Dissecting the seed-to-seedling transition in *Arabidopsis thaliana* by gene co-expression networks

Anderson Tadeu Silva

Thesis committee

Promotor
Prof. Dr Harro J Bouwmeester
Professor of Plant Physiology
Wageningen University, The Netherlands

Co-promotor
Dr Henk WM Hilhorst
Associate Professor, Laboratory of Plant Physiology
Wageningen University, The Netherlands

Dr Wilco Ligterink
Researcher, Laboratory of Plant Physiology
Wageningen University, The Netherlands

Other members
Prof. Dr G.C. Angenent, Wageningen University
Prof. Dr D. de Ridder, Wageningen University
Prof. Dr J.C.M. Smeekens, Utrecht University, The Netherlands
Dr S.P.C. Groot, Wageningen University

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences (EPS)

Dissecting the seed-to-seedling transition in *Arabidopsis thaliana* by gene co-expression networks

Thesis

submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Wednesday December 9th, 2015
at 1.30 p.m in the Aula.

Anderson Tadeu Silva
Dissecting the seed-to-seedling transition in *Arabidopsis thaliana* by gene co-expression networks
179 pages.

PhD thesis, Wageningen University, Wageningen, NL (2015)
With references, with summaries in English and Dutch

ISBN 978-94-6257-592-9

CONTENTS

To my beloved family
Dedico esta tese à minha amada família

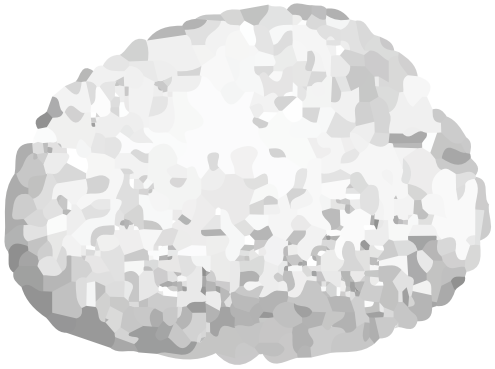
on writing a PhD thesis

"A ambição torna os homens audazes; a audácia sem ambição é privilégio de poucos."

Carlos Drummond de Andrade

Chapter 1

General Introduction



Phase transitions of plants

Plants undergo a number of developmental phase transitions during their life cycle. The transitions between phases are controlled by distinct genetic circuits that integrate endogenous and environmental cues (Rougvie, 2005; Amasino, 2010; Huijser and Schmid, 2011). The correct timing of events occurring in the post-embryonic developmental phase transitions (i.e. germination, the heterotrophic-to-autotrophic transition, juvenile vegetative to adult vegetative and vegetative to reproductive) is critical for plant survival and reproduction. For example, in *Arabidopsis thaliana*, glucose plays an important role in controlling the developmental phase transition from heterotrophic to photoautotrophic seedlings (Xiong *et al.*, 2013). Glucose controls the TARGET OF RAPAMYCIN (TOR) signalling transcriptional network, which represses programs associated with seed nutrient metabolism that is required for germination and stimulates root meristem activation. Evolutionarily, TOR is conserved in almost all life forms from yeasts to plants and humans (Wullschleger *et al.*, 2006; Dobrenel *et al.*, 2011; Russell *et al.*, 2011). Also in insects, phase transitions are important and TOR also plays an important role in the larval/pupal phase transition of *Drosophila*, where the developmental delay caused by nutritional restriction is reversed by activating TOR (Layalle *et al.*, 2008).

Another considerable advance in understanding the networks controlling phase transitions in Arabidopsis was made by studying developmental stages during seed development (Belmonte *et al.*, 2013), the networks regulating the phase transition from dormant to non-dormant seeds (Bassel *et al.*, 2011) and from quiescent seed to seed germination (Dekkers *et al.*, 2013). However, little is known about the regulation of networks involved in the developmental phase transition from seed to seedling.

It is advantageous for a species to keep the period of seedling establishment as short as possible since young seedlings are highly sensitive to biotic and abiotic stresses. Seedling death is one of the main causes of yield losses in crops growing under sub-optimal conditions (Zhang *et al.*, 2012; Cook *et al.*, 2014). In depth analysis of transcriptional and post-transcriptional changes during the seed-to-seedling transition and their implication for stress resistance should provide insight into the genetic regulation of these complex traits. Such insight can then be used to understand processes that are involved in the transition from seed to seedling and the effect on tolerance of stress (i.e. dehydration) during this phase transition. Understanding the regulatory networks that control the seed-to-seedling transition is a first step towards developing strategies for the improvement of crop yield.

Developmental switches between the quiescent dry seed stage and emergence of a photoautotrophic seedling

Germination is a complex process in which the seed must shift from a quiescent state (dry seed) to germination in order to prepare for vigorous seedling establishment. Regulation of this critical transition from a quiescent seed to a photoautotrophic seedling is controlled, partly, by environmental cues such as light (Beligni and Lamattina, 2000; Eastmond *et al.*, 2000), and temperature (Lu and Hills, 2002; Divi and Krishna, 2010; Zhao *et al.*, 2010), together with the antagonistic interaction between abscisic acid (ABA) and gibberellins (GAs) (Bewley, 1997; Koornneef *et al.*, 2002). Basically, germination and early seedling establishment can be described as a sequential time course (Bewley *et al.*, 2013), and can be divided in several morphological distinct stages, marking the seed-to-seedling transition. These are: (*i*) seed rehydration (imbibition), germination-related initiation of metabolic activity and embryo swelling, (*ii*) testa rupture, followed by (*iii*) protrusion of the radicle through the endosperm which marks the completion of seed germination *sensu stricto*. Subsequently, a further increase in water uptake occurs and the embryonic root extends further as cell division commences, followed by (*iv*) the appearance of root hairs, (*vi*) greening of the cotyledons and (*vii*) complete opening of the cotyledons (Nonogaki *et al.*, 2010; Bewley *et al.*, 2013). The developmental phase transition from seed to seedling is also marked by differentiation of etioplasts into chloroplasts in the cotyledons, allowing the seedling to acquire photosynthetic competence (Shimada *et al.*, 2007). Moreover, progression through the seed-to-seedling developmental stages in Arabidopsis is accompanied by mobilization of stored reserves (Penfield *et al.*, 2005; Fait *et al.*, 2006) and by the loss of desiccation tolerance (DT) (Maia *et al.*, 2011).

Once germination has commenced, the consumption of reserves accumulated during seed maturation is necessary for energy production to ensure heterotrophic growth (Fait *et al.*, 2006; Carrera *et al.*, 2007; Bassel *et al.*, 2008). This reserve mobilization phase occurs prior to the greening of the cotyledons and results in depletion of the storage reserves, making the shift from heterotrophic to autotrophic metabolism necessary for successful seedling establishment (Mansfield and Briarty, 1996; Allen *et al.*, 2010). Interestingly, a number of studies in Arabidopsis have shown that germinating seeds can still prevent an irreversible shift in development to a next stage (Maia *et al.*, 2011) and to adult growth if the environmental conditions are unfavourable (Sato *et al.*, 2009).

Point of no return in the seed-to-seedling phase transition

The phase transition from seed to autotrophic seedling is characterized by two main events; the first is defined by the emergence of the radicle and the second is defined by seedling formation. During the phase transition, growth of the seedling can be inhibited if external environmental conditions become unfavourable or potentially lethal. This plant developmental arrest responds directly to the osmotic balance in germinating seeds (Buitink *et al.*, 2003; Maia *et al.*, 2011) and to the carbon-nitrogen (C/N) balance in post-germination seedlings (Kang and Turano, 2003). One mutant line, *cni1-D* (carbon/nitrogen insensitive 1-dominant) isolated from the a large collection of FOX (*full*-lenght cDNA *overexpressor* gene) transgenic plants, was able to survive post-germination growth arrest in the presence of a high C/N balance (Sato *et al.*, 2009). These authors demonstrated that the RING-H2-type ubiquitin ligases ARABIDOPSIS TOXICOS EN LEVADURA 31 (ATL31)/CARBON NITROGEN INSENSITIVE 1 (CNI1) and its paralogue ATL6 are the key elements involved in this C/N response during early seedling establishment. High expression of *ATL31* in the *cni1-D* indicated that this gene was responsible for this mutant survive on high C/N stress (300 mM glucose/0.1 mM). Additionally, a knock-out of *ATL31* resulted in hypersensitivity to this C/N conditions during seedling growth (Sato *et al.*, 2009). This result highlights the importance of protein ubiquitination and degradation for the maintenance of seedling growth during the seed-to-seedling transition.

The shift from seed to seedling can also be arrested by ABA (Lopez-Molina *et al.*, 2001). A developmental window of ABA sensitivity after germination was identified, in which seedling growth can be arrested and tolerance to dehydration re-acquired (Maia *et al.*, 2011). During this dehydration, *ABA INSENSITIVE 3* (*ABI3*) and *ABA INSENSITIVE 5* (*ABI5*) are necessary to acquire osmotolerance (Lopez-Molina *et al.*, 2001; Perruc *et al.*, 2007). Expression of *ABI3* and *ABI5* is down-regulated during germination, which suggests that seeds/seedlings lose their ability to respond to dehydration during this process (Perruc *et al.*, 2007). During the early stages of germination, however, ABA exposure maintains *ABI3* and *ABI5* expression (Lopez-Molina *et al.*, 2001; Perruc *et al.*, 2007) which results in dehydration- and osmotolerance and developmental stage arrest if these conditions occur.

Desiccation tolerance

Seeds frequently face unfavourable environmental conditions, such as low water levels (drought and desiccation), which may limit seeds to germinate and further establish as a seedling. Two types of tolerance based on a critical water level are well studied: (*i)* drought tolerance, which is the capacity to tolerate moderate dehydration,

to moisture contents below ~0.3g H_2O per gram of dry weight (Hoekstra *et al.*, 2001; Tripathi *et al.*, 2014); and (*ii*) desiccation tolerance (DT), or anhydrobiosis, which is the tolerance of further dehydration, down to water levels below 0.1g H_2O per gram of dry weight and successive rehydration without permanent damage (Hoekstra *et al.*, 2001; Maia *et al.*, 2011; Dekkers *et al.*, 2015).

The ability of germinating seeds to tolerate low water levels is limited to a short developmental time window (Buitink *et al.*, 2003; Maia *et al.*, 2011), during which plants monitor the environmental osmotic status before initiating vegetative growth (Lopez-Molina *et al.*, 2001). Within the developmental window between completion of germination and seedling establishment, the genetic program for germination can, at least partly, be reverted to the seed maturation program by application of osmotic treatment and/or ABA (Maia *et al.*, 2011; Maia *et al.*, 2014; Dekkers *et al.*, 2015). This reversion is no longer possible once a *point-of-no-return* has been reached. This *point-of-no-return*, therefore, marks a switch between (seed) vegetative phase. This *point-of-no-return* is strictly dependent on the developmental stage during the seed-to-seedling transition (Buitink *et al.*, 2003; Maia *et al.*, 2011). In Arabidopsis and *Medicago truncatula*, for example, this *point-of-no-return* corresponds to a developmental stage prior to the appearance of root hairs, in which DT could be re-established in germinated seeds (Buitink *et al.*, 2003; Maia *et al.*, 2011). The developmental window in which DT can be re-established coincides with the ABA sensitivity window: from the root hair stage onwards, ABA is unable to block development and, concomitantly, the ability to re-induce DT is dramatically reduced (Maia *et al.*, 2014).

DT is a physiological trait acquired during seed development (Verdier *et al.*, 2013) and very important for survival of dry seeds. Moreover, DT, as well as the expression of *ABI3* and *ABI5*, are lost during germination (Lopez-Molina *et al.*, 2001; Buitink *et al.*, 2003; Maia *et al.*, 2011). DT and the expression of *ABI3* and *ABI5*, however, can be re-acquired in germinated seeds prior to root hair formation upon treatment with ABA and/or PEG (Maia *et al.*, 2014). Some studies support the hypothesis that ABA is important for DT induction and maintenance of seeds in the protective state (Buitink *et al.*, 2003; Verdier *et al.*, 2013). In conclusion, the transition from seed to seedling is not only controlled by regulation of germination, but also by a checkpoint after germination until when seedlings can still cope with prolonged dehydration. Ultimate seedling establishment is, thus, controlled by this *point-of-no-return*, prior to which seedling development can be blocked by a lack of water and resumed when water availability is adequate. The re-acquisition of DT in germinated seeds may be an effective strategy to arrest growth under unfavourable conditions, and allow a vigorous seedling establishment when conditions turn optimal. However, this suggestion awaits definitive proof.

Abscisic acid and dehydration stress signalling

The five major hormone classes (i.e. ABA, auxin, cytokinins, ethylene, brassinosteroids and gibberellins) elicit a wide range of responses in plant systems (Ulmasov *et al.*, 1997; Rademacher, 2000; Achard *et al.*, 2006; Umehara *et al.*, 2008; Bari and Jones, 2009; Miransari and Smith, 2014). ABA is one of these and is widely known as a hormone with a role in stress responses (Bartels and Sunkar, 2005). Extensive studies using biochemical and molecular-genetic approaches have revealed the main framework of ABA biosynthesis and catabolism pathways (Finkelstein *et al.*, 2002; Nambara *et al.*, 2002; Nambara and Marion-Poll, 2005) as well as ABA transporters and over hundred ABA signalling components (Cutler *et al.*, 2010).

The role of ABA in seeds (Okamoto *et al.*, 2006; Verdier *et al.*, 2013) and seedlings (Nambara *et al.*, 2002; Kondo *et al.*, 2014) has been studied widely. During seed maturation ABA is responsible for the acquisition of DT (Verdier *et al.*, 2013) and during germination for the re-acquisition of DT upon mild osmotic stress (Maia *et al.*, 2014). The levels of ABA and ABA sensitivity determine the response of plants to the hormone (Xiong and Zhu, 2003). It is commonly accepted that increased ABA levels trigger ABA-mediated stress responses. However, in desiccation sensitive Arabidopsis seeds the application of a mild osmotic stress did not increase ABA content but genes encoding ABA receptor proteins were upregulated (Maia *et al*, 2014). Similarly, Arabidopsis plants expressing a cucumber mosaic virus (CMV) factor showed an increase in drought tolerance, mediated by the 2b protein, without accumulation of ABA (Westwood *et al.*, 2013). This suggests that besides enhancing ABA levels, adjustment of ABA signalling and/or perception can be sufficient to induce a proper stress response (Maia *et al*, 2014).

Measured ABA contents are the result of both ABA biosynthesis and catabolism. ABA catabolism is largely categorized into two types of biochemical reactions, namely hydroxylation and conjugation (Nambara and Marion-Poll, 2005). The most common product of ABA hydroxylation is 8′-hydroxy-ABA (Nambara and Marion-Poll, 2005). ABA-8′-hydroxylases are encoded by the *CYP707A* genes, members of the large cytochrome P450 family (Saito *et al.*, 2004), and the spatial and temporal differences in expression patterns of each *CYP707A* gene suggest different developmental and/or physiological role(s) (Okamoto *et al.*, 2006). For example, *CYP707A1* and *CYP707A2* have a role in ABA catabolism during seed development whereas *CYP707A2* also plays a role during germination (Okamoto *et al*., 2006). It has been shown also that the *cyp707a3-1* mutant contains higher ABA levels and has increased tolerance to dehydration in Arabidopsis (Umezawa *et al.*, 2006) and apple seedlings (Kondo *et al.*, 2014). The key regulatory enzymes for ABA biosynthesis

are the 9-*cis*-epoxycarotenoid dioxygenases (NCEDs). Their expression correlates well with endogenous ABA levels (Schwartz *et al.*, 2003) and their overexpression results in a significant accumulation of ABA, which confers enhanced tolerance to multiple abiotic stresses such as dehydration and salt (Iuchi *et al.*, 2001; Qin and Zeevaart, 2002; Je *et al.*, 2014; Xian *et al.*, 2014).

ABA signalling also plays a critical role in stress response pathways by producing various, yet specific, outputs such as stomatal aperture (Gonzalez-Guzman *et al.*, 2012), and re-establishment of DT (Maia *et al*, 2014). Recently, one of the breakthroughs in ABA biology was the discovery of a core ABA signalling cascade consisting of the ABA receptors PYR/PYL/RCAR, the PP2C protein phosphatases as negative regulators, and SnRK2 protein kinases as positive regulators which, in turn, can activate transcription factors to regulate the expression of downstream targets (Park *et al.*, 2009).

Under normal conditions PP2C inactivates SnRK2 by dephosphorylating its multiple phosphorylation sites, but when endogenous ABA is upregulated due to abiotic stress such as dehydration (i.e. desiccation and drought), ABA binds to PYR/PYL/RCAR, this complex interacts with the PP2C and thus inhibits the protein phosphatase activity (Figure 1). In the absence of PP2C activity, SnRK2 auto-phosphorylates and, in doing so, activates itself. Moreover, in the regulatory network of natural acquisition of DT during seed development, ABA signalling genes such as *ABI3*, *ABA INSENSITIVE 4* (*ABI4*) and *ABI5* are strongly linked to DT in *Medicago truncatula* (Buitink *et al.*, 2003). *ABI3* is an orthologue of maize *VIVIPAROUS1 (VP1)* (McCarty *et al.*, 1991), and belongs to the B3-domain family of transcription factors (Swaminathan *et al.*, 2008). Together with two other B3 proteins, FUS3 and LEC2, ABI3 plays a key role in seed maturation, including the acquisition of dormancy and DT (Holdsworth *et al.*, 2008; Mentzen and Wurtele, 2008). A genome-wide analysis of the ABI3 regulon showed that 98 genes are targets of ABI3 (Mentzen and Wurtele, 2008). One of the ABI3 targets is ABI5, which acts downstream of ABI3 to arrest seedling establishment (Lopez-Molina *et al.*, 2002).

ABI4 belongs to the DREB/CBF subfamily of the AP2/ERF superfamily (Sakuma et al. 2002) and plays an important role in stress response (Wind *et al.*, 2013). ABI5 belongs to the ABF/AREB subfamily, and it plays a role in the ABA response during early stages of seedling establishment and its mutation result in ABA-insensitive germination and seedling growth (Finkelstein, 1994; Lopez-Molina *et al.*, 2002). Regarding the role of ABA signalling in the induction of DT, Maia *et al.* (2014) phenotyped several Arabidopsis mutants for re-establishment of DT in germinated Arabidopsis seeds. As mentioned, DT can be re-established in germinated Arabidopsis seeds upon application of a mild osmotic stress prior to the

appearance of the first root hairs. ABA signalling mutants, such as *abi3-8*, *abi3-9*, *abi4-3* and *abi5-7* which produced desiccation tolerant seeds, displayed a reduced ability to re-establish DT (Maia *et al,* 2014).

Additionally, *abi5* mutants of *M. truncatula* can no longer re-establish DT in germinated seeds (Terrasson *et al.*, 2013). Regulatory pathways leading to DT may act redundantly, as they respond to both seed developmental (Verdier *et al.*, 2013) and environmental cues during germination, such as osmotic stress (Nambara *et al.*, 2002; Buitink *et al.*, 2003; Maia *et al.*, 2011).

