die door de moederplant aangeleverd worden zijn "boodschapper" RNAs (mRNAs). Deze zijn essentieel voor de kieming van het zaad.

Tijdens de zaadontwikkeling en kieming gaat een zaad door verschillende fases van dehydratie en rehydratie: dehydratie (of uitdroging) tijdens de rijpingsfase en rehydratie wanneer het zaad water opneemt voor de kieming (imbibitie). Deze fasen beïnvloeden de vertaling van mRNA naar eiwitten (translatie). Dit maakt een zaad een interessant systeem om translatie in te onderzoeken. Kennis over translationele regulatie in de diverse zaadstadia maakt het mogelijk genexpressie te modificeren en daarmee zaadeigenschappen zoals kieming, kiemrust en bewaarbaarheid te verbeteren.

Normaal gesproken correleert de eiwitproductie met de transcriptie van mRNA. Dit kenmerkt ook de verschillende zaadontwikkelingsstadia (dit proefschrift), echter we zien ook specifieke stadia waarbij transcriptie en translatie ontkoppeld zijn (**Hoofdstuk 2).** In dit proefschrift heb ik ribosoomprofielen geanalyseerd, door ribosomen op grootte te scheiden middels een sucrose gradiënt. De aan de ribosomen gebonden mRNAs zijn vervolgens met microarrays geïdentificeerd (het translatoom). Op deze manier wordt er kennis verkregen over mRNAs die actief afgelezen en in eiwitten worden vertaald door de ribosomen. Deze analyse is voor verschillende stadia, van droog zaad tot kiemende zaailingen, uitgevoerd. Het blijkt dat er slechts in twee van de onderzochte stadia sprake is van translationele regulatie en dus een gebrek aan correlatie tussen de transcriptie en translatie. Dit betreft twee stadia tijdens de kieming, we hebben deze stadia de "Hydration Shift" en de "Germination Shift" genoemd, naar de fysiologische stadia waaraan ze gerelateerd zijn. De grote aantallen genen onder translationele regulatie in kiemende zaden staan in groot contrast met het totale gebrek hieraan tijdens de imbibitie van dormante zaden (**Hoofdstuk 4**) of de regulatie van kiemrust en bewaarbaarheid in zaden van de *ATP binding cassette family member i* (*abci)* mutant (**Hoofdstuk 5**). De analyses

beschreven in dit proefschrift laten duidelijk zien dat translationele regulatie zeer belangrijk is tijdens kieming (**Hoofdstuk 2**), maar dat kiemrust voornamelijk op transcriptioneel niveau geregeld wordt (**Hoofdstuk 4 en 5**).

Sequentiekarakteristieken zoals de lengte van het transcript, secundaire structuren en de aanwezigheid van specifieke motieven van de genen onder translationele regulatie, wijzen erop dat beide fases moleculair gezien op verschillende manieren gereguleerd worden. De genen die meer dan gemiddeld afgelezen worden uit de "Hydration Shift" worden gemarkeerd door een korte transcriptlengte en een 5'UTR met een laag GC%. Genen uit de "Germination Shift" worden minder dan gemiddeld vertaald en hebben over het algemeen een sterkere secundaire structuur (**Hoofdstuk 2**).

Ribosoomstudies richten zich in het algemeen meer op ribosomen in de polysoomvorm en de mRNAs gebonden aan deze polysomen. In droge zaden zien we echter voornamelijk monosomen. In **Hoofdstuk 3** heb ik in meer detail gekeken naar deze monosomen. Eén derde van de mRNAs geassocieerd met de monosomen in droog zaad bleek tijdens de kieming translationeel omhoog gereguleerd te zijn. mRNAs geassocieerd met polysomen zijn voornamelijk restanten van de zaadontwikkeling. Het gaat hierbij vooral om genen die geassocieerd zijn met "processing bodies". Monosomen zijn geassocieerd met verschillende typen mRNA bindende eiwitten, waaronder GLUTATHIONE S-TRANSFERASES die mogelijk betrokken zijn bij de bescherming van mRNAs tegen oxidatie tijdens de droge bewaring van zaden.

