

The mRNA-binding proteome of a critical phase transition during Arabidopsis seed germination

Nikita Sajeev, Anirban Baral, Antoine H.P. America, Leo A.J. Willems, Rémy Merret, Leónie Bentsink

▶ **To cite this version:**

Nikita Sajeev, Anirban Baral, Antoine H.P. America, Leo A.J. Willems, Rémy Merret, et al.. The mRNA-binding proteome of a critical phase transition during Arabidopsis seed germination. New Phytologist, In press, 10.1111/nph.17800 . hal-03377537

HAL Id: hal-03377537

https://univ-perp.hal.science/hal-03377537

Submitted on 14 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

New Phytologist

The mRNA-binding proteome of a critical phase transition during Arabidopsis seed germination

Journal:	*New Phytologist*
Manuscript ID	NPH-MS-2021-37074.R1
Manuscript Type:	MS - Regular Manuscript
Date Submitted by the Author:	24-Sep-2021
Complete List of Authors:	Sajeev, Nikita; Wageningen Universiteit en Research, Plant Physiology Baral, Anirban; Wageningen Universiteit en Research, Plant Physiology America, Antoine H.P. ; Wageningen Universiteit en Research, BU Bioscience Willems, Leo; Wageningen Universiteit en Research, Plant Physiology Merret, Rémy; Universite de Perpignan Via Domitia, LGDP UMR5096 Bentsink, Leónie; Wageningen Universiteit en Research, Plant Physiology
Key Words:	mRNA, RNA binding proteins, Seeds, Translation, Germination

SCHOLARONE™
Manuscripts

The mRNA-binding proteome of a critical phase transition during Arabidopsis seed germination

Nikita Sajeev[1], Anirban Baral[1], Antoine H.P. America[2], Leo A.J. Willems[1], Rémy Merret[3] and Leónie Bentsink[1]*

[1]Wageningen Seed Science Centre, Laboratory of Physiology Wageningen University, 6708PB Wageningen, The Netherlands

[2] BU Bioscience, Wageningen Plant Research, 6700 AP Wageningen, The Netherlands

[3] Laboratoire Génome et Développement des Plantes, CNRS-LGDP UMR 5096, 66860 Perpignan, France

*Author for correspondence:

Leónie Bentsink

Email: leonie.bentsink@wur.nl

Total word count (excluding summary, references, acknowledgements and legends):	6407	Acknowledgements	59
Summary:	189	No. of figures	5
Introduction:	732	No. of tables	0
Materials and Methods:	2449	No. of Supporting Information files:	2
Results and discussion	3186		

Peer Review

Abstract

- *Arabidopsis thaliana* seed germination is marked by extensive translational control at two critical phase transitions. The first transition refers to the start of hydration, the hydration translational shift. The second shift, the germination translational shift (GTS) is the phase between testa rupture and radicle protrusion at which the seed makes the all or nothing decision to germinate.
- The mechanism behind the translational regulation at these phase transitions is unknown. RNA binding proteins are versatile players in the post-transcriptional control of mRNAs and as such candidates for regulating translation during seed germination.
- Here, we report the mRNA binding protein repertoire of seeds during the GTS. Thirty seed specific RBPs and 22 dynamic RBPs were identified during the GTS, like the putative RBP Vacuolar ATPase subunit A and RBP HSP101. Several stress granule markers were identified in this study, which suggests that seeds are prepared to quickly adapt the translation of specific mRNAs in response to changes in environmental conditions during the GTS.
- Taken together this study provides a detailed insight into the world of RNA binding proteins during seed germination and their possible regulatory role during this developmentally regulated process.

Keywords

Germination, mRNA, RNA binding proteins, seeds, translation.

For Peer Review

Introduction

Seed germination is a complex process in which the seeds need to undergo developmental transitions to successfully establish themselves as a plant. The majority of our understanding on how plant development is regulated has been a product of studying gene expression with the main focus on transcription and DNA binding partners. However, recent studies have highlighted that translational regulation plays an important role in regulating plant development (Sorenson & Bailey-Serres, 2014; Merchante *et al.*, 2017; Sablok *et al.*, 2017; Cho *et al.*, 2018; Cho *et al.*, 2019; Jang *et al.*, 2019). The complete switch-off state of translation between seed maturation and seed germination makes seeds a unique system to study developmentally regulated translation (Sajeev *et al.*, 2019). Previously, it has been shown that there is extensive translational control at two temporal shifts during seed germination. These shifts were defined as the hydration translation shift (HTS) and germination translational shift (GTS) (Bai *et al.*, 2017). Interestingly these shifts coincide with important developmental phase transitions during seed germination. The HTS spans the first six hours after imbibition (HAI); the phase at which seeds take up water. Upon imbibition, the dry seed undergoes a drastic transition from a metabolically inactive to a highly active state. The GTS is the developmental phase between seed testa rupture (TR) and radicle protrusion (RP). These phases mark critical physiological stages of seed germination. Upon TR, seeds can still be dried back without hampering its viability which becomes more difficult as germination progresses. This is because desiccation tolerance can be re-introduced into seeds only within a limited time frame which is usually lost once upon RP (Maia *et al.*, 2011). This developmental transition can be viewed as a point of no return, also known as germination *sensu strictu* (Perino & Côme, 1991). The decision to germinate is based on a complex web of environmental and developmental signals to ensure seedling survival. At the GTS, distinct subsets of mRNAs show differential translation which suggests dynamic regulation of germination (Bai *et al.*, 2017). The mechanism behind this selection is yet to be understood. In recent years, several studies have implicated RNA binding proteins (RBPs) can regulate their target mRNAs co- and post-transcriptionally, thereby altering its translation efficiency in plants (Köster *et al.*, 2017; Lou *et al.*, 2020). Furthermore, a recent study reported that certain stored mRNAs in the dry seed are associated with single ribosomes and RBPs which are later

For Peer Review

translationally upregulated during germination (Bai *et al.*, 2020). This led to the hypothesis that certain RBPs could play a role in determining the fate of the regulated mRNAs during seed germination.

Defining features of RBPs are their putative RNA binding domains like the Pumilio (PUM) domain, Zinc-finger domains, K homology (KH) domain or the RNA recognition motif (RRM) (Lorković, 2009). Several studies have demonstrated the role of RBPs in plant development. Some examples include the RBP JULGI that regulates phloem differentiation by translational control of *SUPPRESSOR OF MAX2-LIKE1-4/5 (SMXL4/5*) and *ETHYLENE INSENSITIVE2 (EIN2)* a non-canonical RBP that can regulate hypocotyl elongation by repressing the translation of ethylene responsive mRNAs ((Merchante *et al.*, 2015; Cho *et al.*, 2018; Cho *et al.*, 2019). In seeds, through a transcriptomics study, an RBP belonging to the PUM family, *ARABIDOPSIS PUMILIO (APUM) 9* was shown to play a role in delaying seed germination (dormancy) (Xiang *et al.*, 2014). Although, recent advancements in RNA-protein interactome capture techniques have allowed the identification of classical and novel RNA binding proteins (RBPs) in different plant tissues, their identity and role in seeds has not yet been explored (Marondedze *et al.*, 2016; Reichel *et al.*, 2016; Zhang *et al.*, 2016; Köster *et al.*, 2017; Cho *et al.*, 2019; Bach-Pages *et al.*, 2020).

In the present study, mRNA interactome capture was performed in Arabidopsis embryos at TR and RP, the physiological stages that mark the GTS. Hundreds of high confidence RBPs were identified. Additionally, dynamic RBPs were identified in this study like the putative RBP Vacuolar H+-ATPase subunit A and known RBP HEAT SHOCK PROTEIN 101 (HSP101). These RBPs were also exclusively identified in the seed mRNA interactome capture and not in leaves, protoplasts or etiolated seedlings (Marondedze *et al.*, 2016; Reichel *et al.*, 2016; Zhang *et al.*, 2016; Bach-Pages *et al.*, 2020). Overall, this study provides a valuable resource for future RBP research in seeds and will be the starting point of identifying their possible regulatory role in translation during seed germination.

For Peer Review

Results and discussion

Identification of the mRNA binding proteome at the germination translational shift

The GTS defines the period of translational regulation between TR and RP. The exact moment of RP is genotype and environment dependent, which implies that this has to be determined for every new experiment. In this experiment, TR and RP occurred at 26 and 42 hours after imbibition (HAI) for Arabidopsis ecotype Col-0 seeds (Fig. S1). To unravel the mRNA binding proteome during the GTS, the existing mRNA binding interactome protocol had to be extensively adapted for Arabidopsis embryos (Castello *et al.*, 2013) (Fig. 1a). The mRNA interactome capture was performed on the embryos of three independent biological replicates at TR and RP. To summarize, UV radiation was used to crosslink (CL) the mRNA-RBP complexes while processing the non-crosslinked controls (NCL) in parallel. The embryos were lysed in a denaturing buffer and poly-A mRNA was pulled down using oligo- dT magnetic beads (Fig. 1a). Poly-A mRNA enrichment was seen in the eluates after poly-A pulldown compared to the total input RNA before pulldown using qPCR (Fig. S2). Next, the enrichment of proteins in the CL samples over the NCL was confirmed using silver stained SDS-PAGE gels (Fig. 1b). The samples were then analyzed using label free nano LC MS/MS analysis. Scatter plots of the LFQ (Relative Label-free quantitation) intensities between the replicates showed good reproducibility at both time-points (Fig. S3).