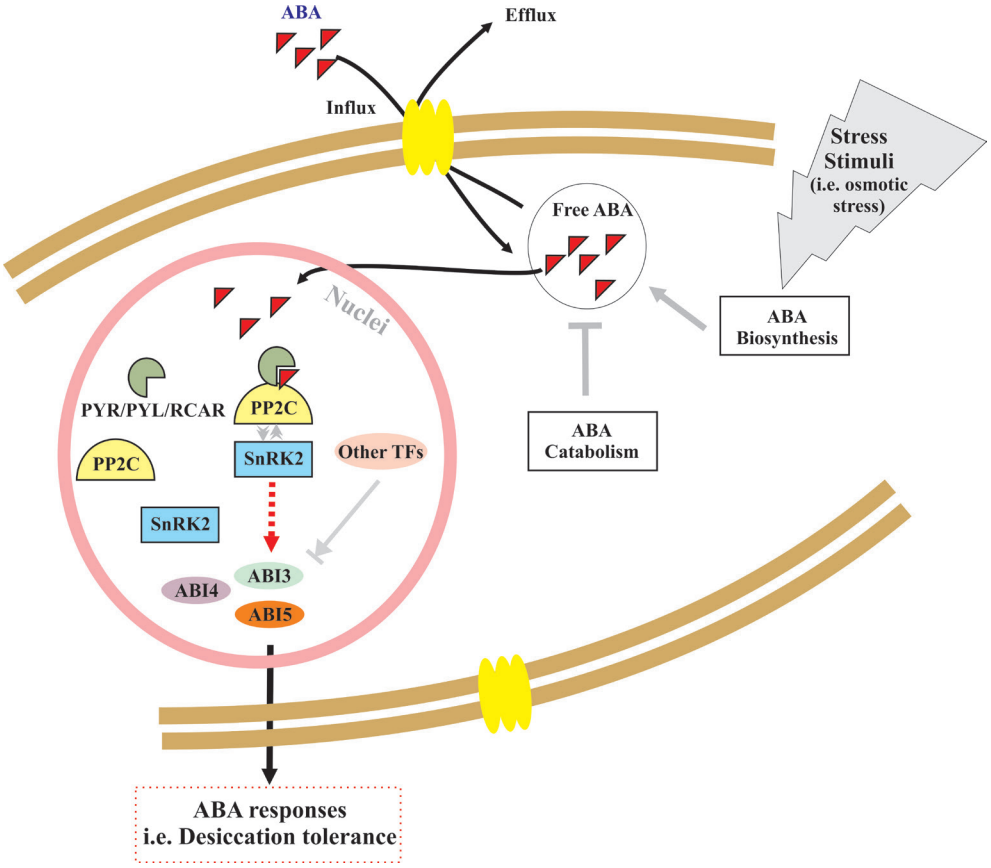


Figure1. Overview of ABA sensing, signalling and transport upon dehydration stress. PYR/PYL/RCAR, PP2C and SnRK2 in a core signalling complex. In the nucleus, the core complex directly regulates ABA-responsive gene expression by phosphorylation of SnRK2 (partially redraw from Umezawa *et al*., 2010).

Outline of this thesis

As described in this General Introduction, the transition from a quiescent metabolic state (dry seeds) to the active state of a vigorously growing seedling is crucial in the plant's life cycle and relatively little is known about the molecular basis underlying this transition. Therefore, the primary objective of the research described in this thesis was to detail the molecular basis underlying the seed-to-seedling phase transition. In **Chapter 2** I describe the global gene expression in seven successive developmental stages across the seed-to-seedling transition in *Arabidopsis thaliana*. With this data I identified several dominant expression patterns associated with transitions across developmental stages. One of these patterns suggests the existence of dedicated signal transduction pathways that regulate seedling establishment. In **Chapter 2** I show that one of the key regulators present in the latter pattern, the homeodomain leucine zipper I transcription factor *ATHB13*, affects root development during late seedling establishment. In **Chapter 3,** I analysed the primary metabolite profile in the same seven developmental stages. I detected two main metabolic shifts during the seed-to-seedling transition. Moreover, I correlated the transcriptome data from Chapter 2 with these metabolite data and this revealed a general framework of the contribution of metabolites and selected transcripts to the seed-to-seedling transition.

Having described the molecular basis of the seed-to-seedling phase transition in Chapters 2 and 3, in Chapters 4 and 5 I addressed the second objective of this study, to unveil molecular aspects of the re-establishment of desiccation tolerance (DT) in germinated seeds. Key here was to do this in such a way that it would mimic 'natural' drying in the soil better than when using osmotic solutions. To this end I developed a novel protocol for studying DT in germinated seeds called Mild Air Drying Treatment (MADT) in C**hapter 4**. With this new protocol I confirmed that an enhanced abscisic acid (ABA) accumulation is part of the DT response. I also found that ABA accumulation takes place beyond the DT window in germinated seeds. This new protocol is further explored in **Chapter 5** where I used the *ABA INSENSITIVE 3* mutant (*abi3-9*) and a microarray approach which enabled me to suggest the existence of crosstalk between ABA-dependent and ABA-independent transcription factors in the re-establishment of DT. Finally, in **Chapter 6** I discuss and integrate the results of my research and describe future perspectives.

References

Achard P, Cheng H, De Grauwe L, Decat J, Schoutteten H, Moritz T, Van Der Straeten D, Peng J, Harberd NP.2006 Integration of plant responses to environmentally activated phytohormonal signals. *Science* **311:** 91-94

Allen E, Moing A, Ebbels T, Maucourt M, Tomos AD, Rolin D, Hooks M.2010 Correlation Network Analysis reveals a sequential reorganization of metabolic and transcriptional states during germination and gene-metabolite relationships in developing seedlings of Arabidopsis. *BMC Systems Biology* **4:** 62

Amasino R.2010 Seasonal and developmental timing of flowering. *Plant J* **61:** 1001-1013

Bari R, Jones JDG.2009 Role of plant hormones in plant defence responses. *Plant Molecular Biology* **69:** 473-488

Bartels D, Sunkar R.2005 Drought and salt tolerance in plants. *Critical Reviews in Plant Sciences* **24:** 23-58

Bassel GW, Fung P, Chow T-fF, Foong JA, Provart NJ, Cutler SR.2008 Elucidating the Germination Transcriptional Program Using Small Molecules. *Plant Physiology* **147:** 143-155

Bassel GW, Lan H, Glaab E, Gibbs DJ, Gerjets T, Krasnogor N, Bonner AJ, Holdsworth MJ, Provart NJ.2011 Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. *Proceedings of the National Academy of Sciences* **108:** 9709-9714

Beligni MV, Lamattina L.2000 Nitric oxide stimulates seed germination and de-etiolation, and inhibits hypocotyl elongation, three light-inducible responses in plants. *Planta* **210:** 215-221

Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, Hashimoto M, Fei J, Harada CM, Munoz MD, Le BH, Drews GN, Brady SM, Goldberg RB, Harada JJ.2013 Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. *Proceedings of the National Academy of Sciences* **110:** E435-E444

Bewley JD.1997 Seed Germination and Dormancy. *The Plant Cell* **9:** 1055-1066

Bewley JD, Bradford K, Hilhorst HM, Nonogaki H (2013) Germination. *In* Seeds. Springer New York, pp 133-181

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

Carrera E, Holman T, Medhurst A, Peer W, Schmuths H, Footitt S, Theodoulou FL, Holdsworth MJ.2007 Gene Expression Profiling Reveals Defined Functions of the ATP-Binding Cassette Transporter COMATOSE Late in Phase II of Germination. *Plant Physiology* **143:** 1669-1679

Cook DR, Stewart SD, Howard JE, Akin DS, Gore J, Leonard BR, Lorenz GM, Davis JA.2014 Impact of simulated threecornered alfalfa hopper (Hemiptera: Membracidae)

induced plant loss on yield of maturity group IV and V soybeans. **49:** 176-189

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR.**2010** Abscisic Acid: Emergence of a Core Signaling Network. *Annual Review of Plant Biology* **61:** 651-679

Dekkers BJW, Pearce S, van Bolderen-Veldkamp RP, Marshall A, Widera P, Gilbert J, Drost H-G, Bassel GW, Müller K, King JR, Wood ATA, Grosse I, Quint M, Krasnogor N, Leubner-Metzger G, Holdsworth MJ, Bentsink L.**2013** Transcriptional Dynamics of Two Seed Compartments with Opposing Roles in Arabidopsis Seed Germination. *Plant Physiology*

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

Divi UK, 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

Dobrenel T, Marchive C, Sormani R, Moreau M, Mozzo M, Montané M-H, Menand B, Robaglia C, Meyer C (2011) Regulation of plant growth and metabolism by the TOR kinase, Vol 39

Eastmond PJ, Germain V, Lange PR, Bryce JH, Smith SM, Graham IA.**2000** Postgerminative growth and lipid catabolism in oilseeds lacking the gloxylate cycle. *Proceedings of the National Academy of Sciences of the United States of America* **97:** 5669-5674

Fait A, Angelovici R, Less H, Ohad I, Urbanczyk-Wochniak E, Fernie AR, Galili G.**2006** Arabidopsis Seed Development and Germination Is Associated with Temporally Distinct Metabolic Switches. *Plant Physiology* **142:** 839-854

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

Finkelstein RR, Gampala SSL, Rock CD.**2002** Abscisic Acid Signaling in Seeds and Seedlings. *The Plant Cell* **14:** S15-S45

Gonzalez-Guzman M, Pizzio GA, Antoni R, Vera-Sirera F, Merilo E, Bassel GW, Fernández MA, Holdsworth MJ, Perez-Amador MA, Kollist H, Rodriguez PL.**2012** Arabidopsis PYR/PYL/RCAR Receptors Play a Major Role in Quantitative Regulation of Stomatal Aperture and Transcriptional Response to Abscisic Acid. *The Plant Cell Online* **24:** 2483-2496

Hoekstra FA, Golovina EA, Buitink J.**2001** Mechanisms of plant desiccation tolerance. *Trends Plant Sci* **6:** 431-438

Holdsworth MJ, Bentsink L, Soppe WJJ.**2008** Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytologist* **179:** 33-54

Huijser P, Schmid M.**2011** The control of developmental phase transitions in plants. *Development* **138:** 4117-4129

Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K.2001 Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. *Plant J* **27:** 325-333

Je J, Chen H, Song C, Lim CO.2014 Arabidopsis DREB2C modulates ABA biosynthesis during germination. *Biochemical and Biophysical Research Communications* **452:** 91-98

Kang J, Turano FJ.2003 The putative glutamate receptor 1.1 (AtGLR1.1) functions as a regulator of carbon and nitrogen metabolism in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* **100:** 6872-6877

Kondo S, Sugaya S, Kittikorn M, Todoroki Y, Mizutani M, Hirai N (2014) Dehydration tolerance in apple seedlings is advanced by retarding ABA 8'-hydroxylase CYP707A. *In* Acta Horticulturae, Vol 1042, pp 151-157

Koornneef M, Bentsink L, Hilhorst H.2002 Seed dormancy and germination. *Curr Opin Plant Biol* **5:** 33-36

Layalle S, Arquier N, Léopold P.2008 The TOR Pathway Couples Nutrition and Developmental Timing in Drosophila. *Developmental Cell* **15:** 568-577

Lopez-Molina L, Mongrand S, Chua N-H.2001 A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factor in Arabidopsis. *Proceedings of the National Academy of Sciences* **98:** 4782-4787

Lopez-Molina L, Mongrand S, McLachlin DT, Chait BT, Chua NH.2002 ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *Plant J* **32:** 317-328

Lu C, Hills MJ.2002 Arabidopsis mutants deficient in diacylglycerol acyltransferase display increased sensitivity to abscisic acid, sugars, and osmotic stress during germination and seedling development. *Plant Physiology* **129:** 1352-1358

Maia J, Dekkers BJW, Dolle MJ, Ligterink W, Hilhorst HWM.2014 Abscisic acid (ABA) sensitivity regulates desiccation tolerance in germinated Arabidopsis seeds. *New Phytologist***:** n/a-n/a

Maia J, Dekkers BJW, Provart NJ, Ligterink W, Hilhorst HWM.2011 The Re-Establishment of Desiccation Tolerance in Germinated *Arabidopsis thaliana* Seeds and Its Associated Transcriptome. *PLos ONE* **6**

Mansfield SG, Briarty LG.1996 The Dynamics of Seedling and Cotyledon Cell Development in Arabidopsis thaliana During Reserve Mobilization. *International Journal of Plant Sciences* **157:** 280-295

McCarty DR, Hattori T, Carson CB, Vasil V, Lazar M, Vasil IK.1991 The Viviparous-1 developmental gene of maize encodes a novel transcriptional activator. *Cell* **66:** 895-905

Mentzen W, Wurtele E.2008 Regulon organization of Arabidopsis. *BMC Plant Biology* **8:** 99

Miransari M, Smith DL.2014 Plant hormones and seed germination. *Environmental and Experimental Botany* **99:** 110-121

Nambara E, Marion-Poll A (2005) Abscisic acid biosynthesis and catabolism. *In* Annual Review of Plant Biology, Vol 56, pp 165-185

Nambara E, Suzuki M, Abrams S, McCarty DR, Kamiya Y, McCourt P.**2002** A Screen for Genes That Function in Abscisic Acid Signaling in Arabidopsis thaliana. *Genetics* **161:** 1247-1255

Nonogaki H, Bassel GW, Bewley JD.**2010** Germination—Still a mystery. *Plant Science* **179:** 574-581

Okamoto M, Kuwahara A, Seo M, Kushiro T, Asami T, Hirai N, Kamiya Y, Koshiba T, Nambara E.**2006** CYP707A1 and CYP707A2, which encode abscisic acid 8'-hydroxylases, are indispensable for proper control of seed dormancy and germination in Arabidopsis. *Plant Physiol* **141:** 97-107

Park SY, Fung P, Nishimura N, Jensen DR, Fujii H, Zhao Y, Lumba S, Santiago J, Rodrigues A, Chow TFF, Alfred SE, Bonetta D, Finkelstein R, Provart NJ, Desveaux D, Rodriguez PL, McCourt P, Zhu JK, Schroeder JI, Volkman BF, Cutler SR.**2009** Abscisic Acid Inhibits Type 2C Protein Phosphatases via the PYR/PYL Family of START Proteins. *Science* **324:** 1068-1071

Penfield S, Graham S, Graham IA.**2005** Storage reserve mobilization in germinating oilseeds: Arabidopsis as a model system. *Biochem Soc Trans* **33:** 380-383

Perruc E, Kinoshita N, Lopez-Molina L.**2007** The role of chromatin-remodeling factor PKL in balancing osmotic stress responses during Arabidopsis seed germination. *Plant J* **52:** 927-936

Qin X, Zeevaart JAD.**2002** Overexpression of a 9-cis-Epoxycarotenoid Dioxygenase Gene in Nicotiana plumbaginifolia Increases Abscisic Acid and Phaseic Acid Levels and Enhances Drought Tolerance. *Plant Physiology* **128:** 544-551

Rademacher W (2000) Growth retardants: Effects on gibberellin biosynthesis and other metabolic pathways. *In* Annual Review of Plant Biology, Vol 51, pp 501-531

Rougvie AE.**2005** Intrinsic and extrinsic regulators of developmental timing: from miRNAs to nutritional cues. *Development* **132:** 3787-3798

Russell RC, Fang C, Guan K-L.**2011** An emerging role for TOR signaling in mammalian tissue and stem cell physiology. *Development* **138:** 3343-3356

Saito S, Hirai N, Matsumoto C, Ohigashi H, Ohta D, Sakata K, Mizutani M.**2004** Arabidopsis CYP707As encode (+)-abscisic acid 8'-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. *Plant Physiol* **134:** 1439-1449

Sato T, Maekawa S, Yasuda S, Sonoda Y, Katoh E, Ichikawa T, Nakazawa M, Seki M, Shinozaki K, Matsui M, Goto DB, Ikeda A, Yamaguchi J.**2009** CNI1/ATL31, a RING-type ubiquitin ligase that functions in the carbon/nitrogen response for growth phase transition in Arabidopsis seedlings. *The Plant Journal* **60:** 852-864

Schwartz SH, Qin X, Zeevaart JAD.**2003** Elucidation of the Indirect Pathway of Abscisic Acid Biosynthesis by Mutants, Genes, and Enzymes. *Plant Physiology* **131:** 1591-1601

Shimada H, Mochizuki M, Ogura K, Froehlich JE, Osteryoung KW, Shirano Y, Shibata D, Masuda S, Mori K, Takamiya K-i.**2007** Arabidopsis Cotyledon-Specific

Chloroplast Biogenesis Factor CYO1 Is a Protein Disulfide Isomerase. *The Plant Cell* **19:** 3157-3169

Swaminathan K, Peterson K, Jack T.**2008** The plant B3 superfamily. *Trends Plant Sci* **13:** 647-655

Terrasson E, Buitink J, Righetti K, Ly Vu B, Pelletier S, Zinsmeister J, Lalanne D, Leprince O.**2013** An emerging picture of the seed desiccome: confirmed regulators and newcomers identified using transcriptome comparison. *Front Plant Sci* **4:** 497

Tripathi P, Rabara RC, Rushton PJ.**2014** A systems biology perspective on the role of WRKY transcription factors in drought responses in plants. *Planta* **239:** 255-266

Ulmasov T, Murfett J, Hagen G, Guilfoyle TJ.**1997** Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **9:** 1963-1971

Umehara M, Hanada A, Yoshida S, Akiyama K, Arite T, Takeda-Kamiya N, Magome H, Kamiya Y, Shirasu K, Yoneyama K, Kyozuka J, Yamaguchi S.**2008** Inhibition of shoot branching by new terpenoid plant hormones. *Nature* **455:** 195-200

Umezawa T, Nakashima K, Miyakawa T, Kuromori T, Tanokura M, Shinozaki K, Yamaguchi-Shinozaki K.**2010** Molecular Basis of the Core Regulatory Network in ABA Responses: Sensing, Signaling and Transport. *Plant and Cell Physiology* **51:** 1821-1839

Umezawa T, Okamoto M, Kushiro T, Nambara E, Oono Y, Seki M, Kobayashi M, Koshiba T, Kamiya Y, Shinozaki K.**2006** CYP707A3, a major ABA 8'-hydroxylase involved in dehydration and rehydration response in Arabidopsis thaliana. *Plant J* **46:** 171-182

Verdier J, Lalanne D, Pelletier S, Torres-Jerez I, Righetti K, Bandyopadhyay K, Leprince O, Chatelain E, Vu BL, Gouzy J, Gamas P, Udvardi MK, 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

Westwood JH, McCann L, Naish M, Dixon H, Murphy AM, Stancombe MA, Bennett MH, Powell G, Webb AA, Carr JP.**2013** A viral RNA silencing suppressor interferes with abscisic acid-mediated signalling and induces drought tolerance in Arabidopsis thaliana. *Mol Plant Pathol* **14:** 158-170

Wind JJ, Peviani A, Snel B, Hanson J, Smeekens SC.**2013** ABI4: versatile activator and repressor. *Trends Plant Sci* **18:** 125-132

Wullschleger S, Loewith R, Hall MN.**2006** TOR signaling in growth and metabolism. *Cell* **124:** 471-484

Xian L, Sun P, Hu S, Wu J, Liu JH.**2014** Molecular cloning and characterization of CrNCED1, a gene encoding 9-cis-epoxycarotenoid dioxygenase in Citrus reshni, with functions in tolerance to multiple abiotic stresses. *Planta* **239:** 61-77

Xiong L, Zhu J-K.**2003** Regulation of Abscisic Acid Biosynthesis. *Plant Physiology* **133:** 29-36

Xiong Y, McCormack M, Li L, Hall Q, Xiang C, Sheen J.2013 Glucose-TOR signalling reprograms the transcriptome and activates meristems. **496:** 181-186

Zhang F, Huang L, Wang W, Zhao X, Zhu L, Fu B, Li Z.2012 Genome-wide gene expression profiling of introgressed indica rice alleles associated with seedling cold tolerance improvement in a japonica rice background. 461

Zhao L, Hu Y, Chong K, Wang T.2010 ARAG1, an ABA-responsive DREB gene, plays a role in seed germination and drought tolerance of rice. *Annals of botany* **105:** 401-409

Chapter 2

A predictive co-expression network for the seed-to-seedling transition in *Arabidopsis thaliana*

Anderson Tadeu Silva, Wilco Ligterink and Henk W. M. Hilhorst
(*Submitted*)

Summary

The transition from a quiescent dry seed to an actively growing photoautotrophic seedling is a complex and crucial trait for plant propagation. The seed-to-seedling transition has been studied intensively but only few genetic factors regulating this phase transition have been identified. This study provides a detailed description of the dynamics of global gene expression in seven successive developmental stages of seedling establishment in Arabidopsis. The results offer the most comprehensive description of gene expression during the seed-to-seedling transition to date. The co-expression gene network highlights interactions between known regulators of this transition and predicts the function of uncharacterized genes in seedling establishment. Analysis of co-expressed gene data sets for the phase transition from germinated seed to established seedling suggests dedicated signal transduction pathways that regulate seedling establishment. One of the identified key regulators, the homeodomain leucine zipper I transcription factor *ATHB13,* is expressed during germination, but affects late seedling establishment. The seedling phenotype of the *athb13* mutant showed increased primary root length as compared with the wild type (Col-0), suggesting that this transcription factor may controls cell division during early seedling formation. We conclude that signal transduction pathways present during the early phases of the seed-to-seedling transition anticipate on controlling root growth in later stages of seedling establishment. This study demonstrates that a gene co-expression network together with its transcriptional modules can provide mechanistic insights that are not likely derived from comparative transcript profiling alone.