Zoals eerder genoemd wordt het behoud van kiemrust in geïmbibeerd zaad voornamelijk bepaald door transcriptie (**Hoofdstuk 4**). Deze conclusie is gebaseerd op de grote transcriptionele verschillen tussen dormant en nagerijpt (niet-dormant) geïmbibeerd zaad, het gebrek aan translationele controle en het feit dat de ribosomale profielen van dormante zaden nagebootst kunnen worden door transcriptie te remmen met cordycepine. De meerderheid van de genen die differentieel tot expressie komen wanneer dormante met niet dormante zaden vergeleken worden zijn genen die een rol spelen bij de tryptofaanafhankelijke auxine- en de indol glucosinolaat biosynthese.

In **hoofdstuk 5** van dit proefschrift heb ik het *ABCI20* gen bestudeerd. Dit gen

speelt een rol bij de regulatie van zowel kiemrust als bewaarbaarheid, zoals blijkt uit kiemexperimenten met de *abci* mutant. Deze mutant is zeer dormant maar slecht bewaarbaar, en is hiermee de eerste mutant waarbij kiemrust en bewaarbaarheid negatief correleren. De mutant is ook overgevoelig voor ABA. Voor de analyse van de *abci20* mutant hebben we dezelfde methodes gebruikt als in de eerdere hoofdstukken. Uit de ribosoomprofilering bleek dat droge zaden van de *abci20* mutant meer polysomen bevatten dan droog zaad van het wildtype. De polysoomcomplexen bleken verrijkt met eiwitten gerelateerd aan metabole processen, die normaal gedurende de late zaadrijping plaatsvinden, en proteasoomcomplexen. Dit suggereert dat het afrijpingsprogramma in de *abci20* mutant verstoord is. Ditzelfde is terug te zien in de transcripten die nog aanwezig zijn in rijp zaad van de *abci20* mutant. Normaal gesproken komen deze tijdens de late zaadontwikkeling tot expressie. Onder deze genen zijn er die gerelateerd zijn aan kiemrust, zoals genen uit de ABA- en ethyleensignalering en GA- metabolisme, maar ook genen die mogelijk een rol spelen bij bewaarbaarheid, met name oxidatie-reductie, stress response, autofagie en metabolisme gerelateerde genen.

Welbeschouwd staat het onderzoek naar translationele regulatie in zaden nog in de kinderschoenen; echter dit proefschrift verschaft nieuwe inzichten in regulatoire processen die mogelijk tot de verbetering van gewassen kunnen leiden.

Acknowledgments

Time flies. This feeling suddenly comes to my mind as time runs out for my PhD. Reflecting about the great personal and scientific life I have experienced in the past four years, I cannot express more gratitude to the people who kindly provided their company and support, either with a sparling smile during coffee breaks or with a brilliant scientific idea exchanged during work discussions. These are the positive emotions that continuously motivate my passion to complete this thesis.

I have been very lucky to get acquainted with two groups of colleagues and friends from Utrecht and Wageningen. All of them have contributed to the completion of my thesis.

It was a great pleasure to work with you, Leónie. Your determination, responsibility and your great passion for science were the key to help me finalizing my PhD. You introduced the term 'translatome' in the project proposal you wrote together with Johannes. You and Johannes bravely applied recently developed translatomic platform and integrated it into the seed research. This creative idea raised my curiosity and drove me to explore in-depth of the mRNA translational machinery occurring in seeds. As a supervisor, you have set an example for a high quality and responsible supervision. Besides your contribution to the development of my scientific performances, you were also present when I needed support in my personal life and for my future career. This way of supervision turned out to be effective to pull me out from the ambiguity about the way of life I would like to choose. I am truly willing to spend more time working with you although for now I have to move forward and try to pursue a further scientific career. However you will be the most important contact for the work already done and to be done in the future.