Over 1300 proteins were identified across all samples. However, only proteins for which two or more unique peptides were detected in at least two biological replicates of the CL samples were taken for further analysis. This resulted in more than 600 proteins that were enriched in the CL samples in both stages. One hundred and six and 112 proteins were identified as high confidence RBPs (FDR<5%) at TR and RP respectively (GTS-RBPs), with an overlap of 54 proteins that were present at both time-points (Supporting Information Table S1a,b). Although, several proteins did not pass these stringent parameters of selection, many proteins were highly enriched in the crosslinked samples over the controls and therefore could be important RBPs that play a role in the GTS. Hence, 228 proteins at TR and 244 proteins at RP with a log2 fold (CL/NCL) enrichment >1 were

For Peer Review

classified into a second set called the candidate RBPs for each time-point in our dataset (Fig. 1c and Table S1a,b).

Next, the GTS-RBPs and candidate RBPs were annotated based on their molecular function. This revealed that approximately 80% of the GTS-RBPs had been previously annotated with known or predicted RNA binding activity, while 47 GTS-RBPs were not and could be putative RBPs (Fig 1d). The candidate RBP set showed a large proportion of RBPs not annotated as mRNA binding and therefore provide a repertoire of putative RBPs in seeds (Fig. 1d, S1a,b). A gene ontology (GO) enrichment analysis for all GTS and candidate RBPs over the two time-points showed common enrichment for GO terms like binding, mRNA binding, heterocyclic compound binding and organic cyclic compound binding (Table S1c). Overall, the GO analysis, revealed that the interactome capture strongly enriched for proteins related to RNA biology.

Protein domain analysis reveals stage specific protein families during the GTS

Both the GTS-RBPs and candidate RBPs at TR and RP were grouped by their protein domain annotations (PFAM or Interpro annotations) (Fig. 2a, Table S1d). At both stages, diverse classical and non-classical RNA binding domains (RBDs) were captured (Fig. 2a). Examples of classical domains include RRM, KH domain, Zinc finger (zf)-CCCH, DEAD box Helicases and PUM. The vast majority of the RBPs identified contained the RRM domain (Fig. 2a). The Arabidopsis proteome consists of 253 proteins containing an RRM domain (Lorković & Barta, 2002). The RRM family is highly diverse in plants and in this study 66 GTS-RBPs and 47 candidate RBPs containing an RRM domain were identified in seeds. Majority of the RRMs have not been investigated for their roles in germination and could be important regulators of germination. An example of such a regulator is an RRM containing glycine rich protein, atRZ-1a, which was identified as a candidate RBP at RP. This RBP has been reported to negatively impact germination under salt and osmotic stress (Kim *et al.*, 2007). The Arabidopsis PUM family contains 25 proteins that are phylogenetically classified into four groups. Interestingly only group 1 APUM RBPs (APUM1,3,5 and 6) were identified as GTS-RBPs at both stages indicating that group 1 APUMs are especially abundant during seed germination.

For Peer Review

Non-classical RBDs like Ribosomal, La and GTP-EFTU were also well represented at both TR and RP (Fig. 1a). The non-classical RBD, HABP4_PAI-RBP1 family was only identified in the candidate RBP set at TR (Fig. 2a). Three Hyaluronan/mRNA binding proteins contained this RBD namely, AtRGGA, AT5G47210 and AT4G17520. AtRGGA has been reported to play a role in abscisic acid (ABA) signalling during stress response in seedlings. Mutants of this RBP are highly susceptible to salt and osmotic stress (Ambrosone *et al.*, 2015). *AT5G47210* was revealed to be highly expressed one day after seed imbibition followed by a reduction at later time-points (Narsai *et al.*, 2011). These time-points closely coincide with the stages of TR and RP and could explain why this protein is no longer identified at RP point. In the present study, one knockout mutant (Fig. S4) and 2 complementation lines of *AT5G47210* have been investigated for seed germination phenotypes. This revealed a dormancy phenotype, measured as DSDS50 (days of seed dry storage required for 50% germination (Alonso-Blanco *et al.*, 2003; Soppe & Bentsink, 2020). The knockout mutant *at5g47210* had a DSDS50 of only 8.5 days in comparison to its wild type Col-0 which required 20 days (Fig. 2c). The complementation lines COMP1 and COMP2, complemented this mutant phenotype (Fig. 2c). Therefore, AT5G47210 could play a role in inhibiting germination. The mechanism by which this RBP regulates germination needs to be further explored.

The domain analysis also revealed many putative RBDs many of which belonged to the elongation Initiation factor 3 (EIF3) family (Fig. 2b). Other protein families such as HSP70, AAA and DUF1264 have also been identified as putative RBDs in previous studies (Reichel *et al.*, 2016; Bach-Pages *et al.*, 2020). Interestingly, many enzyme families like Phosphoglycerate kinase (PGK), thioredoxins, Glutathione-S-transferase (GST), and NAD(P) binding domain (NAD(P)-bd_dom_sf) proteins were pulled down in this study (Fig 2b, Table S1d). There have been more reports on metabolic enzymes with RNA binding functions in eukaryotes (Castello *et al.*, 2015; Marondedze *et al.*, 2016; Reichel *et al.*, 2016; Bach-Pages *et al.*, 2020). PGKs and Thioredoxins have been validated as RBPs in humans and yeast cells (Beckmann *et al.*, 2015). In plants, it has been shown that GSTs are modulated by atRZ-1a, an RRM and Zinc finger domain containing protein also identified as a GTS-RBP in this study. This report concluded that this enzyme among others

play a role in ROS homeostasis during germination (Kim *et al.*, 2007). In another study, some NAD(P) binding domain proteins were identified as RBPs that respond to osmotic stress (Marondedze *et al.*, 2019). Most enzyme families identified in this interactome capture have been known to play a role in ROS homeostasis. However, their discovery as a putative RBDs in this study, suggests novel roles for these metabolic enzymes as RBPs in the translational regulation of seed germination.

Dynamic RBPs identified during the GTS

An in-depth analysis into the non-overlapping GTS- RBPs (106 at TR and 112 at RP) showed that many RBPs were identified as a GTS-RBP at one time-point and as a candidate RBP in the other. However, only 22 RBPs of these GTS- RBPs were exclusively identified in one time-point alone and therefore classified as dynamic GTS-RBPs (Table S2). The dynamic GTS-RBPs included known RBPs like ETHYLENE INSENSITIVE 2 (EIN2) and HSP101 (Merchante *et al.*, 2015; Merret *et al.*, 2017). EIN2 mutants have been shown to have a very strong dormancy phenotype due to high ABA levels in the dry seed (Koornneef *et al.*, 2002). This study demonstrates that EIN2 can also function as an RBP during seed germination. HSP101 was reported to bind and regulate the translation of the internal light-regulatory element (iLRE) of ferredoxin (Fed-1) mRNA in carrot protoplasts (Ling *et al.*, 2000). A recent study further showed that HSP101 is required for the efficient release of ribosomal protein mRNAs from stress granules for the rapid recovery of the translational machinery from heat stress (Merret *et al.*, 2017). Traditionally, HSP proteins are regarded as conserved molecular chaperones involved in protein folding stability and activation. However, several other HSPs such as HSP81.2, HSP70 and HSP70b were identified as part of the candidate RBP set at TR while HSP60, HSP91 chloroplast and mitochondria HSP70.1 were identified in the candidate RBP dataset at the RP stage. HSP101 was the only GTS RBP identified exclusively at the TR point and could function as an RBP involved in the phase transition from TR to RP, however the *hsp101* mutant did not show a germination or dormancy phenotype compared to wild-type (Fig. S5).

We also identified many dynamic putative GTS-RBPs. An example of a dynamic GTS-RBP with no links to RNA biology is the VACUOLAR H+-ATPase SUBUNIT A (V-ATPase SUBUNIT A) identified at the RP stage. V-ATPases are versatile multi-subunit

Manuscript submitted to New Phytologist for review

For Peer Review

proton pumps that control the pH of many intracellular compartments in all eukaryotic cells. In Arabidopsis, V-ATPases play a role in plant defenses against environmental stresses like salt stress. The subunit A gene detected in Arabidopsis can produce at least four different transcripts by using different polyadenylation sites. These transcripts differ only in their 3' untranslated region and produce identical proteins (Magnotta & Gogarten, 2002).