Introduction

The transition from seed to seedling is mediated by germination, which is a complex process that starts with imbibition and completes with radicle emergence. Seed germination is a crucial process in seedling establishment as it marks a functional *point-of-no-return*. Despite the profound impact of seedling performance on crop establishment and yield, relatively little is known about the molecular processes underlying the transition from seed to seedling, or from hetero- to autotrophic growth. This transition is decisive for plants to enter a natural or agricultural ecosystem and is an important basis for crop production.

Once germination has started, mobilization of stored reserves is essential to provide the growing seedling with energy and building blocks before it becomes (photo)autotrophic. The importance of energy metabolism to support germination

and seedling growth is evident from primary metabolite profiling of early germination (Fait *et al.*, 2006) and from studies that show inhibited seedling growth in mutants defective in seed lipid mobilization (Fulda *et al.*, 2004). Moreover, evidence from gene expression profiling studies in *Arabidopis thaliana* suggests that the transcriptome required for seed germination and seedling growth is already present in the mature dry seed that has just completed development and maturation (Cadman *et al.*, 2006; Finch-Savage *et al.*, 2007).

Application of α-amanitin, an inhibitor of transcription that targets RNA polymerase II, appears to allow completion of the germination process until radicle protrusion but inhibits subsequent seedling growth while inhibitors of translation prevent progress of germination from the start (Rajjou *et al.*, 2004). This suggests that transcriptional changes during the germination process are required to accommodate post-germination growth. Thus, it appears that in the successive developmental stages between seed maturation and seedling growth the transcriptome is one developmental step 'ahead' of the proteome and the physiology. This conclusion was corroborated by the observation that light, perceived by phytochrome B in the seed, generated a downstream trans-developmental phase signal (mediated by the *ABI3* gene) which, apparently, preconditions seedlings to their most likely environment (Mazzella *et al.*, 2005). Seedling emergence, therefore, depends, at least partly, on inherent seed characteristics.

Although germination has been studied for many years, a significant advancement of knowledge of the complex germination process was not attained until sequence information and –omics technologies became widely available. In Arabidopsis, a number of studies utilizing, sometimes high-resolution, transcriptomic approaches to investigate the time course of seed germination has made major contributions (Holdsworth *et al.*, 2008; Holdsworth *et al.*, 2008; Narsai *et al.*, 2011; Dekkers *et al.*, 2013). However, there is a general lack of similar studies following the completion of seed germination, *viz.* the beginning of radicle protrusion and subsequent seedling establishment. Similar studies of the transcriptome during this phase of growth may, therefore, not only provide a global view of gene expression patterns, including biological function enrichment, but also a predictive dimension once co-expressed gene sets have been identified. For example, the most comprehensive transcriptional study, thus far, of the time course of seed development, reported predictions of gene regulatory networks, identifying regulators of seed development (Belmonte *et al.*, 2013). In this context, seedling growth stages until appearance of the first root hairs, or beyond, have not been studied in meaningful detail. This implies that potentially regulatory changes in the transcriptome have not yet been associated with seedling establishment.

The main objective of the present study was to identify regulatory factors to reveal signal transduction routes that are involved in the seed-to-seedling transition. Studies using large transcriptome data sets have demonstrated correlation of gene expression (Usadel *et al.*, 2009; Bassel *et al.*, 2011; Dekkers *et al.*, 2013; Verdier *et al.*, 2013). Co-expressed genes have a greater likelihood of being involved in a common biological condition or developmental process (Aoki *et al.*, 2007; Usadel *et al.*, 2009; Bassel *et al.*, 2011; Verdier *et al.*, 2013). In addition, transcriptional modules, such as those identified for seed development (Belmonte *et al.*, 2013), are likely to identify regulatory circuits of a process by association of overrepresented DNA sequence motifs with co-expressed transcription factors (Belostotsky *et al.*, 2009). Here we show that the assessment of global transcript changes across developmental stages from the mature dry seed to the seedling stage of fully opened cotyledons provides a comprehensive view of biological processes involved in seedling development and establishment. Analysis of these developmental stages enabled us to identify informative gene sets, such as stage peak-transcripts and dominant expression patterns (DPs). The identification of these robust expression patterns will provide an essential resource to better understand the seed-to-seedling transition.

Experimental procedures

Plant material collection

Seeds of *Arabidopsis thaliana*, accession Columbia (Col-0 [N60000]) were cold stratified at 4 ºC in the dark for 72 h in Petri dishes using two layers of moistened blue filter paper (Anchor paper Co, Saint Paul, Mn, USA.) to break residual dormancy. Germination tests were performed in a growth chamber at 22 ºC under constant white light. To elucidate the changes in the transcriptome related to the transition from a seed to a photo-autotrophic seedling, seven developmental stages were identified: (DS) mature dry seed; (6H) seeds, imbibed for 6h and germination-related initiation of metabolic activity; (TR), embryo swelling and testa rupture; (RP) protrusion of the radicle through the endosperm, primarily through cell elongation, followed by further embryonic root extension and beginning of (RH) root hair formation, succeed by the appearance of (GC) greening cotyledons and (OC) cotyledons that are fully opened (Figure 1).

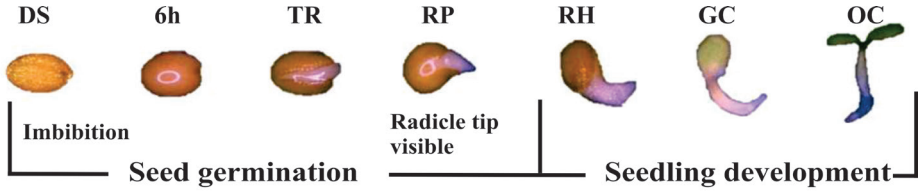


Figure 1. Sub-division of the seed-to-seedling developmental stages.

RNA extraction

Total RNA was extracted using the hot borate method according to Wan and Wilkins (1994) with some modifications as described previously (Maia *et al.*, 2011). RNA quality and concentration were measured by agarose gel electrophoresis (0.1 $g.mL^{-1}$) and NanoDrop®.

Microarray hybridization

The quality control, RNA labeling, hybridization and data extraction were performed at ServiceXS B.V. (Leiden, The Netherlands). Labelled ss-cDNA was synthesized using the Affymetrix NuGEN Ovation PicoSL WTA v2 kit and Biotin Module using 50 ng of total RNA as template. The fragmented ss-cDNA was utilized for hybridization on an Affymetrix ARAGene 1.1ST array plate. The Affymetrix HWS Kit was used for the hybridization, washing and staining of the plate. Scanning of the array plates was performed using the Affymetrix GeneTitan scanner. All procedures were performed according to the instructions of the manufacturers (nugen.com and affymetrix.com). The resulting data were analyzed using the R statistical programming environment and the Bioconductor packages (Gentleman *et al.*, 2004). The data was normalized using the RMA algorithm (Irizarry *et al.*, 2003) with the TAIRG v17 cdf file (http://brainarray.mbni.med.umich.edu). Expression data are hosted in the NCBI GEO database (GSE65394). *(http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=onyxsyycjtaxxux&acc=GSE65394). Validation of the seed-to-seedling transcriptome data set was performed by comparison with previously published expression patterns of genes known to be differential expressed across seed to seedling developmental stages (Supplemental Table S1) (Ding *et al.*, 2006; Tepperman *et al.*, 2006; Hammani *et al.*, 2011; Narsai *et al.*, 2011; Feng *et al.*, 2014).

Identification of co-expression gene sets and transcriptional modules

Stage-specific gene sets. For selection of differentially expressed transcripts specific for a given developmental stage, a two-step approach was used. First, Bayesian Estimation of Temporal Regulation - BETR ($p < 0.001$) was used to identify differentially expressed transcripts in at least two different developmental stages (Aryee *et al.*, 2009). At the next step differentially expressed transcripts were further filtered to determine whether they were specific for a particular stage. The Limma package (Gentleman *et al.*, 2005) was used to check whether the average expression values in a specific stage were significantly ($p < 0.01$) larger than expression at other stages. The extracted datasets were hierarchically clustered and visualized in GeneMaths XT®.

Dominant expression pattern and transcriptional module prediction. Dominant expression patterns (DPs) were identified as previously described by Belostotsky *et al.* (2009). DPs were identified of the 50% most variant transcripts, corresponding to 9.565 mRNAs. The R function FANNY (http://cran.r-project.org/web/packages/cluster/cluster.pdf) with a minimum Pearson correlation of 0.85 was used to evaluate the number of clusters (K) choices from 1 to 50 with a cut-off for cluster membership of 0.4. The K choice that yielded the greatest number of transcriptional modules was ten. These transcriptional modules were used in the ChipEnrich software package developed by Brady et al. (2007) and modified by Belmonte et al. (2013). ChipEnrich determines the significance of GO terms, metabolic processes, DNA motifs and transcription factors (TFs) using *p* values calculated from their hypergeometric distribution (Belostotsky *et al.*, 2009; Belmonte *et al.*, 2013). For the hypergeometric distribution lists of GO terms, metabolic processes, DNA motifs and TFs were used based on Arabidopsis Gene Regulatory Information Server – AGRIS (http://arabidopsis.med.ohio-state.edu/AtTFDB/). An optimized ChipEnrich by Belmonte *et al.* (2013) was used to identify significantly enriched DNA motifs, associated TFs and GO terms. Tables generated by ChipEnrich were imported into Cytoscape (version 2.8.2) and the transcriptional module networks were visualized using the yFiles Organic layout.

Seed germination, seedling establishment and root growth phenotype

ATHB mutant lines were obtained from Cabello *et al.* (2012) (*athb13-1*) and Barrero *et al.* (2010) (*athb20-1*). Seeds of these mutants were sown on 5x5 cm Rockwool® blocks in a climate cell (20ºC day; 18ºC night) with a photoperiod of 16h light and 8h dark. Each Rockwool® block was watered with Hyponex® solution (1 g/L). Seeds

were harvested in four replicates of at least three plants. In order to measure root growth, seeds were sterilised with commercial bleach (20% v/v) and placed on solid medium containing 0.5x Murashige and Skoog (MS) (Murashige and Skoog, 1962) without sucrose. Thereafter, seeds were stratified for 72h at 4ºC to remove residual dormancy and transferred to a germination cabinet at 22ºC with constant white light. Twenty-seven seeds were selected at the RP stage for Col-0 and for each of the mutants (*athb13-1* and *athb20-1*). These seeds were placed on new plates with 0.5x solid MS medium without sucrose. Plates were placed vertically in a climate cell in the same conditions as described above. Root growth was scored 15 days later for each plant. Plates were scanned using an Epson document scanner and root lengths were determined by SmartRoot® 4.1 using ImageJ® software.

Results

Transcriptomic changes during the seed-to-seedling transition

Transcript abundance of *Arabidopsis thaliana* (Col-0) in each developmental stage of the seed-to-seedling transition was analysed, using a high-density *Affymetrix*® array (Aragene.st1.1). This array encompasses the complete Arabidopsis transcriptome. Principal component analysis (PCA) was used to compare the overall variation in gene expression levels among the seven developmental stages, using the entire transcriptome with three biological replicates for each developmental stage (Figure 2). Each developmental stage was clearly distinct from the other stages. Proximity of the replicates in the PCA plot highlights the robustness of the experimental set-up and data processing steps and shows that this is a powerful data set to study the seed-to-seedling transition.

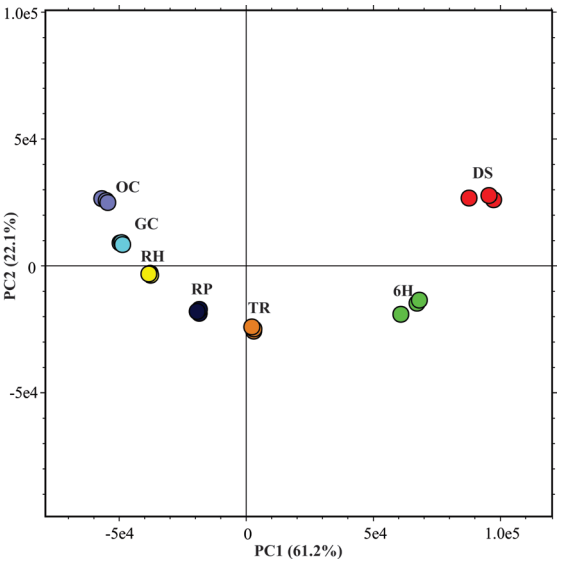


Figure 2. PCA plot of transcript abundance of the seven stages of seed. germination and seedling. development.

Of the complete transcriptome, 19.130 (69%)

transcripts were differentially expressed in at least two developmental stages during the seed-to-seedling transition. Pearson correlation was applied to these transcripts and this indicated distinct temporal profiles (Figure S1 and Dataset S1). When successive developmental stages were compared, considerable changes could be observed. The greatest change in number of transcripts was observed in the comparison between DS and 6H (over 7.000 genes up-regulated and 4.000 down-regulated). Other sets of transcripts showed more moderate changes: from 6H to TR (5600 genes up and 4768 down), TR to RP (840 genes up and 1.107 down), RP to RH (1596 genes up and 1824 down), RH to GC (1543 genes up and 1407 down), and between GC and OC (2515 up and 3272 down). (Figure S2). Highly expressed genes in DS were enriched for GO terms related to *heat response* and *lipid storage*, whereas high expression at 6H was related to *nucleotide binding* and *structural constituent of ribosome*, suggesting activation of translational activity. Gene sets of the other comparisons were associated with such processes as cell cycle, protein synthesis, DNA processing and transcription (Supplemental Table S2). For example, GO terms such as *RNA processing, nucleic acid binding* were enriched for genes highly expressed in the TR, RP and RH stages, whereas *chloroplast envelope*, *endomembrane system* and *ribosome biogenesis* were enriched for genes highly expressed in GC and OC stages.

Transcriptomic analysis identifies sets of developmentally regulated processes during the seed-to-seedling transition

Ramification of gene clusters suggested the temporal expression of developmentally regulated transcripts. The results show transcripts that specifically peaked at each developmental stage and ten dominant patterns (DPs) across all seed-to-seedling developmental stages.

Stage-peaking gene sets. The complexity of the data sets suggests a coordinated shift in gene expression at the developmental stages of the transition. Because of this complexity, we identified genes that peaked ($p < 0.01$, Bonferroni adjusted) at a particular stage, derived from the subset of 19.130 transcripts. This analysis illustrates that the different sets of genes display peaks of expression at different developmental stages, which is suggestive of their relevance for stage-specific developmental functions (Supplemental Table S3). Interestingly, the clusters of developmentally regulated transcripts grouped into specific stages and formed a "wave" of transcript abundance, moving from a quiescent dry seed to a growing seedling (Figure 3).

These clusters may thus govern the progression of the genetic program towards seedling establishment. Analysis of the peaking genes resulted in 6.384 transcripts that showed significant levels of differential expression with a single peak

across the seed-to-seedling development stages. Of 6.384 transcripts, 50% showed a maximum transcript expression in mature dry seeds and 24% in seeds imbibed for six hours, whereas in TR, RP and RH less than 2% displayed maximum expression (0.6% at testa rupture, 0.3% at radicle protrusion and 0.5% at the appearance of the first root hair). GC and OC displayed maximum expression of around 22% and 2%, respectively (Figure 3 and Supplemental Table S3).