Johannes, although the geographic distance decided that I could not meet you as frequently as Leónie, without your contribution the project could not have progressed smoothly. You always provided the most enlightening ideas which sometimes were beyond my understanding, but later they turned out to be wise. This process stimulated my curiosity to explore the deeper insights of the unknown. During the past four years, Skype meetings became the official way to discuss scientific questions regarding the progress of the project and almost all the important decisions

have been made through this modern way of meetings. Of course, I would like to see you in person more often because of your friendly smile and excellent sense of humour, which were for me important elements for a smooth communication between student and supervisor and a successful project completion. I appreciate your early offer for a post-doc position in Sweden. This was especially important for me as it ensured the continuation of the scientific output from my PhD to be completed and in the meanwhile I could concentrate on writing my thesis during the final stage of my PhD. This leads to the accomplishment of my thesis writing in the most efficient way.

I would like to thank my promoters Prof. Sjef Smeekens and Prof. Harro Bouwmeester for spending time on proofreading and giving suggestions on my final thesis manuscript. Sjef, you are the essential player in the project initiation. And you have always been friendly and helpful for everything I needed. You kindly took care of most documentation issues that were necessary for the completion of my PhD. Harro, thank you for accepting me as a guest PhD in the Laboratory of Plant Physiology (PPH) at the Wageningen University and it is great pleasure to work here. Acknowledgement also goes to my external supervisor Prof. Maarten Koornneef. Our annual meetings were always short but efficient. I could always get important input from your experience and knowledge on the work I was conducting. Henk, you always gave me your friendliest hugs and smiles which I deeply appreciate. Wilco, your friendly presence makes me always feel like home. Sander, you are a great scientist with a wide knowledge and lots of new ideas which I always enjoyed listening. Thanks Dick for chairing Monday morning seminars every week and hope you will enjoy your retirement soon. Thanks Richard and Harm for your suggestions on my work and assistance with my data analysis. Bas, you were definitely the greatest desk mate ever, which guaranteed the problem solving in an efficient way thanks to your rich experience working in the lab. Leo and Juriaan, thanks for your helpful technical supports in the lab. I was really glad to work with you. Jolanda and Evelien, working with you has made my working time in Utrecht the most efficient. This was really essential for me to finish the project in a limited time due to the fact

that my daily work was based in Wageningen University.
I am also grateful for the tedious administration work that has been done by PPH secretaries Rina and Margaret as well as by the EPS staff members. Without your effort, the management of such international community would not be that successful. The Wageningen Seed Lab (WSL) is an essential component of the PPH department and represents the biggest seed lab in the world. I am very proud to be a family member. Working in the WSL was definitely the best experience in my scientific career. Everyone here works in a coordinative way since we were free to exchange ideas and seek for support when you get entangled in your work and people always spare no effort to help. One of the most important persons I would like to thank is Hanzi. We met during your visit in Israel and this was also part of the bridge which got me connected to the Netherlands. I am very glad that you found a new job in China and look forward to meet your baby soon. Also I hope that Junwei will also soon find a job which fits him well. In the Netherlands, I enjoyed every moment in such an international group of colleagues and friends. Life was never boring thanks to the PPH activity committee, the PhD trip organizers (Beatriz, Esmer, Maria Cecilia and Melissa) and supporters (Harro and Wilco), Spanish and Dutch lunch organizers (Maria Cecilia, Beatriz, Natalia, Gonda, Esmer). You have made my PhD life colourful and cheerful. I give my sincere thanks to my paranymphs Maria Cecilia and Elise for their company at the final and essential part of my PhD. Thank you Phuong for providing information regarding the defence arrangement at the Utrecht University. Thank you all my faithful sports and games competitors Giovanni, Imran, Qing Liu, Wei Song, Jun He, Ke Lin, Jiao Long, Tao Zhao, Li Xie, Wei Qin, Yang Yu, Xiao Lin, Mingtian.
I am glad I have met a group of old and new Chinese friends-Hanzi, Yanxia, Hui Li, Wei Song, Jun He, Qing Liu, Bo Wang, Yunmeng, Xi Cheng, Yuanyuan, Yuting, Jianhua, Yanli, Guiling, Xu Cheng, Tingting, Tao Zhao, Feng Zhu, Shuhang, Kaile, Zhen Wei, Xuan Xu, Chunxu, Wei Qin, Yang Yu, Huchen, Defeng, Fengjiao and many more. This short period spent together with you is of a great fortune in my life and wish you all the best wherever you go in the future.