The dynamic nature and the RBP identity for HSP101 and V-ATPase SUBUNIT A was validated using western blotting (Fig. 3a). ARGONAUTE 1 (AGO1) being a well-established RBP also identified in this study was used as a positive control and ACTIN 7 was used as a negative control (Fig. 3a). The results confirmed the dynamic nature of HSP101 and V-ATPase SUBUNIT A which were highly abundant in the CL samples at TR and RP respectively (Fig. 3a, Fig. S6). Although AGO1 showed similar LFQ intensities at TR and RP in this study, the western blot showed some dynamics for this protein indicating the qualitative rather than quantitative nature of label free proteomics. The negative control ACTIN 7 was only present in the total protein of TR and RP and not after the poly-A pulldown, demonstrating the stringency of the mRNA interactome procedure. To confirm that the changes observed after the interactome capture were not due to differences in total protein abundance, an additional proteomics analysis on the total input protein fractions was performed at both stages. As highly abundant proteins can limit the identification of less abundant proteins, we were able to identify only 11 out of the 22 dynamic GTS-RBPs in the total input protein samples (Table S1f, Table S2). The data confirmed that there were no significant differences in protein abundance for HSP101 and V-ATPase SUBUNIT A at TR and RP before the interactome capture. This further supports our hypothesis that HSP101 and V-ATPase SUBUNIT A are dynamic RBPs at TR and RP respectively.

HSP101 plays a role in releasing ribosomal RNAs from stress granules for heat stress recovery (Maia *et al.*, 2011; Merret *et al.*, 2017). In the case of V-ATPase SUBUNIT A, its function as an RBP is unclear. It has been previously reported that vacuoles from tomato protoplasts can contain RNA oligonucleotides (Abel *et al.*, 1990) and a recent study demonstrated that, RNAse T2 ribonucleases are targeted to vacuoles for rRNA degradation

For Peer Review

and maintenance of cellular homeostasis in Arabidopsis (Floyd *et al.*, 2017). Both these studies show that RNAs can be targeted to vacuoles. A plausible hypothesis could be that putative GTS-RBP V-ATPase SUBUNIT A is involved in the sequestration of RNA to the expanding vacuoles at RP to maintain cellular RNA homeostasis. However, further research is required to establish the RBP identity and roles of both these RBPs during the GTS.

Comparison with other plant interactome captures reveal seed specific RBPs

Due to technical advancements in the recent years, mRNA interactome capture has gained a momentum in plant research. In the last five years, four different studies have published the mRNA interactome of Arabidopsis seedlings (300 RBPs), leaves (717 and 230 RBPs) and protoplasts (325 RBPs) (Marondedze *et al.*, 2016; Reichel *et al.*, 2016; Zhang *et al.*, 2016; Bach-Pages *et al.*, 2020). Although the previous studies identified much larger sets of statistically enriched RBPs, a comparative analysis of all GTS-RBPs identified in this study with previously performed interactome captures revealed 30 GTS-RBPs that were only identified in germinating seeds and 5 RBPs that were common to all datasets (Fig. 3b, Table S1e and S3). This shows that Arabidopsis RBPs are highly versatile, tissue and developmental stage specific. Eleven out of these 30 seed specific GTS RBPs had been previously annotated with an mRNA binding function and contained classical RNA binding domains. Many previously unknown RBPs in this set were enzymes like H (+)-ATPase 1, pyruvate orthophosphate di-kinase and hydroxysteroid dehydrogenase 5 (Fig. 3b, Table S3). Interestingly, nine out of the 30 seed specific RBPs are also part of the dynamic GTS-RBPs set identified in this study.

Stress granule markers enable quick responses to the environment

During the GTS the seed makes an all or nothing decision to germinate or not. In a biological context, germination must only proceed when the environmental conditions allow the successful establishment of the seedling. Several proteins that have been previously described to be part of cytoplasmic stress granules were identified at both stages with similar LFQ intensities like, RNA BINDING PROTEIN 47 A (RBP47), RBP47B, OLIGOURIDYLATE-BINDING PROTEIN 1C (AtUBP1c) and POLY-A BINDING PROTEIN 2 (PABP2). Stress granules are cytoplasmic foci which are formed in response

For Peer Review

to various environmental stresses like salt stress, hypoxia and heat stress (Chantarachot & Bailey-Serres, 2018). Stress granules can transiently store mRNAs until the stress resolves, allowing cells to quickly repress the translation of specific mRNAs in a stressful situation. To show that stress granule markers quickly respond to stressful conditions, a reporter line of stress granule marker RFP-PABP2 was imaged at TR and RP in response to heat stress (Fig. 4b). At control conditions PABP2 was expressed throughout the cytoplasm in the radicle tip of embryos and did not show any clear foci formation. Interestingly, after a short heat stress, PABP2 is clearly localized into cytoplasmic stress granules at both TR and RP (Fig. 4a). Further, to explore whether dynamic GTS-RBP HSP101 could regulate translation at TR, PABP2 was imaged in *hsp101* background after a short heat stress. (Maia et al., 2011; Merret et al., 2017). As expected, the number of stress granules in the *hsp101* seeds were significantly higher than wild-type at the TR stage (Fig. S7) suggesting that HSP101 is a GTS-RBP that could play a role in the translational control of germination via stress granules. This suggests that seeds during GTS possibly express certain stress granule markers in preparation for a quick adaptation of translation in response to changed environmental conditions.

P-bodies are also cytoplasmic granules in which translationally repressed mRNAs can be decayed or stored for development or stress responses (Narsai *et al.*, 2011; Hubstenberger *et al.*, 2017). P-bodies can contain several RBPs, 5′ to 3′ exoribonucleases, de-adenylation factors and factors involved in nonsense-mediated mRNA decay (Maldonado Bonilla, 2014). Some examples of P-body components are DECAPPING PROTEIN 1 (DCP1), DCP2, DCP5 and EXORIBONUCLEASE4 (XRN4) (Xu & Chua, 2009). Although, many of these P-body makers are known to be expressed in seeds, we only identified DCP5 as a dynamic GTS-RBP at the TR point. Previously, DCP5, has been shown to play a role in the translational repression of mRNAs via P-bodies in seedlings and in dark/light phase translation (Xu & Chua, 2009). To explore the localization of DCP5 during the GTS, a DCP5-GFP reporter line was imaged at TR and RP (Fig 4a). At TR, DCP5 forms more cytoplasmic granules than at the RP stage in the epidermal cells of the radicle tip (Fig. 4b). This differential localization could explain why DCP5 was identified as a dynamic GTS-RBP in the present study. Interestingly, DCP5 was the only well-established P-body marker

For Peer Review

identified in the interactome capture of leaves and in seedlings (Reichel *et al.*, 2016; Bach-Pages *et al.*, 2020) while for example, DCP1 was not. This indicates that the mRNA interactome capture method may not be an ideal to pull down all types of cytoplasmic granules. This could be explained by the fact that P-bodies contain deadenylation factors that degrade the poly-A tails of the mRNAs and in the present interactome capture, only poly-A mRNAs were pulled down (Maldonado Bonilla, 2014) or it could be that the mRNAs present in these bodies are not easily accessible to the oligo-dT beads used in this study.

In summary, the GTS spans a critical phase during germination at which extensive translational regulation takes place in which 195 and 717 mRNAs are translationally up and down regulated respectively (Bai *et al.*, 2017) (Fig. 5). The mechanism behind this selection is yet to be elucidated. The fate of the regulated mRNAs could be controlled by RBPs present during this shift. Over 600 GTS and candidate RBPs were identified. Among these, 228 and 244 GTS-RBPs were identified with high confidence at TR and RP respectively, 22 revealed to be dynamic GTS-RBPs and 30 were seed specific RBPs. Several GTS-RBPs have been previously reported to play a role in Arabidopsis seed germination. GTS-RBP EIN2 plays a role in reducing seed dormancy, possibly by repressing the translation of mRNAs that promote dormancy via P-bodies, while HSP101 and COLD SHOCK PROTEIN 2 (CSP2) promote germination under abiotic stresses (Hong & Vierling, 2001; Koornneef *et al.*, 2002; Park *et al.*, 2009; Li *et al.*, 2015). As mentioned above, the GTS is marked by mRNAs that are translationally down regulated (Bai *et al.*, 2017). These could be mRNAs that are remnants from maturation, storage proteins or proteins that inhibit germination and thus needs to be degraded (Xu *et al.*, 2006). DCP5 may play a role in the decay of these mRNAs via P-bodies during the GTS especially as the RP as larger granules were observed at this stage (Fig 4b) (Xu & Chua, 2009). Additionally, several stress granule markers were identified including the TUDOR-SN protein (TSN1/2) and PAPB2. TSN1/2 has been implied to promote seed germination under salt stress by modulating the mRNA levels of the key GA biosynthesis enzyme *GA20ox3* (Liu *et al.,* 2010). Stress granule marker PABP2 formed stress granules after a short heat stress. Other GTS-RBPs, like APUM5, COLD SHOCK PROTEIN 1 (CSP1) and

For Peer Review

335 AtRZ1 have been reported to negatively regulate germination under abiotic stress
336 conditions (Kim *et al.*, 2007; Park *et al.*, 2009; Huh & Paek, 2014). The presence of RBPs
337 that repress translation or inhibit germination during the GTS, may indicate that during
338 germination, seeds are prepared for quick responses to environmental changes. All together
339 this study provides the first step towards understanding the role of RPBs in the translational
340 control of mRNAs during the GTS, which is important to ensure successful radicle
341 protrusion and thereby completion of germination.