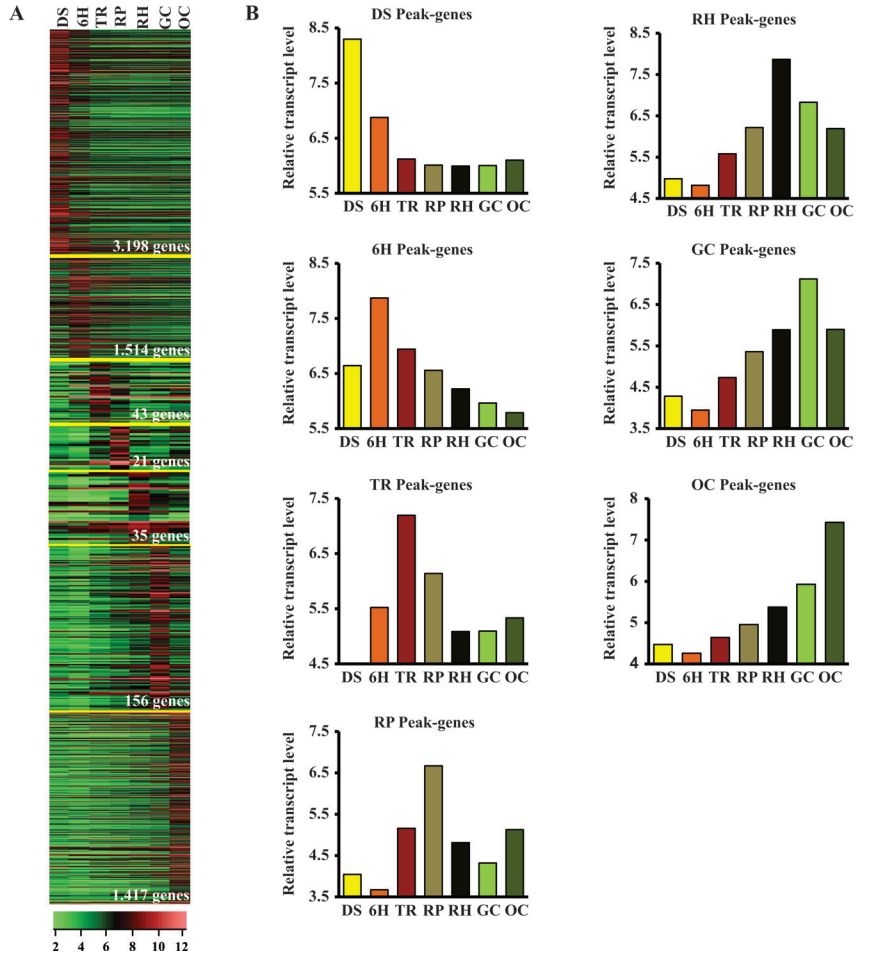


Figure 3. Overview of the expression patterns of transcripts peaking at selected seed-to-seedling transition stages. **A.** Differentially expressed gene clustering based on the peak of expression at the selected seed-to-seedling stage. **B.** General tendency and average abundance of transcripts at each stage. This analysis shows the maximum expression of transcripts at different stages, indicating the stage-specific maximum activity of the transitionally regulated transcripts.

The number of peaking transcripts for each developmental stage indicated that transcript abundance can be grouped in three distinct clusters: (*i*) DS and 6H; (*ii*) TR, RP and RH, and (*iii*) GC and OC, implying two major transitions. This complex pattern of gene activity observed during the seed-to-seedling transition can help to determine the fundamental molecular processes involved in seedling establishment and, hence, to predict seed and seedling quality by monitoring gene expression during seed germination and seedling establishment.

Dominant patterns. To determine how transcript abundance changes during the seed-to-seedling transition, we also clustered transcripts from the subset of 19.130 transcripts into ten dominant expression patterns (DPs) (Figure 4 and Supplemental Table S3), using the *Fuzzy K-Means* clustering method (Belostotsky *et al.*, 2009). Five of the co-expressed gene sets consisted of transcripts with high expression at only one stage (DP3, DP4, DP5, DP8, and DP9), whereas the other five co-expressed gene sets were expressed across several developmental stages. These expression patterns suggest the occurrence of processes related to specific stages of the seed-to-seedling transition.

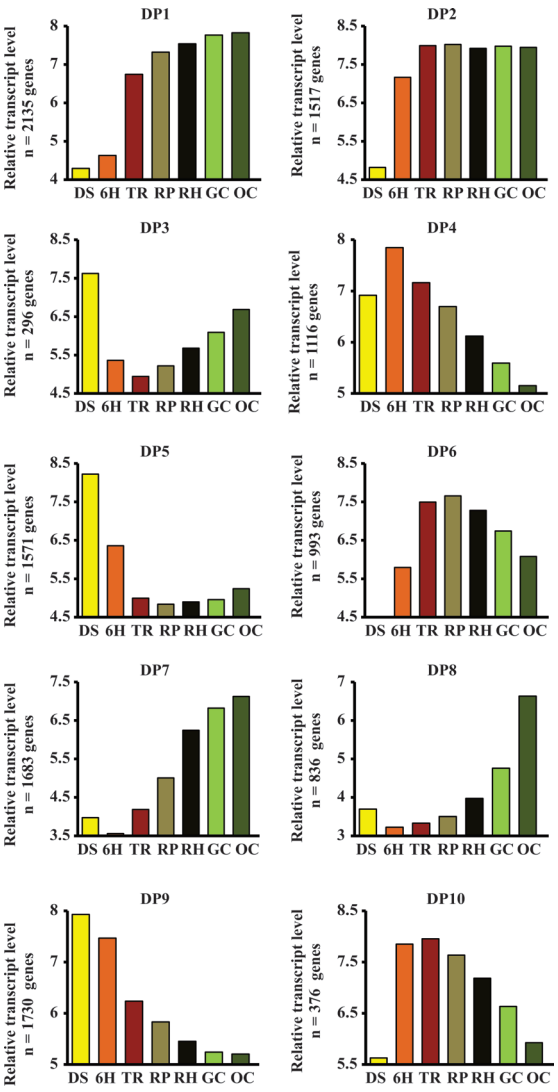


Figure 4. Dominant patterns of gene expression during the seed-to-seedling transition. Ten DPs were found using *Fuzzy K-means* clustering of the 50% most variant transcripts from the data set of 19.130 transcripts that showed significant expression difference in at least one developmental stage. Bar graphs represent averages of mRNA expression levels in each stage (left to right, mature seeds to cotyledons fully opened).

Predictive functions of these processes were determined for each of the DPs by analysis of enriched GO terms ($p < 0.0001$) and metabolic processes ($p<0.001$) (Figure 5 and Supplemental Table S3). For example, the DP1 gene set was overrepresented for genes with GO terms such as *fatty acid activity*, *chlorophyll biosynthesis* and *photosynthesis* (Figure 5). Another gene set, DP7, was significantly enriched for GO terms and metabolic processes associated with photosynthesis and carbon metabolism, including *chloroplast* and *thylakoid structure*, *photorespiration*, *calvin cycle* and *gluconeogenesis* (Figure 5 and Supplemental Table S3). DP6 and DP10 consisted of transcripts highly expressed in the TR, RP and RH stages. DP6 and DP10 were significantly enriched for *DNA unwinding* and *ribosome related transcripts* (Figure 5).

Figure 5. Heat map showing the *p* value significance of GO enrichment terms for DPs (Supplemental Table S3). GO terms listed are for biological process or/and cellular components that were overrepresented for each DP gene set.

The dominant expression patterns displayed a transient variation of transcript abundance across the seed-to-seedling developmental stages, which suggests a complex network of transcripts that control the major transitions in seedling development.

In summary, the predicted pathways associated with the peaking-genes and DP data sets can be described by GO terms and metabolic processes (Supplemental Table S3). DS peak-genes were significantly enriched for GO terms known to be associated with events that occur during seed development, maturation and abscission including *ubiquitin-protein ligase activity* and *response to heat*. 6H peak-genes were overrepresented for processes associated with high levels of transcriptional activity, such as *RNA binding* and *nucleic acid binding*.

Gene sets of TR, RP and RH were not significantly enriched for GO terms and metabolic processes because only few transcripts peaked at these stages. GC-specific transcripts, however, were highly enriched for microtubule motor activity genes (*mitotic kinesins*) that are associated with the cell cycle. Transcripts peaking at OC were enriched for processes supporting tissue growth, such as *transport* and *lipid binding* (Figure S3). Moreover, dominant pattern gene sets associated with transitory regions (DP1, DP2, DP6, DP7 and DP10) were significantly enriched for *cell differentiation*, *proliferation*, and *photosynthesis*. DP2, DP6 and DP10 probably undergo cell divisions with subsequent cellularization, but no photosynthesis. *Photosynthesis* and *chlorophyll binding*, however, were enriched in DP1 and DP7. Gene expression related to photosynthesis might start already at the TR stage (DP1) with a second wave of transcript abundance starting at RP (DP7) (Figure 4).

A co-expression gene regulatory network of the seed-to-seedling transition

To identify regulatory processes that are controlling the seed-to-seedling transition an unweighted gene co-expression network analysis was carried out. Edge adjacency threshold was set at 0.98, resulting in a network consisting of 6.896 nodes with 99.762 edges. Network visualization was carried out in Cytoscape using the organic layout (Figure 6). Temporal expression profiles corresponding to the different co-expression gene sets, were clearly separated. These distinct regions are associated with the three main stages that dominate the transition: mature dry seeds (DS), early imbibition (6H) and the final stage of seedling establishment (OC). As we were interested in the mechanisms that govern the transition from a quiescent to a photo-autotrophic state, we investigated which dominant patterns were most representative for this transition. For this purpose, each DP previously identified was mapped on the seed-to-seedling network (Figure 7).

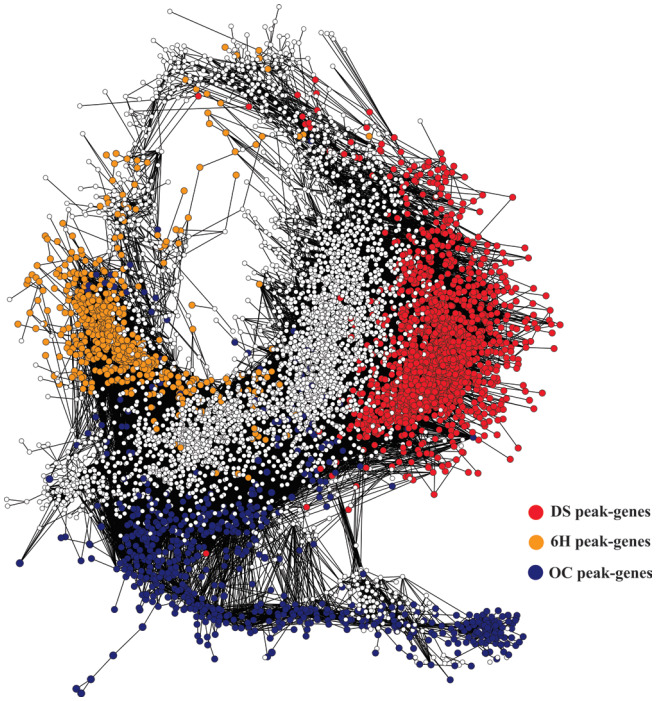


Figure 6. Unweighted gene co-expression network for the seeds-to-seedling transition. Circles (nodes) represent transcripts and lines (edges) represent significant transcriptional interactions between the transcripts. Temporal analysis of network clusters was performed by colouring each gene by its maximum expression across seed-to-seedling transition developmental stages (peaking-gene sets).

Regions of transcriptional interaction associated with phase transitions were observed for the DP1, DP2, DP6, DP7 and DP10 clusters. DP2 and DP10 showed gene expression patterns increasing from DS to 6H, whereas DP1 and DP6 showed an increase from 6H to TR (Figure 4). This suggests that these four DPs are associated with phase transitions in seed germination rather than early seedling growth.Therefore, of these five DPs (DP1, DP2, DP6, DP7 and DP10) DP7 is predicted to bridge the gap between germination and the seedling stage (Figure 7). Therefore, DP7 is the most representative gene set associated with the transition from germination (6H) to early seedling establishment (OC). The DP7 set is unique in that the transcript abundance increased after RP and kept increasing to higher levels thereafter (Figure 4). The expression pattern of DP7 suggests that a common set of transcripts is uniformly up-regulated across the seedling developmental stages. It may also indicate that a distinct regulatory process during RP affects the abundance of over 1.600 transcripts which subsequently control seedling establishment.

Predicted regulatory circuitry controlling transcripts expressed during the transition from germinated seed to established seedling

To define a gene regulatory network that might control the transition from a quiescent to an autotrophic photosynthesizing state, we inferred a predicted transcriptional module that links transcription factors (TFs) with their potential co-expressed target transcripts, using ChipEnrich and Cytoscape. Transcriptional modules of stage-specific transcripts (DS and OC) link enriched ABRE, G-box, ABFs, and ATHB5 DNA sequence motifs with known or predicted TFs that are present in mature seeds and seedlings (Figure 8). The transcriptional module for DS is enriched for G-box, ABRE and ABF DNA motifs, which are associated with bZIP TFs known to function in seed maturation (Gutierrez *et al.*, 2007).

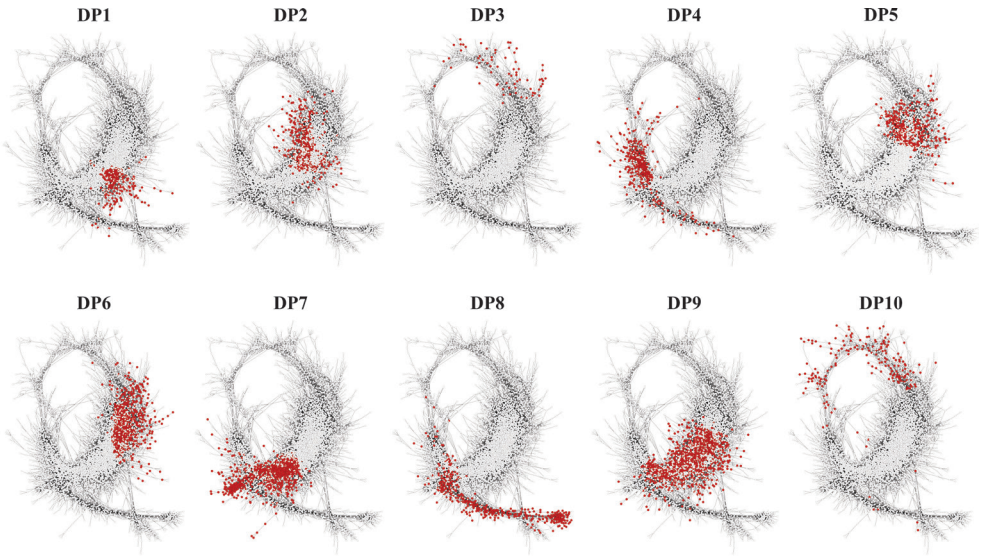


Figure 7. Seed-to-seedling network with nodes (transcripts) coloured according to the DP to which they belong, showing expression regions on the network for the different DPs (Figure 4).

The transcriptional module for OC showed enrichment for the ATHB5 DNA binding motif known to be associated with seedling development (Johannesson *et al.*, 2003). Transcriptional modules built from the other seed-to-seedling transition gene sets may give additional input to identify potential regulatory circuits that govern phase transitions within the network. We also built transcriptional modules for each DP (Figure S4 and S5). These showed an evident overlap in enriched DNA motifs and TFs identified in the stage-specific gene sets. This observation suggests

that seedling establishment processes are regulated in a dynamic manner as shown by the DPs. The transcriptional module for the co-expressed gene sets identified as being in the transitory region from germinated seed to the photo-autotrophic seedling (DP7) was predicted to be enriched for the ATHB5 DNA binding motif associated with the three homeodomain-leucine zipper TFs *ATHB13*; *ATHB20* and *ATHB23*. These TFs are members of the subclass-α (*ATHB3*, *-13*, *-20* and *-23*) within class I homeodomain-leucine zipper TFs, with a putative role in leaf development (Henriksson *et al.*, 2005). However, none of these have been previously related to a seedling-related developmental phase transition.

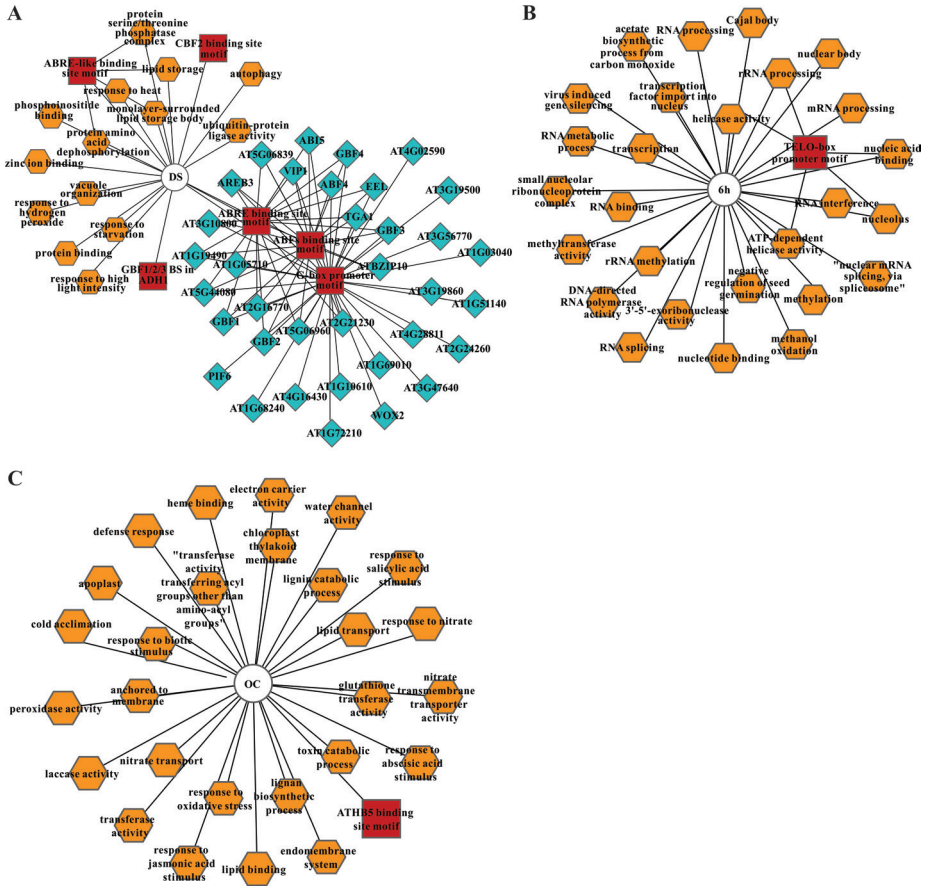


Figure 8. Transcriptional modules predicted to regulate the seed-to-seedling transition. DNA motifs (red rectangles) and GO terms (orange octagons) that are significantly overrepresented (p <0.001) within the peak–transcript sets (open ellipses) together with co-expressed transcription factors (light-blue diamonds). Transcriptional modules were predicted for DS (**A**), 6H (**B**) and OC (**C**) peak-transcripts (Supplemental Table S3).