Due to the special work arrangement, I also have a group of colleagues to whom I would like to express my gratitude. These persons are from the Molecular Plant Physiology (MPF) group at the Utrecht University. Magdalena, you are definitely an excellent researcher. Your perseverance to set up and optimize the ribosome fractionation system has undeniably helped me to achieve this project. I still remember the time I spent with you and Jeroen studying basics of translation research and improving our skills. Jeroen, I hope you will soon find the ideal job that you are looking for. Alessia let me thank you and your excellent student, Sjors, for assisting me with data analysis. Wouter, you are one of the few PhDs I could meet and talk frequently in Utrecht. Good luck with finalizing your PhD and your Japanese study. Monica and Jeba, good luck with finalizing your PhD.

I am the luckiest man because I have met my girlfriend Yinyin in the Netherlands. Your passion for combining nature and emotions in an artistic fashion is truly attractive. This is also an intersecting point where our interests meet. Your artistic vision of life in harmony with nature perfectly completes and even enriched my scientific vision and understanding of nature. Most importantly, your long-lasting company and support during my PhD were the most important part of my motivation to look towards to a future full of hopes and I will cherish these memories. I believe the efforts we have made for this relationship are truly precious and I am looking forward to the fruits of it in a near future.

I would like to give my special thanks to my parents. Your emotional support full of love and care during my time abroad is always encouraging. Although I choose an exploratory way of life rather than staying with you, in my hometown, in my country, the Chinese dream that fallen leaves will eventually return to the roots (落叶归根) is always buried in my heart wherever I go in the world. I send you best wishes for your well-being and health. I appreciate all the cares from the relatives of my family given to my parents during my absence from home.

(请让我对在背后默默支持我国外求学的父母表示由衷地感谢。你们无声的支持满载着感情与关爱，这是对我在求学征途中最大的鼓励。虽然我告别父母，背井离乡，追寻了一条科学与探索之路，但是落叶归根的游子之心无论

走到哪里始终都深藏我心。我深深的祝福您们，祝您幸福安康。同时感谢亲人们在我游走他乡之时给予我的父母无微不至的的关照)

Last but not least, my particular appreciation goes to my thesis defence opponents Prof. Julia Bailey-Serres, Prof. Corné Pieterse, Prof. Berend Snel, Dr. Hiro Nonogaki and Dr. Loïc Rajjou. Your presence and witness of my defence will make my PhD complete.

Curriculum Vitae

Curriculum Vitae

Bing Bai was born on 22 of January 1986 in Zhengzhou, Henan, China. In 2005, he started his Bachelor degree in the Department of Life Sciences at the Henan Agricultural University in Zhengzhou, China. He did his Bachelor thesis under the supervisor of Prof. Wei Wang on the molecular basis of seed vigor and longevity at the Laboratory of Plant Proteomics. After obtaining a bachelor degree in June 2009, he continued his Master study in the Albert Katz International School for Desert Studies, Ben-Gurion University, Israel. There, he worked under the supervision of Dr. Aaron Fait on stress-related plant metabolomics. After two years of study, in January 2012, he obtained his MSc degree with the thesis titled 'Metabolic Characterization of Plant Dehydration Tolerance in Seeds and Seedlings of Model plant *Arabidopsis thanliana* and Desert Plant *Schismus arabicus*'. This thesis yielded three publications. From 2012, he started his PhD at the Department of Molecular Plant Physiology of Utrecht University, the Netherlands. In the meanwhile, he worked as a guest PhD at the Laboratory of Plant Physiology of Wageningen University, the Netherlands, under the supervision of Dr. Leónie Bentsink and Dr. Johannes Hanson. The aim of the work was to investigate the role of seed stored mRNAs and translation in the control of seed dormancy. On 30 of March 2016, Bing defended his PhD thesis, presented in this book. After accomplish his PhD, Bing will continue developing his scientific career as a post-doc in Umeå Plant Science Centre (UPSC), Department of Plant Physiology at Umeå University, Sweden.