For Peer Review

Materials and Methods

Plant materials

Fully after ripened seeds of *Arabidopsis thaliana* accession Columbia-0 (Col-0) were used for all assays described in this manuscript (NASC N60000). The mutant line of Hyaluronan/mRNA binding protein (AT5G47210) was obtained from NASC (SALKseq_055953). The complementation lines contain the genomic fragment (forward primer AGGAGGAGGAGGAGAGAA and reverse primer: TCGCAGAAAAGACCTTCA) with its native promoter transferred to the mutant backgrounds using entry vector pDONR207 and the destination vector pKGW-RedSeed (https://gatewayvectors.vib.be/collection/pkgw-redseed). The pPABP2-PABP2-RFP reporter lines in wild type and *hsp101* background were described in Merret et al., 2017. The pUBQ-DCP5-GFP like was a kind gift from the lab of Scheer, Hélène (Scheer *et al.*, 2021) while the *hsp101* mutant and complementation line used for the germination phenotypes of Fig S5 were a kind gift from Elizabeth Vierling (McLoughlin *et al.*, 2019).

Germination condition and assays

Seeds were sowed on two layers of blue blotter paper (Anchorpaper company, www.seedpaper.com) were equilibrated with 48ml of demineralized water in plastic trays (15X21cm). Each replicate contained 1.2g of seeds which were wrapped in a closed transparent plastic bag and placed at 22°C in continuous light (143µm m-2s-1) for germination. The time-points for the germination translational shift were selected based on the physiological stage of the seeds described previously (7). In this study TR occurred at 26 hours after imbibition (HAI) and RP at 42 HAI.

To determine the DSDS50 values, germination assays were carried from 3 days until 5 weeks after harvest, when the seeds were fully after-ripened (100% germination). The germination experiments were performed as described above, however at 26°C instead of at 22°C, since these suboptimal germination conditions allowed to also identify smaller differences in dormancy level (Alonso-Blanco *et al.*, 2003). The germination percentages were calculated using the germinator software package (Joosen *et al.*, 2010) and the DSDS50 levels were calculated using the statistical program R version 2.14 (R Development Core Team, 2009; www.r-project. org) (He *et al.*, 2014).

For Peer Review

Embryo isolation and UV crosslinking

For embryo isolation, the imbibed seeds were scraped from the tray and pressed between two microscope slides. Due to the pressure applied, the embryos were expelled out of the seed coat. The embryo-seed coat mixture was separated in a 40% sucrose solution. Using centrifugation, the mixture was separated and the top layer containing the pure embryos was collected. The embryos were spread evenly over a germination tray containing white Whatmann filter papers to absorb the sucrose solution (Lopez-Molina, personal communication).

For in-vivo crosslinking (CL), the trays were placed on ice and irradiated in a Stratalinker (Stratagene) with 254nm UV light at 1J/cm2. The crosslinking was performed twice with 30 seconds pause in between treatments. The controls were processed simultaneously. The embryos were harvested immediately after irradiation and frozen in liquid N2.

The frozen embryo tissue was ground into fine powder in liquid N2 and resuspended in tubes with 24 ml of a modified seed RBP extraction buffer (1.25% Sucrose, 400mM Tris-HCl pH 8, 0.5% LiDS, 200mM LiCl, 35mM MgCl2, 1mM EGTA, 5mM DTT, 20U/ml RNasin, 1X EDTA-free Complete Protease cocktail inhibitor tablet). The tubes containing the lysate were placed on ice for 10 minutes following which they were centrifuged for 20 minutes at 14000 rpm to precipitate the cell debris. The supernatant (20ml) from each tube was transferred to fresh RNase free tube. Aliquots from the lysate were taken for quality controls (silver stain, western blots) and for mRNA enrichment check.

The mRNA-protein complexes were isolated using 1.5 ml of oligo(dT)25 magnetic beads (New England Biolabs) per tube. The beads were equilibrated using 5ml of wash buffer 1 (20mM Tris-HCl pH 7.6, 0.1% LiDS, 500mM LiCl, 1mM EDTA, 5mM DTT) and incubated for 2 minutes with gentle rotation at 4°C. The tubes were placed on the magnetic rack, which resulted in the magnetic capture of the beads and a clear suspension. Thereafter the supernatant of the magnetic beads was discarded and the cell lysate was immediately added to the tubes and incubated at 4°C for 1 hour by applying gentle rotation. Beads were collected on the magnet and washed twice with 15mL of ice-cold wash buffer1, buffer 2 (20mM Tris-HCl pH 7.6, 500mM LiCl, 1mM EDTA, 5mM DTT) and wash buffer 3 (20mM Tris-HCl pH 7.6, 200mM LiCl, 1mM EDTA, 5mM DTT) for 5 mins at room

For Peer Review

temperature. Finally, the beads were incubated with 500µL of elution buffer (20mM Tris-HCl pH 7.6, 1mM EDTA) at 50°C for 3 minutes to release the poly(A)-tailed RNAs from the beads. Two additional rounds of pulldown were performed for each sample, and the three eluates were combined in a new RNase free tube (total volume 1.5 ml).

mRNA enrichment check using qRT-PCR

Aliquots taken of the total Input and after poly-A pulldown samples were spiked with a mix of the four eukaryotic poly(A) RNAs (Affymetrix, Santa Clara, CA, USA; Ambion, P/N900433), and purified with TriPure Isolation Reagent (Roche, Basel, Switzerland). cDNA was synthesized using the iScript™ cDNA synthesis kit (Bio-Rad, Hercules, CA, USA) according to the manufacturer's protocol. qRT-PCR was performed using Power SYBR Green (Applied Biosystems, Waltham MA, USA) in a 10µl reaction using the standard program of the ViiA™ 7 instrument (Applied Biosystems). To quantify RNA levels, the comparative Ct method, namely the $2^{-\Delta\Delta Ct}$ method was used and normalized to the geometric mean of the spike-in standards (Livak & Schmittgen, 2001).

RNA Quantification and Normalization for SDS page loading

The pooled eluates were quantified using the NanoDrop spectrophotometer (260/280 ratios between 1.7-2.0). All samples were normalized for mRNA quantity in both time-points and for each replicate using the elution buffer.

RNase treatment and Protein concentration

The mRNA was digested by adding 100 units of the commercially available RNase cocktail containing RNase A and T1 to the eluates. The samples were mixed and incubated at 37°C for 1 hour along with a negative control sample. After the RNase digestion the samples were concentrated using Amicon® centrifugal filter units (0.5mL, 3kDa). Each sample was concentrated to approximately 40 µL in low-binding Eppendorf tubes.

SDS-PAGE, Silver Staining and Immunoblot

20 µL of the concentrated protein samples mixed with of 5x SDS loading dye were loaded on a 12% Bis-Tris protein gel (Thermo Fisher). The gel was run at 100V until the loading dye reached the end of the resolving gel. The SDS page gel was washed twice with ultra-pure water for 5 minutes each time. The silver staining was performed using LCMS-MS compatible silver staining protocol (34).

For Review

For western blotting, following the SDS-PAGE, the gels were electroblotted on to PVDF membranes (Trans-Blot Turbo Mini 0.2 um PVDF transfer packs, BIO-RAD). The membranes were blocked with 5% nonfat milk in 1x TBST (1x TBS with 0.1% Tween 20) for 1 hour at room temperature, followed by an overnight incubation at 4°C with primary antibodies in 3% nonfat milk with rotation. The membrane was incubated with secondary in 3% nonfat milk in 1x TBST for 1 hour at room temperature. Protein signals were detected using a high sensitivity ECL substrate and visualized using the Chemidoc (BIO-RAD). The primary antibodies used were Anti-AGO1 (Agrisera; AS09-527), Plant Anti-Actin (Agrisera; AS13 2640), Anti-HSP101 (Kind gift from Elizabeth Vierling, Amherst, Massachusetts) and Anti- V-ATPase subunit A (Agrisera; AS09467). The secondary Antibody used was HRP-conjugated Anti-Rabbit IgG concentrate (Item I1) (Sigma Aldrich; RABHRP1).