Homeodomain-leucine zipper *ATHB13* regulates primary root growth

To investigate whether *ATHB* transcription factors are present in the transitory region, we mapped these transcription factors on the seed-to-seedling network (Figure 9). *ATHB13* and *ATHB20* were in the transitory region, whereas *ATHB23* was not. Since the inference of co-expression networks and DPs represent two independent methods (*Materials and Methods*), this result is plausible. To generate a seed-to-seedling co-expression network a high cut-off for Pearson correlation (0.98) was used, whereas for the DPs a lower cut-off was used (0.85). We therefore conclude that *ATHB23* has a weaker interaction with transcripts of the transitory region of the seed-to-seedling network than *ATHB13* and *ATHB20*, which showed strong interactions with many transcripts of this region (Figure 9). In our data-set 58 transcripts were found to be co-expressed with *ATHB13* and *ATHB20*. Within the transitory region, many key regulators of photosynthesis and root development interact in the regulation of seeding establishment. Seven genes previously characterized with a role in photosynthesis and five with a role in root development were identified from this co-expression analysis (Figure 9). Photosynthesis-related genes included *PHOTOSYSTEM I SUBUNIT F* (*PSAF*; At1g31330); *GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE B SUBUNIT* (*GAPB*; At1g42970); *CHLOROPHYLL A-B BINDING FAMILY PROTEIN* (*CP22*; At1g44575); *NDH-DEPENDENT CYCLIC ELECTRON FLOW 1* (*NDF4*; At3g16250); *PLASTID TRANSCRIPTIONALLY ACTIVE 5* (*PTAC5*; At4g13670); *GLUTAMINE SYNTHETASE 2* (*GS2*; At5g35630) and *PS II OXYGEN-EVOLVING COMPLEX 1* (*PSBO1*; At5g66570). Root-related genes were: *AUXIN RESISTANT 3* (*IAA17*; At1g04250); *DUF538* (At1g09310); *WRKY DNA-BINDING PROTEIN 36* (*WRKY36*; At1g69810); *TONOPLAST INTRINSIC PROTEIN 1* (*TIP1*; At2g36830) and *MIZU-KUSSEI 1*(*MIZ1*; At2g41660). We phenotyped T-DNA insertion mutants for *ATHB13* and *ATHB20* (*athb13-1* and *athb20-1*) with no previously described function in the seed-to-seedling phase transition. A major effect of the *athb13-1* knock-out was a significantly increased primary root length relative to the wild type, whereas *athb20-1* did not show a similar root phenotype (Fig. 9 C, D). This suggests that *ATHB13* negatively regulates root growth during seedling establishment. This example demonstrates how a transcriptional module together with a co-expression network can generate insights that are not immediately apparent from simple transcript profile comparisons.

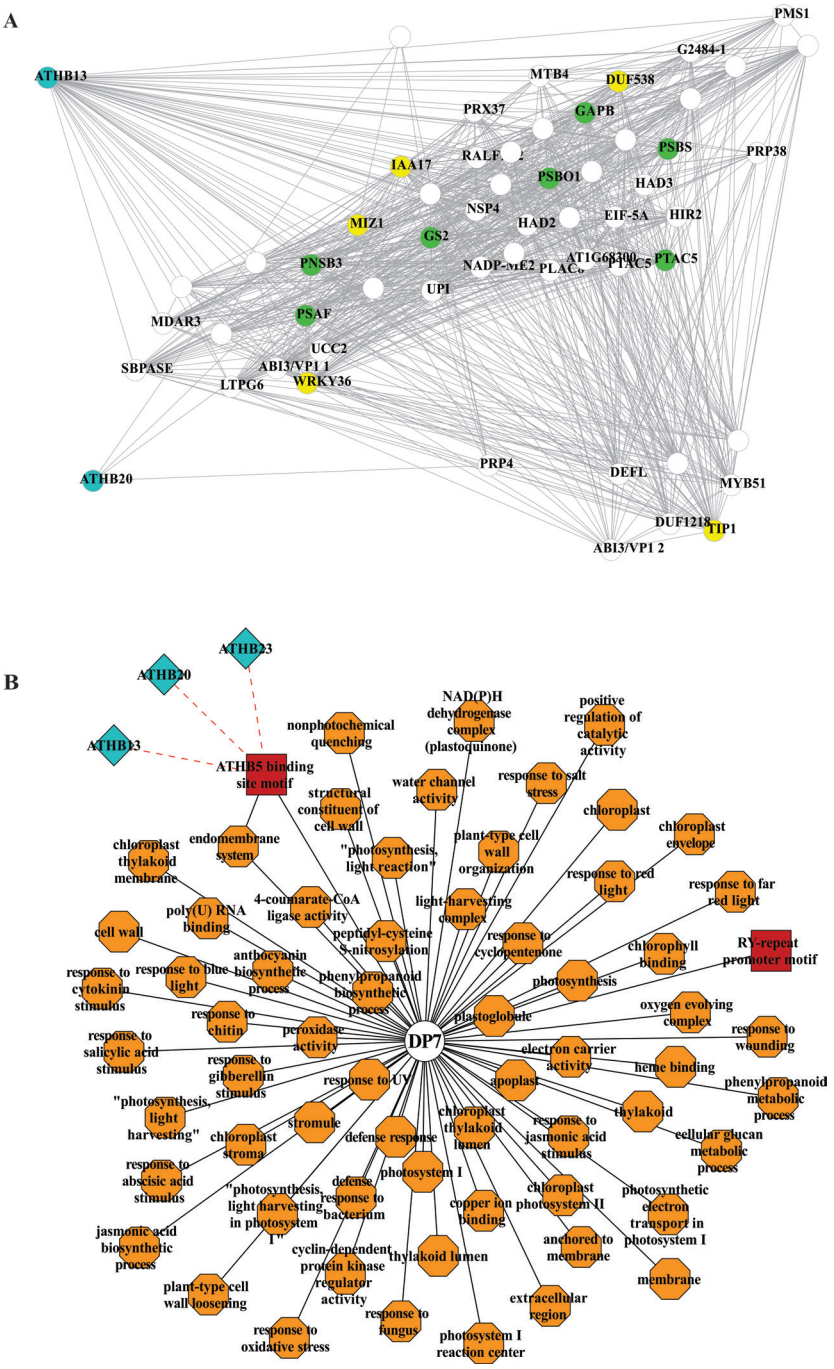


Figure 9. Continued

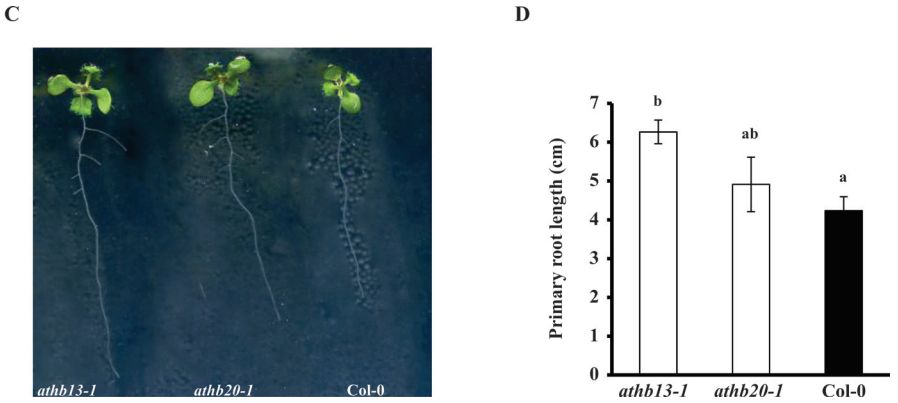


Figure 9. A. Gene network interactions of *ATHB13* and *ATHB20* as derived from the seed-to-seedling network. Green nodes represent genes known to be associated with photosynthesis, whereas yellow nodes represent genes associated with root development. **B.** Transcriptional modules predicting regulators of the transitory region (DP7) of the seed-to-seedling transition. DNA motifs (red rectangle) and GO terms (orange octagons) that are significantly overrepresented (p <0.001) within the transitory gene set (open ellipse) together with co-expressed transcription factors (light-blue diamonds). The transitory region gene set possesses a transcriptional module, of which ATHB5 is known to be associated with seedling development. All DPs with DNA motif, TF and GO term enrichments are listed in Supplemental Table S3. **C.** Root growth phenotype of mutant lines (*athb13-1* and *athb20-1*) and wild type (Col-0). Roots were scanned at 15 days after transfer to MS medium. **D**. Primary root length comparison at 15 days after transfer to MS medium. Different letters above error bars indicate statistically significant differences between lines (p < 0.01, n=27).

Discussion

To reach an autotrophic state, the seed-to-seedling transition is assumed to temporally and spatially employ various regulatory factors. These regulatory factors modulate the controlling genes and proteins, which will eventually govern seedling establishment. A detailed transcriptome analysis with the inference of co-expression networks and transcriptional modules of the seed-to-seedling transition resulted in the identification of several co-regulating modules of gene expression, coinciding with specific developmental stages. Each developmental stage was clearly distinguishable in a PCA plot, which suggests that changes in gene expression are already occurring within six hours of imbibition. The changes in the number of transcripts which were highly abundant during the first six hours of imbibition demonstrates that transcripts that had accumulated during seed maturation are degraded and that the induction of translation-associated transcripts, involved in germination, may commence directly upon imbibition. This observation confirms that of Dekkers *et al.* (2013), who

showed that the majority of seed maturation-repressed genes are reactivated in the first phase of Arabidopsis seed germination (dry to imbibed seed). Moreover, the enriched GO terms for DS peak-transcripts corroborate a previous study of gene co-expression networks of seed maturation in *Medicago truncatula* (Verdier *et al.*, 2013), in which genes expressed during seed maturation were also associated with GO terms such as *lipid binding* and *heat response*. The overrepresented GO terms for 6H peak–genes are associated with *RNA binding* and *nucleic acid binding*, and these are in agreement with a previous study of mitochondrial biogenesis during seed germination (Law *et al.*, 2012). This previous study showed that transcripts that increased in abundance in early germination were overrepresented for *RNA processing* and encode mitochondrial biogenesis functions, which precede crucial bioenergetic and metabolic functions. This observation correlates with our 6H peak-gene results and it postulates that an early transcript expression (6H) may act as a signal to prompt the expression of genes involved in various metabolic functions required for germination and seedling establishment. Clusters of TR, RP and RH showed only few peaking genes and a minor change in transcript expression. The GC peak-gene set, however, is overrepresented for microtubule motor activity genes (*mitotic kinesins*), which are associated with the cell cycle (Vanstraelen et al., 2006). Thus, this cluster has genes playing roles in cellular organization and dynamics, chromosome movement and cytokinesis (Vanstraelen et al., 2006), which prompts us to conclude that, at the GC stage, high cell division activity occurs. Genes peaking at OC are enriched for processes supporting tissue growth, such as *transport* and *lipid binding*. Taken together, overrepresentation results for peak-gene sets, reflect two major transitions in gene activity. These results, furthermore, imply that a transcript preparation for germination and early seedling development is initiated already at early imbibition (6H peak-genes) and an efficient seedling establishment is ensured by additional distinct transcript expression of GC and OC peak-genes.

Besides overrepresentation of GO terms for the peak–gene sets, changes in transcriptional modules (Belmonte *et al.*, 2013) of these gene sets link to TFs that control the functions of the regulatory networks. The transcriptional module for transcripts that accumulated in DS include DNA motifs (G-box, ABRE and ABFs) linking to bZIP transcription factors such as *ABA INSENSITIVE 5* (*ABI5*). *ABI5* is known to play a role in maturing seeds (Finkelstein *et al.*, 2005) and to be a major regulator of LATE EMBRYOGENESIS ABUNDANT PROTEINS (LEAs) (Nakashima *et al.*, 2006). Additionally, *ABI5* has been shown to play an important role in the expression of an ABA-responsive gene *RESPONSIVE TO DESSICATION 29B* (*RD29B*) which contains the ABRE motif (Nakashima *et al.*, 2006). Thus, the occurrence of DS peak-genes are in agreement with the known regulatory network in

maturing seeds. Transcripts of key regulators of germination are associated with 6H peak-genes and they are rapidly down-regulated during early seedling establishment. The 6H peak-genes data set is enriched for the TELO-box motif which is associated with genes involved in DNA replication (Wang *et al.*, 2011), whereas the transcriptional module of OC peak-genes is enriched for the ATHB5 DNA binding site motif known to be associated with the regulation of ABA-responsiveness and seedling development (Johannesson *et al.*, 2003).

The transcriptional modules identify potential regulatory circuits that control processes associated with the GO terms of seed-to-seedling phase transitions. The transcriptional interactions in each of the stage peak-genes are distinct, with the DS peak-genes, 6H peak-genes, and OC peak-genes showing greater numbers of highly expressed genes, and a transcriptional coordination according to the modules. Interestingly, TR, RP, and RH peak-genes with only a few highly expressed transcripts, suggest that a common transcriptional mechanism may be responsible for seedling establishment which continues by activation of additional transcriptional mechanisms (GC peak-genes and OC peak-genes).

To better understand the seed-to-seedling phase transition, we examined the distribution of these peak-genes over the seed-to-seedling co-expression network. Co-expression analysis suggests that stage peak-genes capture transcriptional interactions associated with developmental stages from dry seed to fully open cotyledons. The seed-to-seedling network shows three regions of interaction (dry, germination and the seedling state) with a transitory region between the germination and seedling stages. It suggests that functional differentiation within the transitional developmental stages occurs mainly through two distinct processes, namely germination (6H peak-genes) and seedling establishment (OC peak-genes) state.

Our analysis highlights this network as a powerful tool to understand the regulation of the transition from a quiescent dry state to a photo-autotrophic one. Moreover, the DP analysis reduces the complexity of co-expression data (Brady *et al.*, 2007; Belostotsky *et al.*, 2009) and helps to understand how transcript abundance changes over successive developmental stages (Brady *et al.*, 2007; Belmonte *et al.*, 2013). For example, the DP1 gene set is overrepresented for GO terms associated with chlorophyll biosynthesis and photosynthesis. Chlorophyll biosynthesis is controlled during the critical initial emergence of seedlings from darkness into light (Huq *et al.*, 2004; Moon *et al.*, 2008). Thus, the DP1 co-expression pattern may serve as a blueprint for chlorophyll biosynthesis and photosynthesis during the seed-to-seedling transition. Two other gene sets (DP6 and DP10) show clusters of transcripts highly expressed in the TR, RP and RH stages. DP6 and DP10 are overrepresented for *DNA unwinding*, *ribosome* and *mitochondrion*. Expression of these dominant

patterns point to rapid re-establishment of various biochemical activities to support seedling establishment. Biogenesis of mitochondria is known to provide energy for cellular processes, which is critical for seed germination and seedling establishment (Law *et al.*, 2012; Jiang *et al.*, 2013). Consistent with these observations, transcripts associated with mitochondrion display high expression already after six hours of imbibition (related to DP10), which suggests that mitochondrial biogenesis is required already at 6H to ensure successful seedling establishment. The predicted regulatory circuitry controlling transcripts expressed during the seed-to-seedling transition appears to be essential for successful early seedling establishment.

The transitory region between germination and the seedling stage in the seed-to-seedling network is determined by essentially one dominant expression pattern, namely DP7. The three TFs identified in DP7 are members of the HDZip gene family which are classified into four groups (Henriksson *et al.* (2005). Of the three identified family members, only two (*ATHB13* and *ATHB20*) displayed strong interactions within the seed-to-seedling network. These TFs are paralogs, and have similarities and differences in their response. *ATHB13* is involved in cold tolerance (Cabello *et al.*, 2012), whereas *ATHB20* has been assigned roles in ABA-, salt and cold responses (Henriksson *et al.*, 2005), as well as seed dormancy (Barrero *et al.*, 2010). *ATHB23* and *ATHB20* play a role in light signalling (Barrero *et al.*, 2010; Choi *et al.*, 2014), but none of these have previously been related to the seed-to-seedling phase transition.

These TFs represent high-confidence candidates of previously described regulatory factors present in DP7 and they probably have a role in coordinating gene partners within this dominant pattern (Belmonte *et al.*, 2013). The transcriptional module for DP7, which shows a significant increase from the last stage of germination (RP) to the first stage of seedling establishment (RH) and keeps increasing thereafter, suggests a regulatory process involving over 1.600 transcripts. The co-expression regulatory circuit of *ATHB13* suggests that this TF plays a role in the regulation of transcripts related to photosynthesis and root growth. However, there is no evidence for a connection among *ATHB13*, root growth and photosynthesis but a correlation has been reported between photosynthesis and *ATHB2*, another ATHB family member (Carabelli *et al.*, 1996; Hu *et al.*, 2013), and between root growth and *ATHB8* (Baima *et al.*, 1995). In addition, *ATHB13* has been described recently to play a crucial role in Arabidopsis development (Ribone *et al.*, 2015). This transcription factor appeared to be vital for pollen germination, the mutant showed defected siliques, and it may also play a negative role in stem elongation. Thus, this suggests that *ATHB13* is part of the same regulatory pathway leading to root growth or either photosynthesis for successful seedling establishment. This is corroborated by the root phenotype of

athb13-1 (Fig. 9). To our knowledge, the role of *ATHB13* in the regulatory circuit for root development is unknown. Increased root growth of *athb13-1* as compared to Col-0 suggests a negative feedback regulation in root development regulation, as previously shown for *ATHB8* in auxin signalling (Baima *et al.*, 2014) and *ATHB13* in stem growth (Ribone *et al.*, 2015). This repressing activity in root development by *ATHB13* may indicate antagonistic activities of modulating signals towards the completion of seedling establishment.

Conclusion

Our seed-to-seedling gene expression network describes global transcriptional interactions for two distinctive developmental states, namely germination and seedling development. Evidently, the seed-to-seedling transition is related to the agronomic trait of seedling establishment and vigour, and understanding the associated transcriptional regulatory network will facilitate studies to ultimately enhance seedling establishment in agriculture. The seed-to-seedling gene expression network provides a template to postulate new hypotheses about transcriptional regulators and their interactions. Many previously described transcription factors interact in the regulation of the seed-to-seedling transition. How seedling establishment may be improved by this transition, depends on the interaction activity of factors promoting signalling across the germination and seedling stages. The present data suggest that previously unknown regulators identified in the transitory region act through known regulatory components to promote information controlling development later in seedlings. Therefore, further investigation of co-expressed genes in the transitory region might lead to an answer for the connection between co-expressed genes and seedling establishment.

Acknowledgements

This work was funded by “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior” (CAPES, Brazil). We thank Raquel L. Cabello for providing the *athb13-1* and José M. Barrero for the *athb20-1* mutants. We thank Bas Dekkers for valuable discussions and critical reading of the manuscript.