Publications:

Bing Bai, Alessia Peviani, Sjors van der Horst, Magdalena Gamm, Berend Snel, Johannes Hanson, Leónie Bentsink. Extensive translational regulation during seed germination revealed by translational Profiling. (Submitted)

Bing Bai, Johannes Hanson, Leónie Bentsink. Combined transcriptome and translatome analyses reveal a role for transcriptional inhibition of the tryptophan dependent auxin biosynthesis pathway in the control of seed dormancy. (Submitted)

Bing Bai, Sjors van der Horst, Jan H.G. Cordewener, Antoine H.P. America, Leónie Bentsink, Johannes Hanson. Fate specific mRNA storage in seeds. (In preparation)

Bing Bai, Jan H.G. Cordewener, Antoine H.P. America, Johannes Hanson, Leónie Bentsink. Role for *ABCI20* in regulating seed maturation. (In preparation)

Bing Bai, David Toubiana, Tanya Gendler, Asfaw Degu, Yitzchak Gutterman and Aaron Fait. Metabolic patterns associated with the seasonal rhythm of seed survival after dehydration in germinated seeds of *Schismus arabicus*. BMC Plant Biology. 2015. 15(37)

Airong Liu, Shuangchen Chen, Mengmeng Wang, Dilin Liu, Rui Chang, Zhonghong Wang, Xiaomin Lin, **Bing Bai**, Golam Jalal Ahammed. Arbuscular mycorrhizal fungus alleviates chilling stress by boosting redox poise and antioxidant potential of tomato seedlings. Journal of Plant Growth Regulation. 2015. DOI 10.1007/s00344-015-9511-z

Cristina Martínez-Andúja, Wioletta E. Pluskota, George W. Bassel, Masashi Asahina, Piotr Pupel, Theresa Nguyen, Noriko, Takeda-Kamiya, David Toubiana, **Bing Bai**, Ryszard Gorecki, Aaron Fait, Shinjiro Yamaguchi, Yuji Kamiya, Hiroyuki Nonogaki. Regulation mechanisms of endosperm cap-specific gene expression in

tomato seeds during germination. Plant Journal. 2012. 71: 575-586

Bing Bai, Noga Sikron, Tanya Gendler, Yanna Kazachkova, Simon Barak, Gideon Grafi, Inna Khozin, Aaron Fait. Ecotypic variability of the metabolic rearrangement in response to diurnal hydration-dehydration cycles in Arabidopsis seeds. Plant and cell physiology. 2011. 53(1): 38-52

Jihong Wang, Shaoning Chen, **Bing Bai**, Jihua Tang, Wei Wang. Effects of artificial ageing on proteome of maize seeds. Journal of Henan Agricultural University. 2009, 43 (3): 232-235

Education Statement of the Graduate School
Experimental Plant Sciences

Issued to: **Bing Bai**
Date: **March 30, 2016**
Group: **Molecular Plant Physiology**
University: **Utrecht University**

1) Start-up phase		date
►	**First presentation of your project**	
	The role of seed stored mRNAs and translation in the control of seed dormancy	Mar 09, 2012
►	**Writing or rewriting a project proposal**	
►	**Writing a review or book chapter**	
►	**MSc courses**	
►	**Laboratory use of isotopes**	
	Subtotal Start-up Phase	*1.5 credits**