Sample Preparation for Proteomics

The gel lanes were cut out per sample. The lanes were cut such that it did not include the RNAse enzyme bands present in the lane. Each lane was cut into tiny pieces and divided equally over 3 Eppendorf tubes. The gel pieces were washed with milliQ water and 100% Acetonitrile (ACN). For reduction and alkylation, the gel pieces were incubated with 100 µL of 10mM DTT in 50mM Ammonium bicarbonate (pH 7.6) at 56°C for 45 minutes. The samples were brought back to room temperature. Supernatant was removed and gel pieces were washed with 50% ACN. Following this, 100 µL of 54mM iodoacetamide was added to the gel pieces. The tubes were incubated at room temperature for 20 minutes in the dark. The gel pieces were washed three times using 100% ACN and Ammonium bicarbonate alternatively. After the last wash with ACN, the gel pieces were shortly dried on air and then incubated overnight with 10ng of trypsin in 50mM Ammonium bicarbonate (pH 7.6) at 37°C for protein digestion. The next day the peptides were extracted from the gel twice with 50 µL 50% ACN and 100% ACN. Next, the pooled extracts were vacuum dried for 2h and the dried pellets were dissolved in 0.1% formic acid and used for MS.

For the Input total protein samples, 50µg of the lysate was used in a total volume of 25µL. Following this step 3µL of iodoacetamide was added to lysate and incubated at room

For Peer Review

temperature for 20 minutes in the dark. Next, 3µl of 12% phosphoric acid was added to the sample. To prepare lysates that contain detergents like Lithium dodecyl sulphate for nana-LCLMS/MS, the the S-Trap™ Micro spin columns= digestion protocol was used (https://protifi.com/pages/s-trap) according to the manufacturer's protocol. An overnight on column protein digestion was performed on the samples, using 1.5µg of trypsin in 50mM Ammonium bicarbonate (pH 7.6) at 37°C. The next day the peptides were eluted from the columns with 35 µL of 50% ACN and 0.2% formic acid. Next, the pooled extracts were vacuum dried for 2h and the dried pellets were dissolved in 0.1% formic acid and used for MS.

Liquid Chromatography-Tandem MS Analysis

Samples were analyzed on an LTQ-Orbitrap Velos Pro mass spectrometer (Thermo Scientific) coupled to a nanoAcquity UPLC system (Waters). Peptides were loaded onto a trapping column (nanoAcquity Symmetry C18, 5 µm, 180 µm × 20 mm) at a flow rate of 15 µL/min with solvent A (0.1% formic acid). Peptides were separated over an analytical column (nanoAcquity BEH C18, 1.7 µm, 75 µm × 200 mm) at a constant flow of 0.3 µL/min using the following gradient: 3% solvent B (acetonitrile and 0.1% formic acid) for 10 min, 7 to 25% solvent B within 210 min, 25 to 40% solvent B within 10 min, and 85% solvent B for 10 min. Peptides were introduced into the mass spectrometer using a Pico-Tip Emitter (360 µm outer diameter × 20 µm inner diameter, 10 µm tip; New Objective). MS survey scans were acquired from 300 to 1700 m/z at a nominal resolution of 30,000. The 15 most abundant peptides were isolated within a 2D window and subjected to tandem MS (MS/MS) sequencing using collision-induced dissociation in the ion trap (activation time, 10 ms; normalized collision energy, 40%). Only 2+/3+ charged ions were included for analysis. Precursors were dynamically excluded for 30 s (exclusion list size was set to 500)

Peptide and protein Identification

Raw data were processed using MaxQuant (version 1.6.1) (Cox & Mann, 2008).MS/MS spectra were searched against the Araport11 Arabidopsis database (input proteome version 11/07/2015 including 50.164 entries) concatenated to a database containing protein sequences of common contaminants. Enzyme specificity was set to trypsin/P, allowing a

maximum of two missed cleavages. Cysteine carbamidomethylation was set as fixed modification, and methionine oxidation and protein N-terminal acetylation were used as variable modifications. The minimal peptide length was set to six amino acids and a minimum of one unique peptide was required for the identification. The mass tolerances were set to 20 ppm for the first search, 6 ppm for the main search, and 0.05 Da for product ion masses. FDRs for peptide and protein identification were set to 1%. Match between runs (time window 2 min) and requantify options were enabled, as well as the IBAQ function.

Definition of GTS-RBPs and Candidate RBPs

The proteinGroups.txt output from MaxQuant was further processed in Perseus version 1.6.12 from MaxQuant (Tyanova *et al.*, 2016). Proteins that were identified in at least 2 or more biological replicates of the CL treatment and with a minimum of 2 unique peptides identified the proteins were selected for further analysis. To be able to perform statistics between the NCL and CL samples, all normalized LFQ intensities were log2 transformed and the missing values were replaced by a constant minimum value of 10. Next, t-tests were performed with a Benjamin-Hochberg correction for multiple t-testing and a false discovery rate of 5% between the NCL and CL for each stage. Proteins that were statistically enriched in the CL samples were defined as the GTS-RBPs per time-point and the ones that were not statistically enriched but had a log2 fold (CL/NCL)>=1, were defined as the candidate RBP set per stage. Similar analysis was performed for the Input total protein. Here only the LFQ intensities of all proteins identified in the CL Input total protein samples were compared between the TR and RP stages.

GO and PFAM Annotation and Analysis

GO and PFAM Annotation for the proteins was performed using the Perseus tool (version 1.6.12) (Tyanova *et al.*, 2016) using the GO database and PFAM database plugins. Proteins that contained the term RNA binding in their GO annotation were categorized as the 'RNA binding' set. PFAM classification was done on the RNA binding and Not binding set by counting the number of proteins per protein family in each stage. The proteins were classified as classical or non-classical RBPs based on previous reports. GO enrichment analysis was performed using the g:Profiler tool (http://biit.cs.ut.ee/gprofiler/) (Raudvere *et al.*, 2019) using the Arabidopsis genome as a reference dataset. For statistical t-tests,

Benjamin Hochberg correction for multiple testing was chosen with 0.05 as the significance level.

Confocal Image Analysis

For visualization of all reporter lines used in this study, epidermal cells from embryonic root tips (at testa rupture and radical protrusion stages imbibed in water) were imaged with a Leica SP8 laser scanning confocal microscope equipped with 63X oil immersion objective (NA 1.4). For the heat stress treatment, embryos were excised and exposed to a short heat stress of 42°C in water for 30 minutes before loading onto a slide for visualization under the Leica SP8 confocal microscope. YFP and RFP fluorophores were excited with 488nm and 552 nm laser lines, respectively and their fluorescence emissions were collected in 515-550 nm and 580-650 nm windows respectively. For each category, 30 epidermal cells from 5 seedlings were measured (n = 30). The number of granules were quantified using Image J plugin 3D object counter (Du *et al.*, 2011). Maximum intensity of a Z projection covering a depth of 5µm deep from the cell surface was quantified. Particles with mean intensity in the upper 10^{th} percentile and within a diameter range of 20-100 pixels were measured. The data was plotted as number of granules per 1000μm^3 volume, the data normality was checked by Kolmogorov-Smirnov test and variance equality was checked by Levene's test. (***) indicate $p < 0.001$ (Student t test).

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Germination curve of the Col-0 seeds used to determine the GTS time-points

Fig. S2 qRT-PCR depicting mRNA enrichment after poly-A pulldown of mRNAs at the radicle protrusion stage of seed germination

Fig. S3 Correlation plots between replicates for CL samples at Testa Rupture and Radicle protrusion stages of Arabidopsis seed germination.

Fig. S4 Confirmation of knockout mutant *at5g47210* using qRT-PCR

Fig. S5 Germination of hsp101 mutant under control conditions**.**

For Peer Review

549 **Fig. S6** Confirmation of dynamic GTS-RBPs by western blotting.

550 **Fig. S7** Visualization of heat stress granule marker PABP2 in Col-0 and the *hsp101* mutant
551 at Testa Rupture**.**

552 **Table S1a** RNA binding proteins identified at testa rupture (TR) stage of the germination
553 translational shift

554 **Table S1b** RNA binding proteins identified at Radicle protrusion (RP) stage of the
555 germination translational shift.

556 **Table S1c** GO enrichment analysis based on Molecular function of the GTS and candidate
557 RNA binding proteins identified at Testa rupture and Radicle protrusion stages of
558 Arabidopsis seed germination

559 **Table S1d** Protein family classification of RNA binding proteins at testa rupture and
560 radicle protrusion stages of the germination translational shift

561 **Table S1e** Dataset showing RBPs that are unique and/or overlap between 5 different
562 interactome captures as shown in Figure 3

563 **Table S1f** Dataset showing all proteins identified in Input total protein samples at TR and
564 RP

565 **Table S2** Dynamic GTS-RBPs at testa Rupture and radicle protrusion during the
566 germination translational shift of seed germination

567 **Table S3** Seed specific RNA binding proteins identified by comparison with previously
568 performed interactome captures in *Arabidopsis thaliana*

569 **Acknowledgments**

570 This work is part of the research domain Applied and Engineering Sciences, project
571 number 15228, which is financed by the Dutch Research Council (NWO). This study is set
572 within the framework of the "Laboratoires d'Excellences (LABEX)" TULIP (ANR-10-
573 LABX-41) and of the "École Universitaire de Recherche (EUR)" TULIP-GS (ANR-18-

For Peer Review

574 EURE-0019). We thank Dominique Gagliardi (IBMP, France) for providing GFP-DCP5
575 transgenic line.

576 **Author contribution**

577 NS and LB planned and designed the research. NS, AB, AHPA, LAJW performed
578 experiments. RM provided transgenic lines. NS and LB wrote the manuscript. All authors
579 commented on the manuscript.