References

Aoki K, Ogata Y, Shibata D.2007 Approaches for Extracting Practical Information from Gene Co-expression Networks in Plant Biology. *Plant and Cell Physiology* 48: 381-390

Aryee M, Gutierrez-Pabello J, Kramnik I, Maiti T, Quackenbush J.2009 An improved empirical bayes approach to estimating differential gene expression in microarray time-course data: BETR (Bayesian Estimation of Temporal Regulation). *BMC Bioinformatics* **10:** 409

Baima S, Forte V, Possenti M, Peñalosa A, Leoni G, Salvi S, Felici B, Ruberti I, Morelli G.2014 Negative Feedback Regulation of Auxin Signaling by ATHB8/ACL5–BUD2 Transcription Module. *Molecular Plant* **7:** 1006-1025

Baima S, Nobili F, Sessa G, Lucchetti S, Ruberti I, Morelli G.1995 The expression of the Athb-8 homeobox gene is restricted to provascular cells in Arabidopsis thaliana. *Development* **121:** 4171-4182

Barrero JM, Millar AA, Griffiths J, Czechowski T, Scheible WR, Udvardi M, Reid JB, Ross JJ, Jacobsen JV, Gubler F.2010 Gene expression profiling identifies two regulatory genes controlling dormancy and ABA sensitivity in Arabidopsis seeds. *The Plant Journal* **61:** 611-622

Bassel GW, Lan H, Glaab E, Gibbs DJ, Gerjets T, Krasnogor N, Bonner AJ, Holdsworth MJ, Provart NJ.2011 Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. *Proceedings of the National Academy of Sciences* **108:** 9709-9714

Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, Hashimoto M, Fei J, Harada CM, Munoz MD, Le BH, Drews GN, Brady SM, Goldberg RB, Harada JJ.2013 Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. *Proceedings of the National Academy of Sciences* **110:** E435-E444

Belostotsky DA, Gadbury G, Garrett K, Allison D (2009) Challenges and Approaches to Statistical Design and Inference in High-Dimensional Investigations. *In* Plant Systems Biology, Vol 553. Humana Press, pp 181-206

Belostotsky DA, Orlando D, Brady S, Koch J, Dinneny J, Benfey P (2009) Manipulating Large-Scale Arabidopsis Microarray Expression Data: Identifying Dominant Expression Patterns and Biological Process Enrichment. *In* Plant Systems Biology, Vol 553. Humana Press, pp 57-77

Brady SM, Orlando DA, Lee J-Y, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN.2007 A High-Resolution Root Spatiotemporal Map Reveals Dominant Expression Patterns. *Science* **318:** 801-806

Cabello JV, Arce AL, Chan RL.2012 The homologous HD-Zip I transcription factors HaHB1 and AtHB13 confer cold tolerance via the induction of pathogenesis-related and glucanase proteins. *The Plant Journal* **69:** 141-153

Cadman CSC, Toorop PE, Hilhorst HWM, Finch-Savage WE.2006 Gene expression

profiles of Arabidopsis Cvi seeds during dormancy cycling indicate a common underlying dormancy control mechanism. *The Plant Journal* **46:** 805-822

Carabelli M, Morelli G, Whitelam G, Ruberti I.1996 Twilight-zone and canopy shade induction of the Athb-2 homeobox gene in green plants. *Proceedings of the National Academy of Sciences of the United States of America* **93:** 3530-3535

Choi H, Jeong S, Kim DS, Na HJ, Ryu JS, Lee SS, Nam HG, Lim PO, Woo HR.2014 The homeodomain-leucine zipper ATHB23, a phytochrome B-interacting protein, is important for phytochrome B-mediated red light signaling. *Physiologia Plantarum* **150:** 308-320

Dekkers BJW, Pearce S, van Bolderen-Veldkamp RP, Marshall A, Widera P, Gilbert J, Drost H-G, Bassel GW, Müller K, King JR, Wood ATA, Grosse I, Quint M, Krasnogor N, Leubner-Metzger G, Holdsworth MJ, Bentsink L.2013 Transcriptional Dynamics of Two Seed Compartments with Opposing Roles in Arabidopsis Seed Germination. *Plant Physiology*

Ding Y-H, Liu N-Y, Tang Z-S, Liu J, Yang W-C.2006 Arabidopsis GLUTAMINE-RICH PROTEIN23 Is Essential for Early Embryogenesis and Encodes a Novel Nuclear PPR Motif Protein That Interacts with RNA Polymerase II Subunit III. *The Plant Cell* **18:** 815-830

Fait A, Angelovici R, Less H, Ohad I, Urbanczyk-Wochniak E, Fernie AR, Galili G.2006 Arabidopsis Seed Development and Germination Is Associated with Temporally Distinct Metabolic Switches. *Plant Physiology* **142:** 839-854

Feng CZ, Chen Y, Wang C, Kong YH, Wu WH, Chen YF.2014 Arabidopsis RAV1 transcription factor, phosphorylated by SnRK2 kinases, regulates the expression of ABI3, ABI4, and ABI5 during seed germination and early seedling development. *Plant J* **80:** 654-668

Finch-Savage WE, Cadman CSC, Toorop PE, Lynn JR, Hilhorst HWM.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, Gampala SS, Lynch TJ, Thomas TL, Rock CD.2005 Redundant and distinct functions of the ABA response loci ABA-INSENSITIVE(ABI)5 and ABRE-BINDING FACTOR (ABF)3. *Plant Mol Biol* **59:** 253-267

Fulda M, Schnurr J, Abbadi A, Heinz E, Browse J.2004 Peroxisomal Acyl-CoA Synthetase Activity Is Essential for Seedling Development in Arabidopsis thaliana. *The Plant Cell Online* **16:** 394-405

Gentleman R, Carey V, Bates D, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini A, Sawitzki G, Smith C, Smyth G, Tierney L, Yang J, Zhang J.2004 Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* **5:** R80

Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S, Smyth GK (2005) limma: Linear

Models for Microarray Data. *In* Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer New York, pp 397-420

Gutierrez L, Van Wuytswinkel O, Castelain M, Bellini C.2007 Combined networks regulating seed maturation. *Trends in Plant Science* **12:** 294-300

Hammani K, Gobert A, Small I, Giegé P.2011 A PPR protein involved in regulating nuclear genes encoding mitochondrial proteins? *Plant Signaling & Behavior* **6:** 748-750

Henriksson E, Olsson ASB, Johannesson H, Johansson H, Hanson J, Engström P, Söderman E.2005 Homeodomain Leucine Zipper Class I Genes in Arabidopsis. Expression Patterns and Phylogenetic Relationships. *Plant Physiology* **139:** 509-518

Holdsworth MJ, Bentsink L, Soppe WJJ.2008 Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytologist* **179:** 33-54

Holdsworth MJ, Finch-Savage WE, Grappin P, Job D.2008 Post-genomics dissection of seed dormancy and germination. *Trends in Plant Science* **13:** 7-13

Hu W, Franklin KA, Sharrock RA, Jones MA, Harmer SL, Lagarias JC.2013 Unanticipated regulatory roles for Arabidopsis phytochromes revealed by null mutant analysis. *Proceedings of the National Academy of Sciences* **110:** 1542-1547

Huq E, Al-Sady B, Hudson M, Kim C, Apel K, Quail PH.2004 PHYTOCHROME-INTERACTING FACTOR 1 Is a Critical bHLH Regulator of Chlorophyll Biosynthesis. *Science* **305:** 1937-1941

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP.2003 Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research* **31:** e15

Jiang L, Liu Y, Sun H, Han Y, Li J, Li C, Guo W, Meng H, Li S, Fan Y, Zhang C.2013 The Mitochondrial Folylpolyglutamate Synthetase Gene Is Required for Nitrogen Utilization during Early Seedling Development in Arabidopsis. *Plant Physiology* **161:** 971-989

Johannesson H, Wang Y, Hanson J, Engström P.2003 The Arabidopsis thaliana homeobox gene ATHB5 is a potential regulator of abscisic acid responsiveness in developing seedlings. **51:** 719-729

Law SR, Narsai R, Taylor NL, Delannoy E, Carrie C, Giraud E, Millar AH, Small I, Whelan J.2012 Nucleotide and RNA Metabolism Prime Translational Initiation in the Earliest Events of Mitochondrial Biogenesis during Arabidopsis Germination. *Plant Physiology* **158:** 1610-1627

Maia J, Dekkers BJW, Provart NJ, Ligterink W, Hilhorst HWM.2011 The Re-Establishment of Desiccation Tolerance in Germinated *Arabidopsis thaliana* Seeds and Its Associated Transcriptome. *PLos ONE* **6**

Mazzella MA, Arana MV, Staneloni RJ, Perelman S, Rodriguez Batiller MJ, Muschietti J, Cerdán PD, Chen K, Sánchez RA, Zhu T, Chory J, Casal JJ.2005 Phytochrome Control of the Arabidopsis Transcriptome Anticipates Seedling Exposure to Light. *The Plant Cell Online* **17:** 2507-2516

Moon J, Zhu L, Shen H, Huq E.2008 PIF1 directly and indirectly regulates chlorophyll

biosynthesis to optimize the greening process in Arabidopsis. *Proceedings of the National Academy of Sciences*

Murashige T, Skoog F.1962 A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* **15:** 473-497

Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K.2006 Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of Arabidopsis. *Plant Molecular Biology* **60:** 51-68

Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K.2006 Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of Arabidopsis. *Plant Mol Biol* **60:** 51-68

Narsai R, Law SR, Carrie C, Xu L, 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

Rajjou L, Gallardo K, Debeaujon I, Vandekerckhove J, Job C, Job D.2004 The Effect of α-Amanitin on the Arabidopsis Seed Proteome Highlights the Distinct Roles of Stored and Neosynthesized mRNAs during Germination. *Plant Physiology* **134:** 1598-1613

Ribone PA, Capella M, Chan RL.2015 Functional characterization of the homeodomain leucine zipper I transcription factor AtHB13 reveals a crucial role in Arabidopsis development. *Journal of Experimental Botany*

Tepperman JM, Hwang YS, Quail PH.2006 phyA dominates in transduction of red-light signals to rapidly responding genes at the initiation of Arabidopsis seedling de-etiolation. *Plant J* **48:** 728-742

Usadel B, Obayashi T, Mutwil M, Giorgi FM, Bassel GW, Tanimoto M, Chow A, Steinhauser D, Persson S, Provart NJ.2009 Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. *Plant, Cell & Environment* **32:** 1633-1651

Verdier J, Lalanne D, Pelletier S, Torres-Jerez I, Righetti K, Bandyopadhyay K, Leprince O, Chatelain E, Vu BL, Gouzy J, Gamas P, Udvardi MK, 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

Wan CY, Wilkins TA.1994 A Modified Hot Borate Method Significantly Enhances the Yield of High-Quality RNA from Cotton (Gossypium hirsutum L.). *Analytical Biochemistry* **223:** 7-12

Wang J, Wang Y, Wang Z, Liu L, Zhu X-G, Ma X.2011 Synchronization of cytoplasmic and transferred mitochondrial ribosomal protein gene expression in land plants is linked to Telo-box motif enrichment. *BMC Evolutionary Biology* **11:** 161

Supporting information

If not displayed bellow, supplementary information can be downloaded from : http://www.wageningenseedlab.nl/thesis/atsilva/SI/

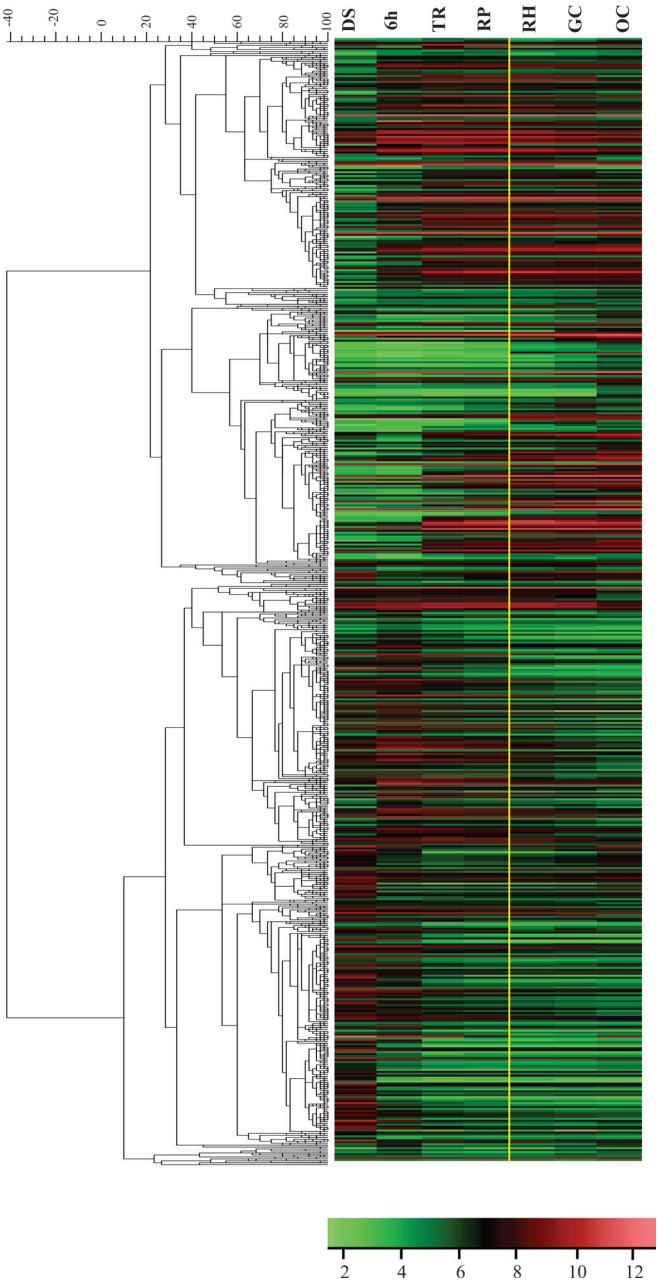


Figure S1. Hierarchical clustering analysis of differentially expressed Arabidopsis seed-to-seedling transition transcripts. 19.130 transcripts, which showed a significant level of differential expression, at least two stages, were clustered hierarchically, following the Bayesian Estimation of Temporal Regulation method. Horizontal lines point out the expression pattern of each gene, with the seed-to-seedling transition stages in the different columns. Seed-to-seedling stages: DS – dry seeds; 6H – six hours imbibed; TR – testa rupture; RP – radicle protrusion; RH – root hair; GC – greening cotyledons; and OC – cotyledons fully opened.

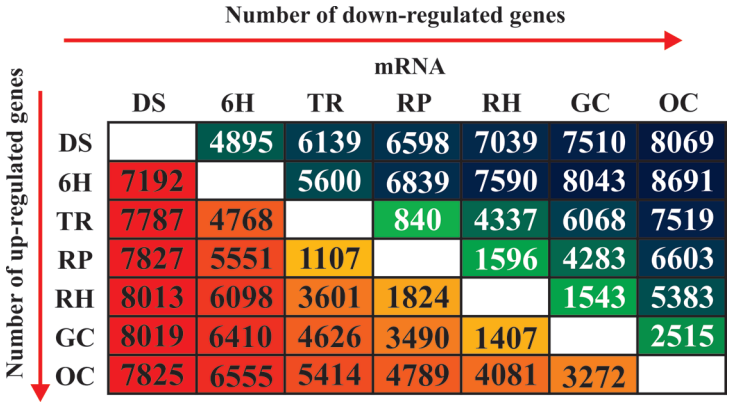


Figure S2. Overview of changes in transcript levels across seed-to-seedling developmental stages. Number of transcripts significantly different in abundance between all combinations of the different morphological developmental stages.

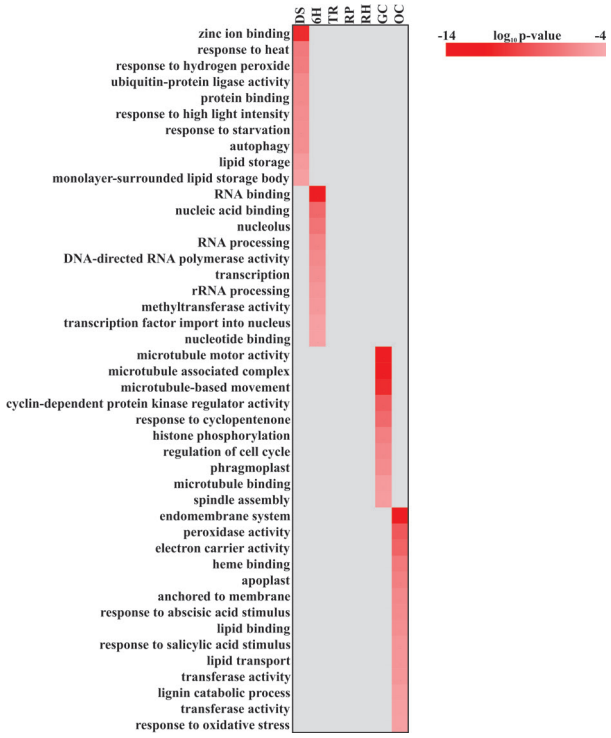


Figure S3. Heat map showing the *p* value significance of GO enrichment terms for peak–transcript sets (Supplemental Table S3). GO terms listed are for biological process or/and cellular components that were overrepresented in peak-transcripts for each stage.