2) Scientific Exposure		date
►	**EPS PhD student days**	
	EPS PhD student days, University of Amsterdam	Nov 30, 2012
	EPS PhD student days, Leiden University	Nov 29, 2013
	EPS PhD student days, Soest (NL)	Jan 29-30, 2015
►	**EPS theme symposia**	
	EPS theme 3 'Metabolism and Adaptation', University of Amsterdam	Mar 22, 2013
	EPS theme 3 'Metabolism and Adaptation', Wageningen University	Mar 11, 2014
	EPS theme 3 'Metabolism and Adaptation', Utrecht University	Feb 10, 2015
►	**NWO Lunteren days and other National Platforms**	
	NWO - ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 02-03, 2012
	NWO - ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 22-23, 2013
	NWO - ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 14-15, 2014
	NWO - ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 13-14, 2015
►	**Seminars (series), workshops and symposia**	
	Invited seminar: 'Tackling natural variance in seed metabolism integrating metabolite profiles via network analysis'	Dec 4, 2012
	Invited seminar: 'ABA signaling networks in Arabidopsis'	Nov 14, 2012
	Invited seminar: 'Use of resurrection plants as models to understand how plants tolerate extreme water loss: A systems biology approach with applications for making drought tolerant crops'	Jun 26, 2012
	Invited seminar: 'Parenting in plants: maternal control of seed dormancy'	Jun 12, 2012
	Invited seminar: 'Using the Nicotiana-TMV system to study resistance gene evolution and plant genome stability'	Sep 11, 2013
	Symposium 'Omics Advances for Academia and Industry - Towards True Molecular Plant Breeding'	Dec 11, 2014
	Invited seminar: 'The evolutionary significance of gene and genome duplications'	Feb 03, 2015
	Invited seminar: 'Hormone-related Functions of LEC2 in Somatic Embryogenesis Induction in Arabidopsis'	Apr 16, 2015
	Invited seminar: 'Regulation of root morphogenesis in tomato species in the face of a changing environment'	Sep 09, 2015
	Invited seminar: 'And yet they oscillate: functional analysis of circadian long non-coding RNAs'	Nov 16, 2015
	Invited seminar: 'The response to cold stress in rice: signaling, transcriptional and metabolic regulation'	Dec 01, 2015
	Invited seminar: 'How Jasmonates provide the key to harness plant chemistry'	Dec 08, 2015
►	**Seminar plus**	
►	**International symposia and congresses**	
	International Conference on Arabidopsis Research ICAR	Jul 03-07, 2012
	EMBO workshop China	Nov 17-22, 2013
	ISSS workshop	Jul 09-12 2013
	2nd Dutch Seed Symposium	Oct 01, 2013
	ISSS conference China	Sep 15-19, 2014
	ICAR conference Paris	Jul 05-09, 2015
	PSRG conference Paris	Jul 10-11, 2015
	4th Dutch Seed Symposium	Oct 06, 2015
►	**Presentations**	
	Plant Physiology Seminar (Talk)	Dec 17, 2012
	NWO - ALW meeting 'Experimental Plant Sciences', Lunteren (Poster)	Mar 22, 2013
	2nd Dutch Seed Symposium (Talk)	Oct 01, 2013
	NWO - ALW meeting 'Experimental Plant Sciences', Lunteren (Poster)	Apr 14-15, 2014
	Plant Physiology Seminar (Talk)	Sep 09, 13
	ISSS conference in China (Talk)	Sep 15-19, 2014
	NWO - ALW meeting 'Experimental Plant Sciences', Lunteren (Talk)	Apr 13-14, 2015
	ETH Zurich (Talk)	Apr 29, 2015
	Plant Research International - PRI (Talk)	Jun, 2015
	ICAR conference Paris (Poster)	Jul 05-09, 2015
	PSRG conference Paris (Talk)	Jul 10-11, 2015
	Plant Physiology Seminar (Talk)	Nov 09, 2015
►	**IAB interview**	
►	**Excursions**	
	PhD trip to Cologne - Tubingen (Germany) and Zurich (Schwitzerland)	April 22-May 02, 2015
	Subtotal Scientific Exposure	*27.5 credits**

3) In-Depth Studies		date
►	**EPS courses or other PhD courses**	
	System biology: -omics data analysis	Dec 10-14, 2012
	Utrecht Summer School	Aug 26-28, 2013
	Microscopy and Spectroscopy in Food and Plant Sciences	May 06-09,2014
	Seed Summer School in Paris	Jun 27-Jul 03, 2015
►	**Journal club**	
	Member of a literature discussion group	2012-2016
►	**Individual research training**	
	Subtotal In-Depth Studies	*8.4 credits**

4) Personal development		date
►	**Skill training courses**	
	Stress Identification and Management	Sep 26, 2013
	Reviewing a scientific paper	Mar 05, 2015
	Grant Proposal Writing	Apr 07-Jun 09, 2015
	Essentials of Scientific Writing & Presenting	Sep 11-23, 2015
►	**Organisation of PhD students day, course or conference**	
	PhD trip for the Lab of Plant Physiology	April 22-May 02, 2015
►	**Membership of Board, Committee or PhD council**	
	Subtotal Personal Development	*5.1 credits**

TOTAL NUMBER OF CREDIT POINTS*	**42.5**

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

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