For Peer Review

References

Abel S, Blume B, Glund K. 1990. Evidence for RNA-Oligonucleotides in Plant Vacuoles Isolated from Cultured Tomato Cells. *Plant Physiology* **94**: 1163-1171.

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

Ambrosone A, Batelli G, Nurcato R, Aurilia V, Punzo P, Bangarusamy DK, Ruberti I, Sassi M, Leone A, Costa A, et al. 2015. The Arabidopsis RNA-Binding Protein AtRGGA Regulates Tolerance to Salt and Drought Stress. *Plant Physiology* **168**: 292-306.

Bach-Pages M, Homma F, Kourelis J, Kaschani F, Mohammed S, Kaiser M, van der Hoorn RA, Castello A, Preston GM. 2020. Discovering the RNA-Binding Proteome of Plant Leaves with an Improved RNA Interactome Capture Method. *Biomolecules* **10**: 661.

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

Bai B, Van Der Horst S, Cordewener JH, America TA, Hanson J, Bentsink L. 2020. Seed-stored mRNAs that are specifically associated to monosomes are translationally regulated during germination. *Plant Physiology* **182**: 378-392.

Beckmann BM, Horos R, Fischer B, Castello A, Eichelbaum K, Alleaume A-M, Schwarzl T, Curk T, Foehr S, Huber W. 2015. The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nature communications* **6**: 1-9.

Castello A, Hentze MW, Preiss T. 2015. Metabolic enzymes enjoying new partnerships as RNA-binding proteins. *Trends in Endocrinology & Metabolism* **26**: 746-757.

Castello A, Horos R, Strein C, Fischer B, Eichelbaum K, Steinmetz LM, Krijgsveld J, Hentze MW. 2013. System-wide identification of RNA-binding proteins by interactome capture. **8**: 491-500.

Chantarachot T, Bailey-Serres J. 2018. Polysomes, Stress Granules, and Processing Bodies: A Dynamic Triumvirate Controlling Cytoplasmic mRNA Fate and Function. *Plant Physiology* **176**: 254-269.

For Peer Review

Cho H, Cho HS, Hwang I. 2019. Emerging roles of RNA-binding proteins in plant development. *Current opinion in plant biology* **51**: 51-57.

Cho H, Cho HS, Nam H, Jo H, Yoon J, Park C, Dang TVT, Kim E, Jeong J, Park S, et al. 2018. Translational control of phloem development by RNA G-quadruplex–JULGI determines plant sink strength. *Nature Plants* **4**: 376-390.

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

Du G, Drexler GA, Friedland W, Greubel C, Hable V, Krücken R, Kugler A, Tonelli L, Friedl AA, Dollinger G. 2011. Spatial dynamics of DNA damage response protein foci along the ion trajectory of high-LET particles. *Radiation Research* **176**: 706-715.

Floyd BE, Mugume Y, Morriss SC, MacIntosh GC, Bassham DC. 2017. Localization of RNS2 ribonuclease to the vacuole is required for its role in cellular homeostasis. *Planta* **245**: 779-792.

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

Hong S-W, Vierling E. 2001. Hsp101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress. *The Plant Journal* **27**: 25-35.

Hubstenberger A, Courel M, Bénard M, Souquere S, Ernoult-Lange M, Chouaib R, Yi Z, Morlot J-B, Munier A, Fradet M. 2017. P-body purification reveals the condensation of repressed mRNA regulons. *Molecular cell* **68**: 144-157. e145.

Huh SU, Paek K-H. 2014. APUM5, encoding a Pumilio RNA binding protein, negatively regulates abiotic stress responsive gene expression. *BMC plant biology* **14**: 75.

Jang G-J, Yang J-Y, Hsieh H-L, Wu S-H. 2019. Processing bodies control the selective translation for optimal development of Arabidopsis young seedlings. *Proceedings of the National Academy of Sciences* **116**: 6451-6456.

For Peer Review

Joosen RV, Kodde J, Willems LA, Ligterink W, van der Plas LH, Hilhorst HW. 2010. GERMINATOR: a software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. *The Plant Journal* **62**: 148-159.

Kim Y-O, Pan S, Jung C-H, Kang H. 2007. A Zinc Finger-Containing Glycine-Rich RNA-Binding Protein, atRZ-1a, Has a Negative Impact on Seed Germination and Seedling Growth of Arabidopsis thaliana Under Salt or Drought Stress Conditions. *Plant and Cell Physiology* **48**: 1170-1181.

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

Köster T, Marondedze C, Meyer K, Staiger D. 2017. RNA-binding proteins revisited–the emerging Arabidopsis mRNA interactome. *Trends in plant science* **6**: 512-526.

Li W, Ma M, Feng Y, Li H, Wang Y, Ma Y, Li M, An F, Guo H. 2015. EIN2-Directed Translational Regulation of Ethylene Signaling in Arabidopsis. *Cell* **163**: 670-683.

Ling J, Wells DR, Tanguay RL, Dickey LF, Thompson WF, Gallie DR. 2000. Heat shock protein HSP101 binds to the Fed-1 internal light regulatory element and mediates its high translational activity. *The Plant Cell* **12**: 1213-1227.

Liu S, Jia J, Gao Y, Zhang B, Han Y. 2010. The AtTudor2, a protein with SN-Tudor domains, is involved in control of seed germination in Arabidopsis. *Planta* **232**: 197-207.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *Methods* **25**(4): 402-408.

Lorković ZJ. 2009. Role of plant RNA-binding proteins in development, stress response and genome organization. *Trends in plant science* **14**: 229-236.

Lorković ZJ, Barta A. 2002. Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant Arabidopsis thaliana. *Nucleic acids research* **30**: 623-635.

Lou L, Ding L, Wang T, Xiang Y. 2020. Emerging Roles of RNA-Binding Proteins in Seed Development and Performance. *International journal of molecular sciences* **21**(18): 6822.

For Peer Review

669 **Magnotta SM, Gogarten JP. 2002.** Multi site polyadenylation and transcriptional
670 response to stress of a vacuolar type H+-ATPase subunit A gene in Arabidopsis
671 thaliana. *BMC plant biology* **2**: 3.
672 **Maia J, Dekkers BJ, Provart NJ, Ligterink W, Hilhorst HW. 2011.** The re-
673 establishment of desiccation tolerance in germinated Arabidopsis thaliana seeds
674 and its associated transcriptome. *PloS one* **6**: e29123.
675 **Maldonado Bonilla LD. 2014.** Composition and function of P bodies in Arabidopsis
676 thaliana. *Frontiers in plant science* **5**:201.
677 **Marondedze C, Thomas L, Gehring C, Lilley KS. 2019.** Changes in the Arabidopsis
678 RNA-binding proteome reveal novel stress response mechanisms. *BMC plant*
679 *biology* **19**: 1-11.
680 **Marondedze C, Thomas L, Serrano NL, Lilley KS, Gehring C. 2016.** The RNA-binding
681 protein repertoire of Arabidopsis thaliana. *Scientific reports* **6**: 29766.
682 **McLoughlin F, Kim M, Marshall RS, Vierstra RD, Vierling E. 2019.** HSP101 Interacts
683 with the Proteasome and Promotes the Clearance of Ubiquitylated Protein
684 Aggregates. *Plant Physiology* **180**: 1829-1847.
685 **Merchante C, Brumos J, Yun J, Hu Q, Spencer KR, Enríquez P, Binder BM, Heber**
686 **S, Stepanova AN, Alonso JM. 2015.** Gene-specific translation regulation
687 mediated by the hormone-signaling molecule EIN2. *Cell* **163**: 684-697.
688 **Merchante C, Stepanova AN, Alonso JM. 2017.** Translation regulation in plants: an
689 interesting past, an exciting present and a promising future. *The Plant Journal* **90**:
690 628-653.
691 **Merret R, Carpentier M-C, Favory J-J, Picart C, Descombin J, Bousquet-Antonelli**
692 **C, Tillard P, Lejay L, Deragon J-M, Charng Y-y. 2017.** Heat shock protein
693 HSP101 affects the release of ribosomal protein mRNAs for recovery after heat
694 shock. *Plant Physiology* **174**: 1216-1225.
695 **Narsai R, Law SR, Carrie C, Xu L, Whelan J. 2011.** In-depth temporal transcriptome
696 profiling reveals a crucial developmental switch with roles for RNA processing and
697 organelle metabolism that are essential for germination in Arabidopsis. *Plant*
698 *Physiology* **157**: 1342-1362.