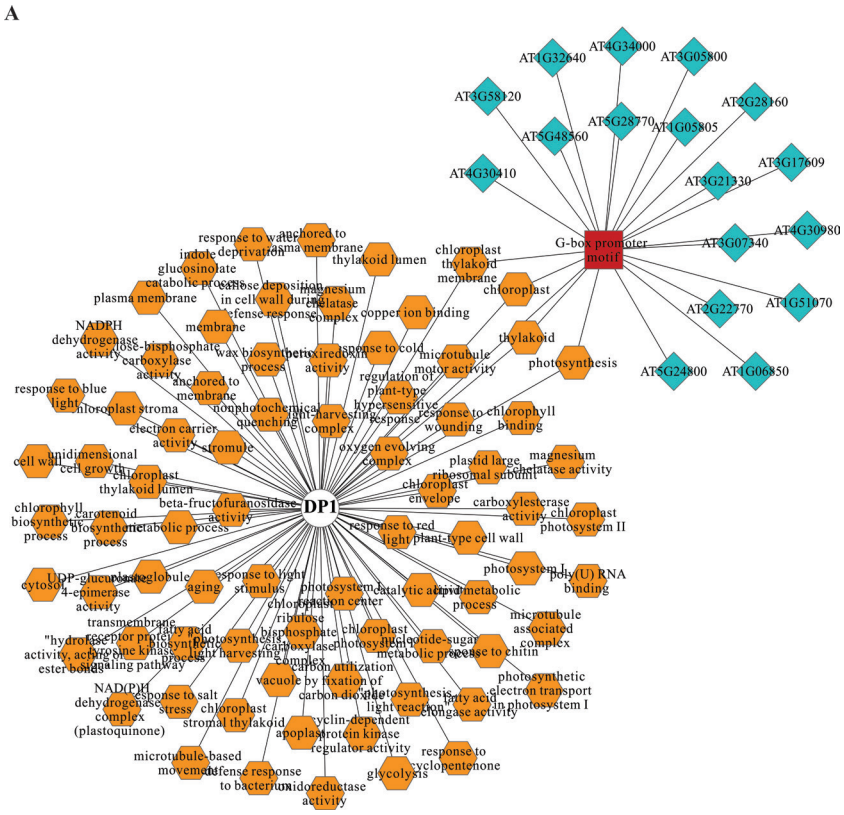


Figure S4. Continued

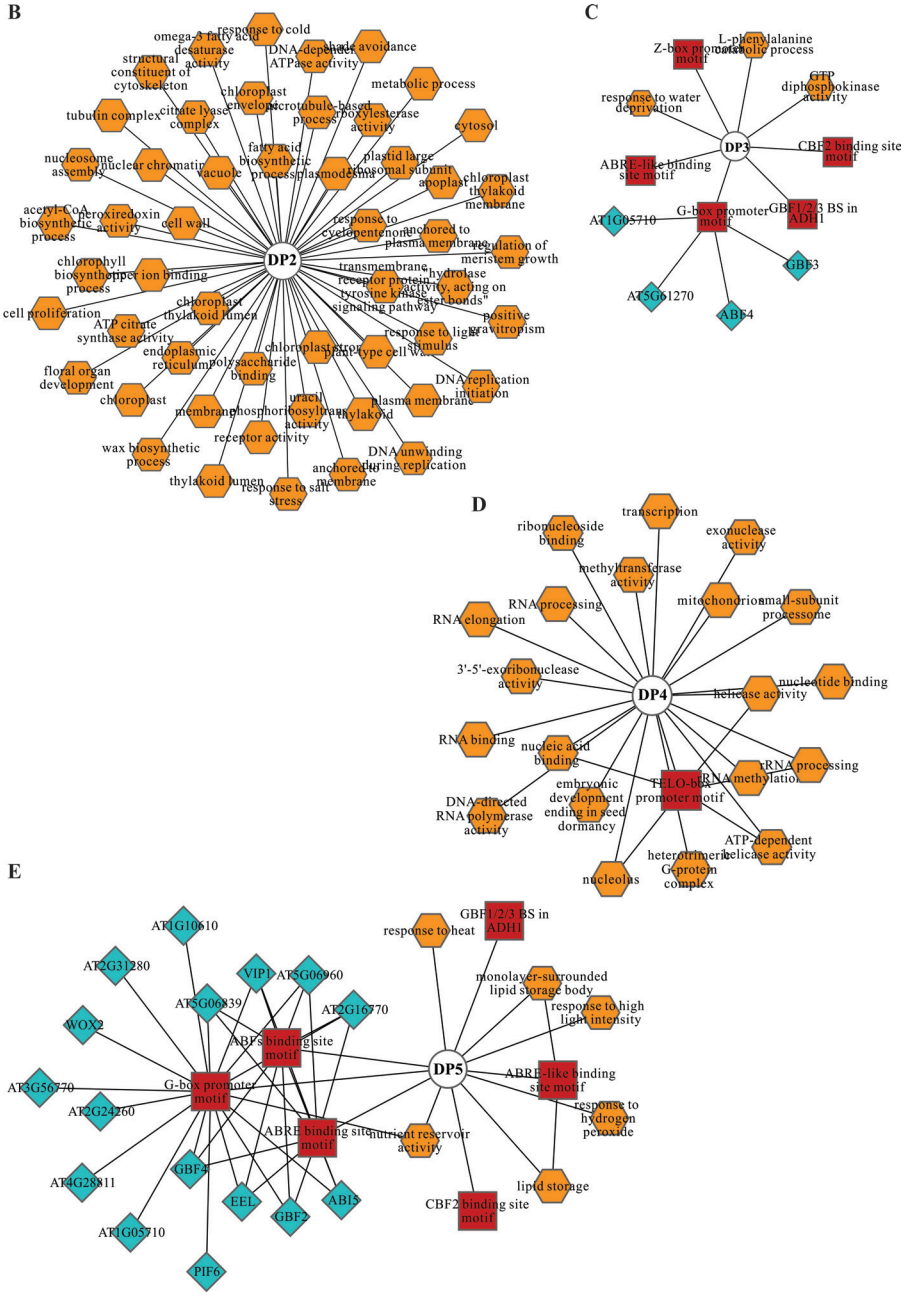


Figure S4. Continued

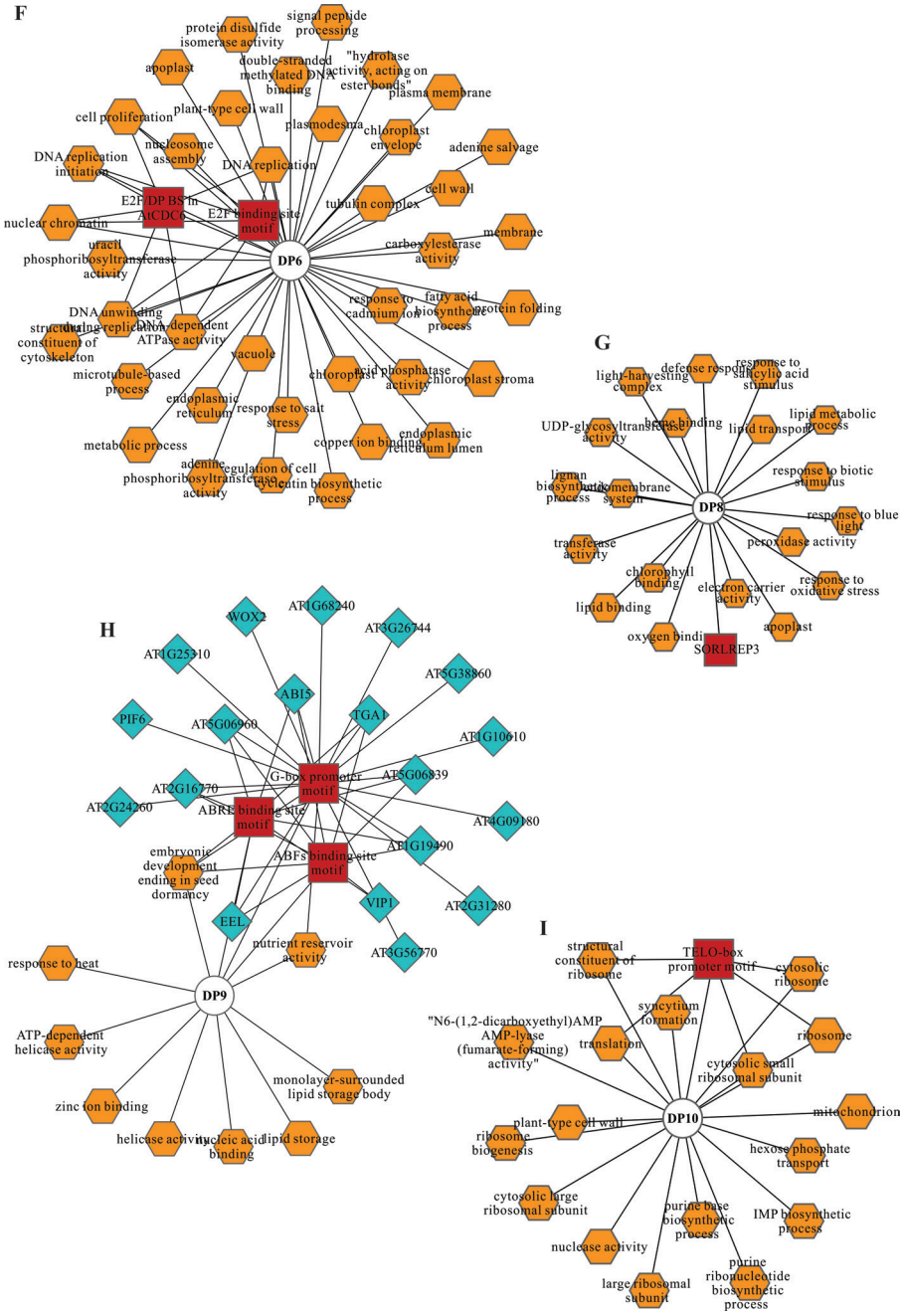


Figure S4. Transcriptional modules predicting regulation of the seed-to-seedling transition. DNA motifs (red rectangles) and GO terms (orange octagons) that are significantly overrepresented (p <0.001) within the peak–transcripts sets (open ellipses) together with co-expressed transcription factors (light-blue diamonds).

Dataset S1: Abundance of all transcripts in the seed-to-seedling transition.

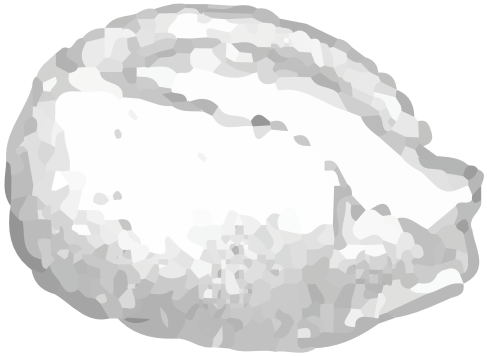
Table S1: Validation of the seed-to-seedling transcriptome data set performed by comparison with previously published expression patterns of genes known to be differential expressed across seed to seedling developmental stages

Table S2: Transcript comparisons between previous and next stages, and their enriched GO terms, DNA motifs, and metabolic pathways.

Table S3: Peak-transcripts and dominant pattern sets. GO terms, DNA motifs and metabolic pathways for each gene set.

Chapter 3

Metabolite profiling reveals two metabolic shifts during the seed-to-seedling transition in *Arabidopsis thaliana*



Anderson Tadeu Silva, Wilco Ligterink and Henk W. M. Hilhorst
(*in preparation for submission*)

Summary

During early seedling establishment, the transition from a quiescent metabolic state (dry seeds) to the active state of a vigorous seedling is crucial in the plant's life cycle. We further analysed this complex physiological trait by measuring the changes in primary metabolism that occur during the transition in order to determine which metabolic networks are operational. The transition involves several developmental stages between seed germination and seedling establishment, i.e. between imbibition of the mature dry seed and opening of the cotyledons, the final stage of seedling establishment. We wanted to determine if the advancement of growth is associated with certain signature metabolite profiles. Metabolite-metabolite correlation analysis underlined two profiles that seem to be implied in the metabolic preparation of seed germination and of efficient seedling establishment, respectively. To understand the regulation of the biochemical pathways involved in the seed-to-seedling transition, we correlated the metabolite profiles with transcription data obtained from the same developmental stages. We demonstrate that the transcriptional changes not always translate one to one to a proportional metabolic response. Nevertheless, this correlation analysis enabled us to identify which metabolic pathways are important during the seed-to-seedling transition, and to identify metabolite levels that seem to be controlled directly by transcript abundance. These results provide a global view of the transcriptional and metabolic changes during the seed-to-seedling transition in Arabidopsis and this opens up new perspectives for understanding the complex regulatory mechanism underlying this transition.

Introduction

Seed germination is a critical stage in the plant's life cycle which starts with the uptake of water by the dry seed, followed by embryo expansion and (commonly) radicle emergence. This process is characterized by a transition from a quiescent- to a metabolically highly active state (Penfield *et al.*, 2005; Fait *et al.*, 2006). Uptake of water by seeds occurs tri-phasic with a rapid initial uptake (phase I, *imbibition*). Phase I of seed germination occurs without visible morphological changes, and is characterized by seed-specific germination metabolism that prepares the seed for radicle protrusion (Fait *et al.*, 2006). Phase I is followed by a plateau phase (phase II) in which water content is constant but metabolic activity increases. Radicle protrusion through the embryo-surrounding structures marks the end of phase II and a further increase in water uptake (phase III) occurs as the embryonic axis elongates and the embryo establishes itself as a young seedling (Schopfer and Plachy, 1984; Nonogaki *et al.*, 2010; Bewley *et al.*, 2013).

Arabidopsis mutants have been extensively used for the functional analysis of genes involved in seed germination (Debeaujon and Koornneef, 2000; Lu and Hills, 2002; Fulda *et al.*, 2004; Yang *et al.*, 2013). Some mutations slow down seed germination but do not significantly arrest it. This phenomenon may occur as a result of a significant reduction in oil reserve content accumulated during seed maturation (Focks and Benning, 1998; Lu and Hills, 2002; Penfield *et al.*, 2005). However, another study with a mutant deficient in plastidic pyruvate kinase (*pkp1*) suggested that delayed seed germination and, consequently, seedling establishment is not caused specifically by a lack of seed oil reserves but may be related to reduced pyruvate kinase activity during germination (Andre and Benning, 2007). These observations suggest that seedling establishment is not only affected by mobilisation of reserves accumulated during seed development, but also by additional metabolic processes across developmental stages during the seed-to-seedling transition.

Although high-throughput functional genomic methods, such as transcriptomics and metabolomics have identified key genes and metabolites involved in seed germination (Fait *et al.*, 2006; Angelovici *et al.*, 2011; Toubiana *et al.*, 2012; Dekkers *et al.*, 2013), there is a lack of integration of these studies, e.g. combining transcriptomics with metabolomics, to further zoom in on key metabolic pathways during successive developmental stages. Integration of *omics* data may constitute an alternative and complementary strategy to identify target genes and their ultimate products (metabolites) regulating either single biochemical pathways or more complex developmental mechanisms (Mercke *et al.*, 2004). Powerful tools are now available for discovering links between transcripts and pathways of plant metabolism (Gutiérrez *et al.*, 2008; Verdier *et al.*, 2013; Cañas *et al.*, 2015). The availability of high-throughput metabolomics approaches produces extensive metabolite data sets which can be combined with gene expression data (Fait *et al.*, 2006). The combination of transcripts and metabolites has also been represented as networks including both transcripts and metabolites (Holdsworth *et al.*, 2008; Fukushima *et al.*, 2011; Lv *et al.*, 2014), which has been successfully used to discover regulatory and biosynthetic genes involved in the control of metabolite production (Verdier *et al.*, 2013).

The present study highlights the dynamics of metabolism across developmental stages during the seed-to-seedling transition in Arabidopsis. Metabolite-metabolite correlation analysis underlined two profiles that seem to be implied in the metabolic preparation of seed germination and of efficient seedling establishment, respectively. In addition, we analysed combined transcriptomics and metabolomics data, which provides a general framework illustrating the significance of metabolites and transcripts in the seed-to-seedling transition.

Experimental procedures

Plant material

Seeds of *Arabidopsis thaliana*, accession Columbia (Col-0 [N60000]), were cold stratified at 4 ºC in the dark for 72 h in Petri dishes on two layers of moistened blue filter paper. Germination was performed in standard plant growth chambers at 22 ºC under constant white light. To elucidate the changes in metabolomes that prepare for and accompany the transition from a seed into a photoautotrophic seedling, seven developmental stages during this transition were identified. Using a stereomicroscope, the successive developmental stages were selected as follows: (DS) mature dry seed; (6H) seed six hours upon imbibition; embryo swelling and (TR) testa rupture; (RP) protrusion of the radicle through the endosperm, followed by embryonic root growth and (RH) root hair formation, followed by (GC) greening cotyledons and (OC) fully opened cotyledons (Chapter 2).

Metabolite extraction, derivatization and GC-TOF-MS analysis

Metabolites were extracted as described before (Ribeiro *et al.*, 2014) with some modifications. Three replicates of approximately 350 seeds/seedlings (~5mg of dry material) from each developmental stage were pulverized in a 2 mL Eppendorf tube with an iron ball (6.3 mm) using a dismembrator (Sartorius®) cooled with liquid nitrogen. Then, 175 µl of methanol/chloroform (4:3 v/v) were added, after which 37.5 µl of the internal standard (1 $mg.mL^{-1}$ ribitol in water) was added. After 10 min in an ultrasonic bath, 50 µL of water was added, followed by 5 min of centrifugation at 17000 g in an Eppendorf centrifuge. After centrifugation the upper phase was transferred to a new tube. To the remaining material, 125 µL of a 1:1 v/v mix of methanol and chloroform was added, vortexed thoroughly and kept on ice for 10 minutes. Then, 50 µL of water was added and centrifuged for 5 min at 17000 g in an Eppendorf centrifuge. The upper phase was transferred to the previously collected upper phase. An aliquot of 30 µL of the joint upper phase was dried overnight by vacuum centrifugation.

Each sample was analysed by gas chromatography coupled to a quadrupole time of flight mass spectrometry system (GC-TOF-MS) as TMS derivatives. TMS derivatives were obtained by online derivatization (Combi PAL autosampler - CTC Analytics) as described previously (Ribeiro *et al.*, 2014). Sample aliquots of 2 µL were injected at a split ratio of 20:1 into an Optic 3 high-performance injector (ATAS) at 70 ºC, after which the injector was rapidly heated to 240 ºC at 6 ºC s^{-1}.

Chromatography was performed in an Agilent 6890 gas chromatograph (Agilent Technologies) coupled to a Pegasus III time-of-flight mass spectrometer (Leco Instruments) using a VF-5 ms capillary column (Varian; 30 m × 0.25 mm × 0.25 µm) including a 10-m guardian column with helium as carrier gas at a column flow rate of 1 mL min^{-1}. The oven temperature program was 2 min at 70 ºC, followed by a 10 ºC min^{-1} ramp to 310 ºC, 5 min at 310 ºC, and 6 min at 70 ºC before the next injection. The transfer line temperature was set at 270 ºC. The column effluent was ionized by electron impact at 70 eV. Mass spectra were recorded at 20 scans s^{-1} within a mass-to-charge ratio range of 50 to 600 at a source temperature of 200 ºC. A solvent delay of 295 s was set. The detector voltage was set to 1,650 V. An alkane mixture (C10–C17 and C19–C33) was used to determine the retention index of the metabolites. Starch was determined as described previously by (Ribeiro *et al.*, 2014). In order to confirm the sucrose trend from GC-TOF-MS, soluble carbohydrates were also determined as previously by (Ribeiro *et al.*, 2014). The supernatant was injected into a Dionex HPLC system (ICS 5000 $^{+}$ DC) to analyse the soluble carbohydrate content, using a CarboPac PA 1, 4- × 250-mm column preceded by a guard column (CarboPac PA 1, 4 × 50 mm), a gradient pump module (ICS 5000 Dual Pump, Dionex). Mono-, di-, and tri-saccharides were separated by elution in an increasing concentration of NaOH (20–350 mM) with a flow rate of 1 mL per minute. Peaks were identified by co-elution of soluble carbohydrate standards. Sugar quantity was corrected by mean of the internal standard (melezitose) and transformed to micrograms of sugar per milligram of dry material.

Data processing. Raw data was processed by ChromaTOF software 2.0 (Leco®), and further baseline correction, accurate mass calculation, data smoothing and noise reduction, followed by alignment between chromatograms was performed using the MetAlign software (Lommen, 2009). MSClust was used to remove metabolite signal redundancy in aligned mass peaks tables and to retrieve mass spectral information of metabolites using mass peak clustering (Tikunov *et al.*, 2012). The mass spectra of the representative masses were used for tentative identification by matching to the spectral libraries (National Institute of Standards and Technology [NIST08]; Golm metabolome database [http://gmd.mpimp-golm.mpg.de/ webcite]) and by comparison of the retention index calculated using a series of alkanes. Authentic reference standards were used to confirm the identity of the metabolites. Levels of identification are presented in Supplementary Dataset S1 according to Sumner *et al.* (2007).

Metabolomic analysis. The processed data was uploaded into MetaboAnalyst software (www.metaboanalyst.ca) according to the user's guide (Xia *et al.*, 2009; Xia and Wishart, 2011). Data normalization was performed by adjustment of

the concentrations based on biological input (dry weight) and reference feature (ribitol). Subsequently, generalized-logarithm transformation was performed, followed by unit scaling (mean-centred and divided by standard deviation of each variable). Multivariate analysis was performed using log transformed and scaled data. Statistically significant differences for the variables between seed-to-seedling transition developmental stages were tested by ANOVA followed by *post hoc* analysis for comparisons in multiple groups. The *p*-value resulting from ANOVA was adjusted to the false discovery rate-adjusted *p*-value (FDR). Principal component analysis (PCA) was performed on the entire data set.