For Peer Review

Park SJ, Kwak KJ, Oh TR, Kim YO, Kang H. 2009. Cold shock domain proteins affect seed germination and growth of Arabidopsis thaliana under abiotic stress conditions. *Plant and Cell Physiology* **50**: 869-878.

Perino C, Côme D. 1991. Physiological and metabolical study of the germination phases in apple embryo. *Seed Science and Technology* **19**: 1-14.

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. 2019. g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic acids research* **47**: W191-W198.

Reichel M, Liao Y, Rettel M, Ragan C, Evers M, Alleaume A-M, Horos R, Hentze MW, Preiss T, Millar AA. 2016. In planta determination of the mRNA-binding proteome of Arabidopsis etiolated seedlings. *The Plant Cell* 2**8**: 2435-2452.

Sablok G, Powell J, Kazan K. 2017. Emerging roles and landscape of translating mRNAs in plants. *Frontiers in Plant Science* **8**: 1443.

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

Scheer H, de Almeida C, Ferrier E, Simonnot Q, Poirier L, Pflieger D, Sement FM, Koechler S, Piermaria C, Krawczyk P. 2021. The TUTase URT1 connects decapping activators and prevents the accumulation of excessively deadenylated mRNAs to avoid siRNA biogenesis. *Nature communications* **12**: 1-17.

Soppe WJ, Bentsink L. 2020. Seed dormancy back on track; its definition and regulation by DOG1. *New Phytologist*.

Sorenson R, Bailey-Serres J. 2014. Selective mRNA sequestration by OLIGOURIDYLATE-BINDING PROTEIN 1 contributes to translational control during hypoxia in Arabidopsis. *Proceedings of the National Academy of Sciences* **111**: 2373-2378.

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J. 2016. The Perseus computational platform for comprehensive analysis of (prote) omics data. *Nature Methods* **13**: 731-740.

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

For Peer Review

730 **Xu J, Chua N-H. 2009.** Arabidopsis decapping 5 is required for mRNA decapping, P-
731 body formation, and translational repression during postembryonic development.
732 *The Plant Cell* **21**: 3270-3279.
733 **Xu J, Yang J-Y, Niu Q-W, Chua N-H. 2006.** Arabidopsis DCP2, DCP1, and VARICOSE
734 form a decapping complex required for postembryonic development. *The Plant*
735 *Cell* **18**: 3386-3398.
736 **Zhang Z, Boonen K, Ferrari P, Schoofs L, Janssens E, van Noort V, Rolland F,**
737 **Geuten K. 2016.** UV crosslinked mRNA-binding proteins captured from leaf
738 mesophyll protoplasts. *Plant Methods* **12**: 42.

739

740

For Peer Review

741

For Peer Review

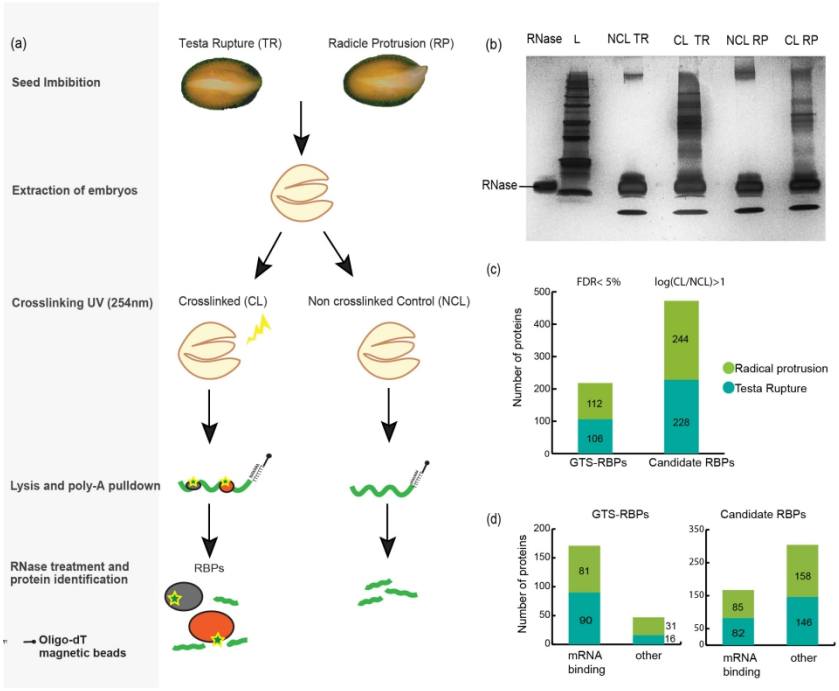


Fig. 1 mRNA interactome capture of the Arabidopsis germination translational shift. (a) Schematic representation of mRNA interactome capture at Testa rupture (TR) and Radicle protrusion (RP), the two stages that define the germination translational shift (GTS). (b) A silver stained SDS page gel showing the RNAse enzyme control at the left side of the protein ladder (L) and to the right side are the mRNA-protein complexes that were isolated from the non-crosslinked (NCL) and crosslinked (CL) samples of the TR and RP stages. Results are representative of three independent interactome capture experiments with three biological replicates. (c) Bar graphs representing GTS-RBPs which were proteins identified with high confidence FDR<0.5 and candidate RBPs that show log2 (CL/NCL)>1 enrichment at TR and RP. (d) Categorization of the GTS-RBPs and candidate RBPs based on the Gene ontology term 'RNA binding'.

212x148mm (300 x 300 DPI)

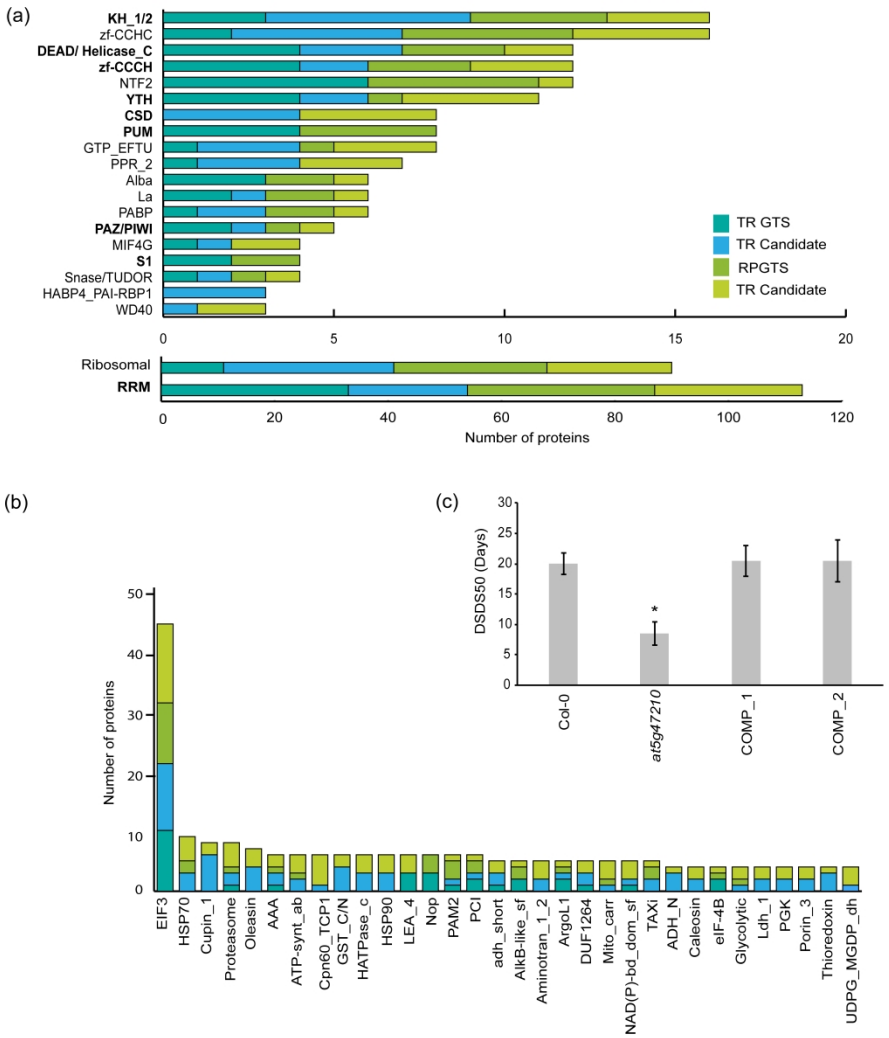


Fig. 2 Proteins domain classification of the identified GTS and candidate RBPs in the germination translational shift of Arabidopsis thaliana. (a) Classical and Non-classical RNA binding domains (RBDs) at Testa Rupture (TR) and Radicle protrusion (RP). The classical RBDs are indicated in bold (families with >=3 proteins depicted in figure). (b) Putative RBDs at TR and RP (families with >4 proteins depicted in figure). (c) Graph representing the days of dry seed storage to reach 50% germination (DSDS50). A mutant of the hyaluronan/mRNA binding protein AT5G47210 (at5g47210) and two complementation lines (COMP_1 and COMP_2) were analysed for their DSDS50 compared to the wild-type Col-0 at 26°C. The results are representative averages of four biological replicates (SE, t-test, p<0,05; error bar).

1979x2101mm (72 x 72 DPI)

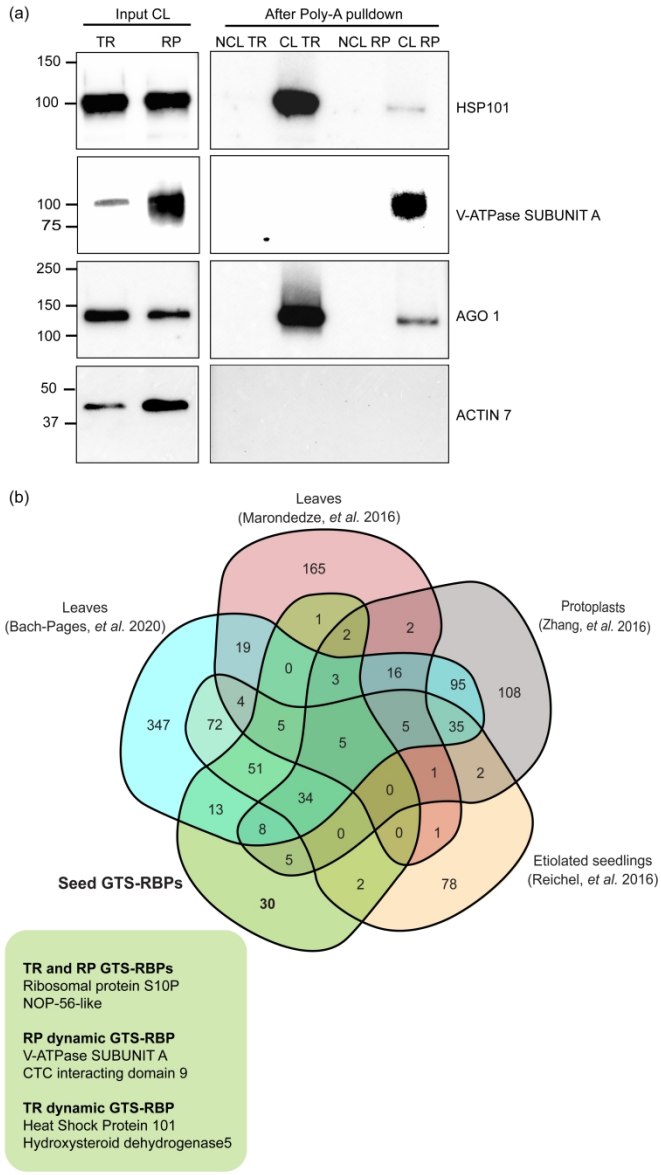


Fig. 3 Dynamic GTS-RBPs and seed specific RBPs identified during the germination translational shift of Arabidopsis thaliana. (a) Western blot image confirming the dynamic nature of GTS-RBPs HSP101 at the Testa rupture (TR) stage and V-ATPase subunit A at the Radicle protrusion (RP) stage after the poly-A pulldown. AGO1 was used as a known RBP control, while ACTIN 7 as a non-RBP negative control. The non-crosslinked (NCL) and crosslinked (CL) samples were normalized based on the mRNA quantity after the poly-A pulldown, while the total protein input for the CL samples were loaded with a fixed volume of the total protein from the CL lysates. (b) Venn Diagram comparing the GTS-RBPs identified in this study and previously performed mRNA interactome captures in different plant tissues. The green box shows representative seed specific GTS-RBPs that were either identified at both stages or dynamic for the Testa rupture (TR) or Radicle protrusion (RP) stages.

1108x1982mm (72 x 72 DPI)

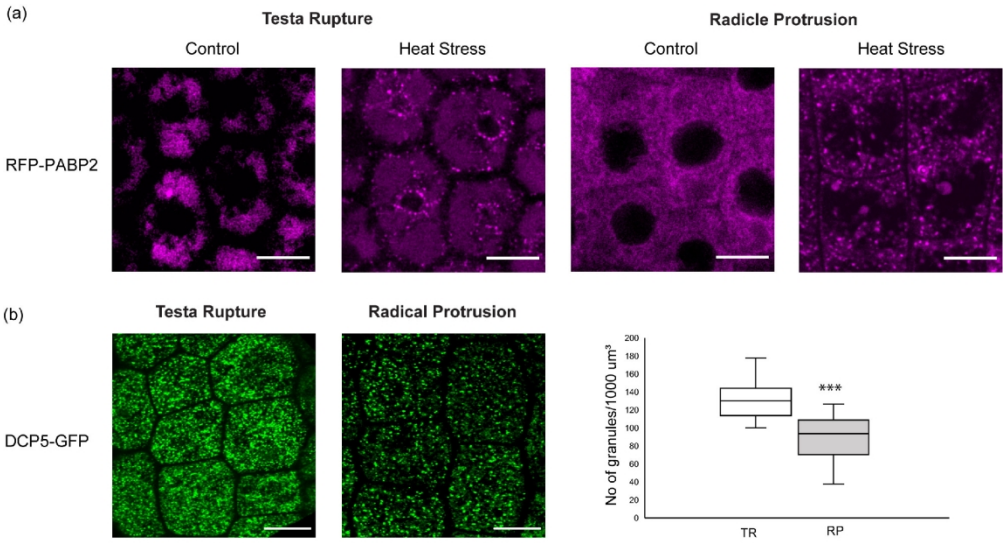


Fig. 4 Visualization of P-bodies and stress granules at the Testa rupture and Radicle protrusion stages of Arabidopsis thaliana seed germination. (a) Visualization of stress granules using dual reporter line pDCP1-YFP-DCP1/pPABP2-tRFP-PABP2 at TR (background bodies are large vacuoles commonly present at this stage) and RP at optimal germination conditions (control) or under short heat stress of 30 minutes at 42°C (b) Visualization of P-bodies using reporter line pUBQ-DCP5-GFP at radicle protrusion (RP) and testa rupture (TR). Box-plot showing the number of granules /1000 um. (n= 30 root epidermal cells and 5 embryos per stage, scale bar = 10 micron).

168x94mm (300 x 300 DPI)

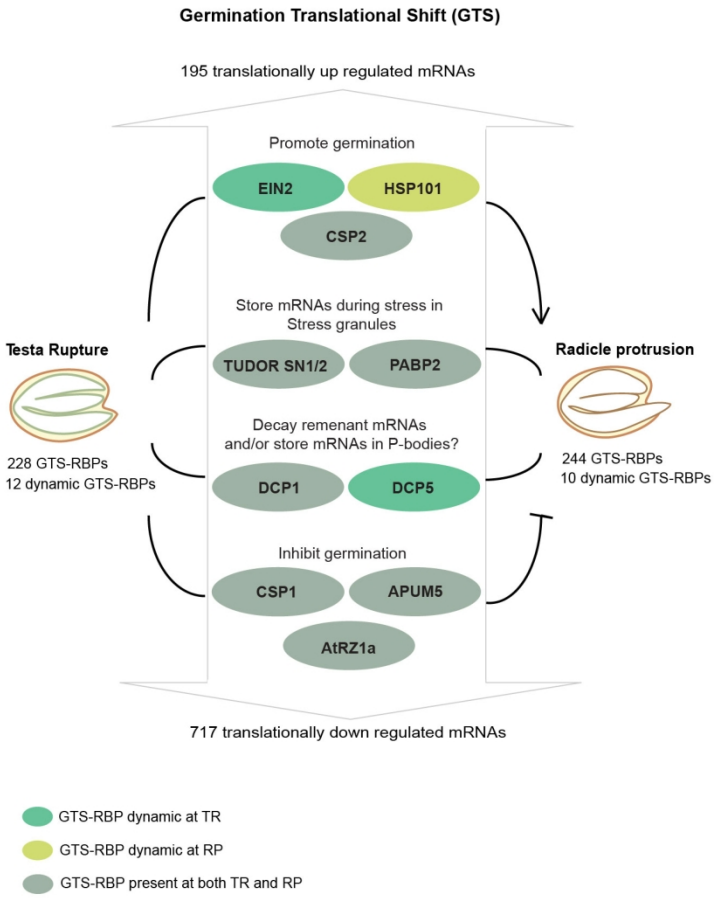


Fig. 5 Summary of the features at the Germination Translational Shift (GTS). The number of RNA binding proteins that are identified at testa rupture and radicle protrusion are indicated at the left and right side of the figure, respectively. The box in the middle of the figure presents the GTS-RBPs (ovals) that play a role in germination and may regulate the translational of mRNAs during the GTS. The numbers indicated in the top and bottom of the figure represent the mRNAs that are under translational control identified by Bai et al., 2017.

156x144mm (300 x 300 DPI)