Dominant profile , network and metabolite pathway enrichment analysis. *Fuzzy K-Means* clustering was used for identification of common patterns (Belostotsky *et al.*, 2009). Pearson correlation was performed for each identified profile. A table for each profile with correlation coefficient values among metabolites was exported to Cytoscape V.2.8.2 (Smoot *et al.*, 2011). Pathway enrichment analysis was performed for each profile using Metaboanalyst 3.0 (Xia *et al.*, 2015).

Results

Metabolic profiling of the seed-to-seedling transition

Upon GC-TOF-MS analysis of the seven developmental stages of the seed-to-seedling transition (see Chapter 2) we detected 144 metabolite peaks. From these, 44 could be identified (Dataset S1). Identified metabolites were subjected to one-way ANOVA and differences in metabolite levels between developmental stages were considered significant if $p < 0.01$ (FDR adjusted). Forty-three metabolites varied significantly between at least two developmental stages of the seed-to-seedling transition (Dataset S2). Principal component analysis (PCA) applied to the entire metabolome data set (identified and unknown compounds) showed an overall variation in metabolite levels among developmental stages (Figure 1A). The high score (66.2%) for principal component one (PC1), which coincides with the seed-to-seedling development shows that there are large changes in metabolite composition during this transition. A biplot, displaying the major contribution of metabolites to the different developmental stages (Figure 1B) suggests that dynamic changes in metabolite levels occur across the developmental stages. Galactinol, sucrose and N-acetylglutamic acid (NAcGlu) represented the most dominant metabolites in DS and 6H, whereas for the TR, RP and RH stages tyrosine, allantoin and urea were the most prominent metabolites. Most of the metabolites particularly accumulated in the final two stages of seedling establishment (GC and OC).

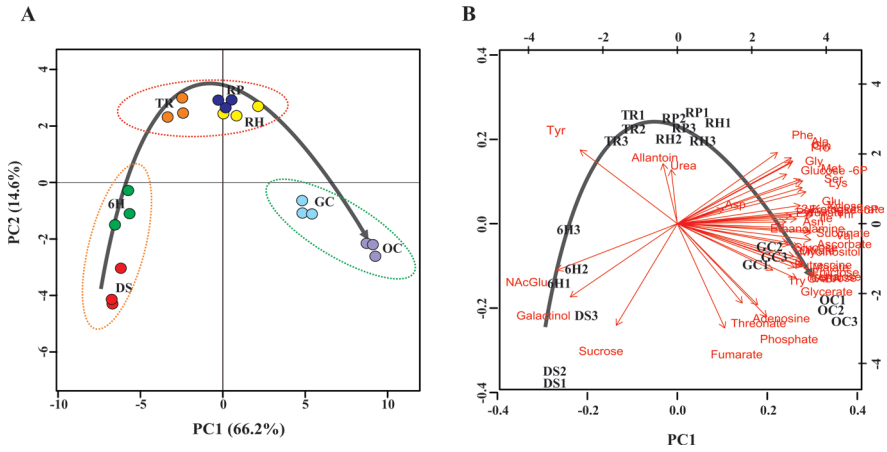


Figure 1. A. Principal component analysis (PCA) of metabolite profiling at seven developmental stages across the seed-to-seedling transition. Data points of the same colour represent sample replicates. Principal component 1 (PC1) explained 66.2% of metabolite variance at developmental stages and principal component 2 (PC2) explained 14.6%. Dashed ellipses show three clusters: initial stages of germination (DS and 6H); early-seedling stages (TR, RP and RH) and last stages of seedling establishment (GC and OC). **B.** Significant changes of the metabolites illustrated in a Biplot derived from the PCA-plot. Developmental stages are represented by: DS – dry seed; 6H – six hours imbibed; TR – testa rupture; RP – radicle protrusion; RH – root hair; GC – greening cotyledons; and OC – cotyledons fully opened. Arrow indicates the development.

Overall changes in carbohydrates. Levels of sorbose (2-fold), fructose (3-fold), glucose-6-phosphate (3-fold) and glucose (16-fold) were enhanced during the first six hours of imbibition and kept increasing thereafter (Figure 2) resulting in an increase in the levels of sucrose-derived monosaccharides at TR, such as fructose-6-phosphate (4 fold) and glucose-6-phosphate (13 fold). Only sucrose and galactinol showed a continuous decrease across the seed-to-seedling developmental stages. In contrast to the latter decrease, the levels of several amino acids increased (Figure 2). An initial reduction in sucrose content was observed at TR to about 50% of the content at 6H, suggesting that in the TR stage carbohydrate mobilization and metabolism are highly active (Figure 2). The opposite trend of changes in amino acid-, sucrose- and galactinol content suggests a shift from carbon- to nitrogen metabolism. In addition, sugars such as fructose, glucose, sorbose, xylose, trehalose, and glucose-6-phosphate increased during the final two stages of seedling establishment from 2-fold (GC) up to 10-fold (OC). These marked changes in sugar levels are predominantly associated with major carbohydrate metabolism. Although sucrose levels decreased, the abundance of all other carbohydrates increased, which suggests a substantial rate of import of the products from reserve mobilization into glycolysis.

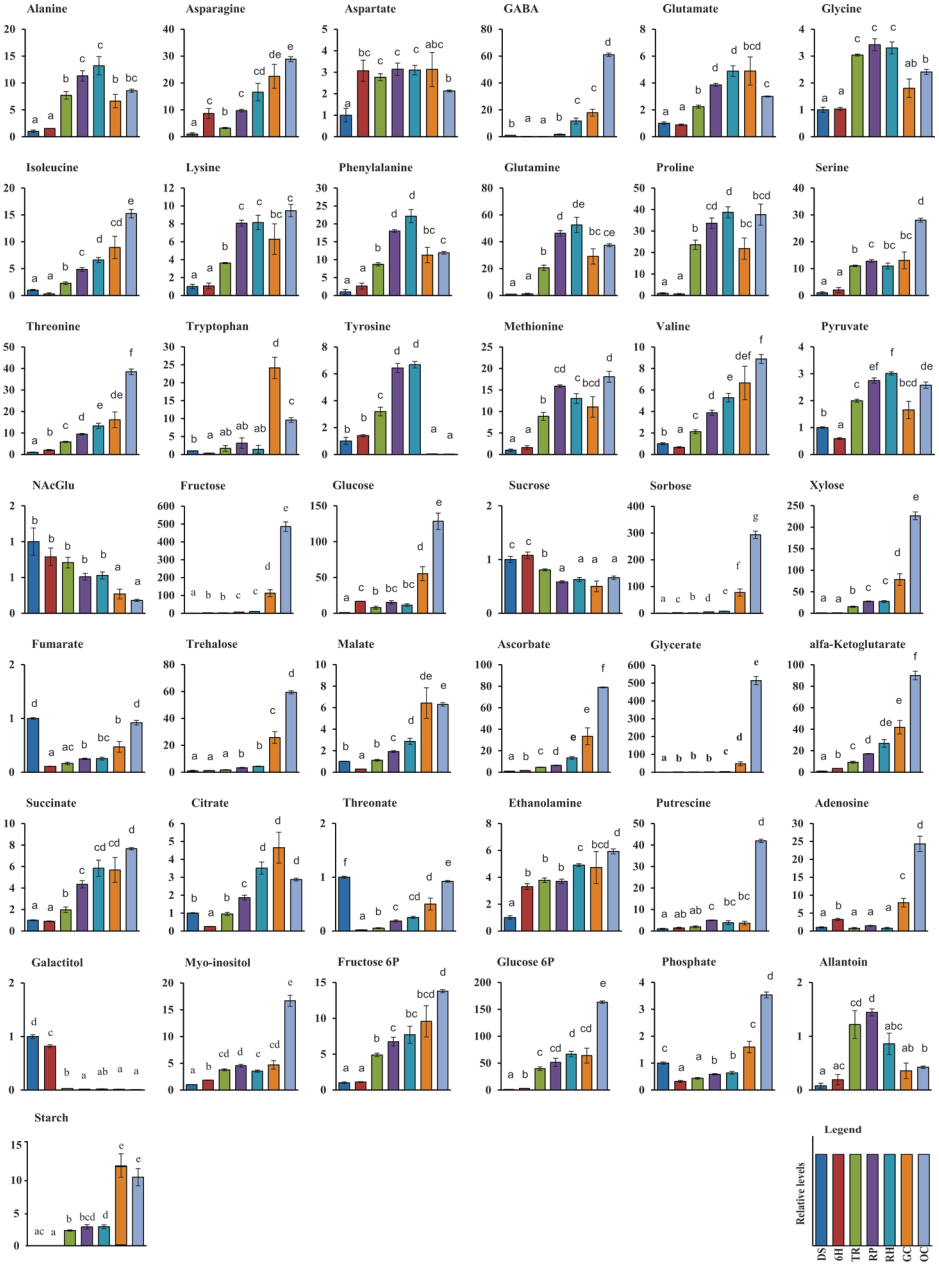


Figure 2. Relative concentration of 42 metabolites and starch which showed statistically significant variation during the seed-to-seedling transition. Different lower case letters above bars represent differences between samples by Tukey's HSD ($p < 0.05$). Colours of bars refer to developmental stages. From left to right: DS – dry seed; 6H – six hours imbibed; TR – testa rupture; RP – radicle protrusion; RH – root hair; GC – greening cotyledons; and OC – cotyledons fully opened.

Overall changes in intermediates of the tricarboxylic acid cycle (TCA). Intermediates of the TCA-cycle such as citrate, fumarate and malate showed levels 3- to 10-fold greater in DS than at 6H (Figure 2). In contrast, levels of α-ketoglutarate increased 3-fold at 6H, as compared to DS. In addition, all identified TCA cycle intermediates showed a constant increase from TR to OC. This suggests the presence of variation in mitochondrial function, including regeneration or *de novo* synthesis of mitochondria, starting between 6H and TR and continuing to increase during the later seed-to-seedling developmental stages.

Overall changes in amino acid contents. Eighteen amino acids were identified of which four changed during the first six hours of imbibition. An increase was detected for asparagine (8-fold), aspartate (3-fold) and threonine (2-fold) from DS to 6H, whereas pyruvate content at 6H was 50% lower than at DS (Figure 2). However, most of the amino acids (15) increased at TR or later stages. Three of these 15 displayed a considerable boost in their levels at TR when compared to 6H; isoleucine (8-fold), glutamine (16-fold) and proline (32-fold) (Figure 2). The shift from the last stage of germination (RP) to the first stage of seedling establishment (RH) was accompanied by an increase in gamma-aminobutyric acid (GABA) levels of around 6-fold. When photosynthesis became active (GC), levels of tryptophan increased vastly to around 17-fold compared with RH. Three other amino acids did not display a strong increase during the RH to GC transition, but increased around 2-fold (serine and threonine) and 3-fold (GABA) in OC compared with GC. Whereas all amino acids attained their highest levels at GC and OC stages, tyrosine was a notable exception as it was 149-fold higher at RP and RH when compared to GC and OC (Figure 2).

This consistent increase in amino acids, sugars, starch and organic acids indicates that central metabolic processes are being maintained or induced in preparation for seedling establishment. In addition, it should be noted that metabolite profiling, PCA and the biplot results suggest that the seed-to-seedling transition developmental stages are represented by three clusters according to the major changes in metabolite levels: germination (DS and 6H); early seedling growth (TR, RP and RH) and seedling establishment (GC and OC).

Metabolic correlation network analysis

We next explored the metabolic shifts between the heterotrophic and photoautotrophic states. We searched for groups of metabolites that display a similar concentration pattern across the different developmental stages by using the *Fuzzy K-Means* clustering method (Belostotsky *et al.*, 2009) with some modifications (*Materials and Methods*). We identified two dominant metabolite level profiles (Figure 3).

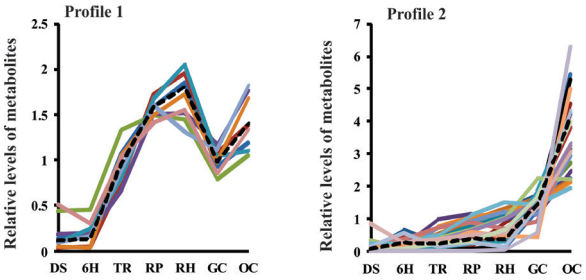


Figure 3. Two profiles of metabolite levels at developmental stages of the the seed-to-seedling transition. Profiles were identified using *Fuzzy K-means* clustering. The black line depicts the relative median of metabolite levels in each profile. Developmental stages are represented by: DS – dry seed; 6H – six hours imbibed; TR – testa rupture; RP – radicle protrusion; RH – root hair; GC – greening cotyledons; and OC – cotyledons fully opened.

Profile 1 included seven amino acids (alanine, glutamine, glycine, lysine, phenylalanine, proline, and methionine) and pyruvate, whereas Profile 2 was comprised of six amino acids (asparagine, GABA, isoleucine, serine, threonine, and valine), five carbohydrates (fructose, glucose, sorbose, xylose and trehalose), five organic acids (malate, α-ketoglutarate, succinate, ascorbate and glycerate), three carbohydrate derivatives (myo-inositol, fructose-6-phosphate and glucose-6-phosphate), adenosine, putrescine and phosphate (Figure 3). Profile 1 represents metabolites that did not change much until 6H but increased steadily up to a maximum at the RH stage, followed by a slight decline to the GC stage and a subsequent increase to OC. Profile 2 represents metabolite levels that gradually increase across the seed-to-seedling developmental stages and reached the highest level at OC. The three metabolites (sucrose, NacGlu and galactinol) that were not included into these two dominant profiles displayed an opposite trend to Profile 2, that is they gradually decreased during development. The difference in numbers of metabolites between the patterns likely reflects variations in biochemical pathways during seedling development.

To characterize the profiles, we built a metabolite-metabolite correlation network for each profile (Figure 4). Each edge of the networks represents a source-target (outgoing-incoming edges) correlation. A high number of incoming edges into a metabolite indicates the dependence of that metabolite on the outgoing metabolite edge (Xue *et al.*, 2013).

In the profile 1, amino acids showed a high correlation with pyruvate; high numbers of incoming edges were present for pyruvate, whereas alanine showed only outgoing edges (Figure 4). This profile 1, which shows an increase in levels of seven amino acids and of pyruvate at TR, seems to represent key metabolites for an initial

energy boost for seed germination. Profile 2 showed a high number of incoming edges for carbohydrates, organic acids, myo-inositol, putrescine and phosphate, whereas the amino acids showed a low number of incoming edges. Six amino acids present in profile 2, besides a few incoming edges, showed high numbers of outgoing edges with asparagine only showing outgoing edges (Figure 4). In profile 2, metabolites such as carbohydrates and TCA cycle intermediates increased across all stages. Thus, increased levels of metabolites in profile 2, especially those for TCA cycle intermediates, may represent a cluster of metabolites required for successful seed germination as well as seedling establishment.

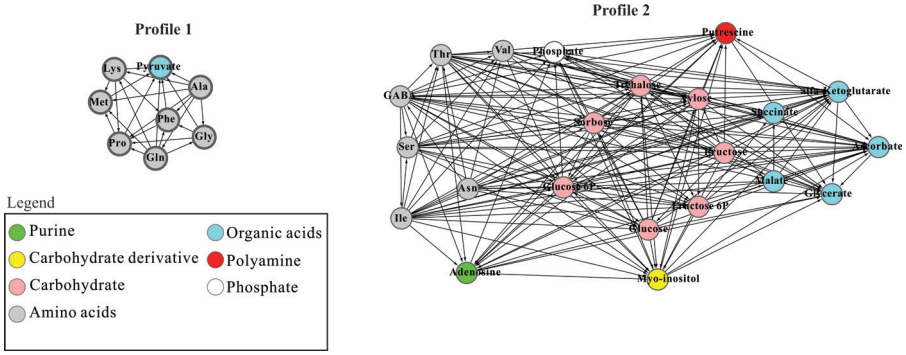


Figure 4. Metabolite profile networks, in which metabolites display significantly higher connectivity to each other. Each edge in the networks indicates a source-target correlation.

To delineate the biochemical pathways of each profile, we performed biochemical pathway enrichment analysis using the MetaboAnalyst3.0 website tool (Table 1).

Profile 1 was significantly ($p < 0.01$) enriched for biochemical pathways such as *aminoacyl-tRNA biosynthesis*, *nitrogen metabolism*, *alanine, asparagine and glutamate metabolism*, and *carbon fixation in photosynthetic organisms*. Profile 2 was enriched for *alanine, aspartate and glutamate metabolism*, *TCA cycle*, *aminoacyl-tRNA biosynthesis*, *galactose metabolism*, and *valine, leucine and isoleucine biosynthesis* (Table 1). These dominant profiles and their pathway enrichments suggest that seed germination and seedling establishment are associated with initial activation of important metabolic processes (profile 2). It is also likely that seed germination is associated with a follow up of additional metabolites (profile 1), which appear later during germination.

Table1. Significantly enriched pathways for profiles 1 and 2.

Pathway Name	Total	Hits	*P**	-log(p)	Holm p	FDR
Profile 1						
Aminoacyl-tRNA biosynthesis	67	7	6E-09	2E+01	5E-07	5E-07
Nitrogen metabolism	15	3	7E-05	1E+01	6E-03	3E-03
Alanine, aspartate and glutamate metabolism	22	3	2E-04	8E+00	2E-02	7E-03
Carbon fixation in photosynthetic organisms	21	2	7E-03	5E+00	6E-01	1E-01
Profile 2						
Alanine, aspartate and glutamate metabolism	22	4	4E-04	8E+00	3E-02	3E-02
Citrate cycle (TCA cycle)	20	3	4E-03	5E+00	4E-01	1E-01
Aminoacyl-tRNA biosynthesis	67	5	4E-03	5E+00	4E-01	1E-01
Galactose metabolism	26	3	9E-03	5E+00	7E-01	2E-01
Valine, leucine and isoleucine biosynthesis	26	3	9E-03	5E+00	7E-01	2E-01

Summary of the metabolites in each profile ranked by their *P* values. **Total** is the total number of compounds in the pathway; **Hits** is the actually matched number from each profile; **Holm p** is the *P* value adjusted by Holm-Bonferroni method and **FDR** is the *P* value adjusted using the False Discovery Rate adjustment method. ***** Significant pathways with $p<0.01$

Comprehensive metabolic pathways

To understand the regulation of the biochemical pathways that operate during the seed-to-seedling transition, we correlated the metabolite changes to transcriptional changes. Hereto, we used the transcriptome data of the same seven successive developmental stages that are described in Chapter 2. This transcriptome dataset provides a comprehensive description of gene expression during the seed-to-seedling transition. We mapped 146 representative transcripts associated with energy and amino acid metabolism on the seven stages (Figure 5). Displaying the transcripts and metabolites in this way revealed that changes in transcript abundance preluded the later changes in metabolite contents. Importantly, 42 out of 146 transcripts were involved in the metabolism of eight carbohydrates and their derivatives (sucrose, glucose, fructose, fructose 6-phosphate, glucose 6-phosphate, threhalose, xylose, myo-inositol and galactinol), whereas 34 were involved in the metabolism of six organic acids (pyruvate, citrate, α-ketoglutarate, succinate, fumarate and malate). Furthermore, metabolic pathways involving 17 amino acids (glycine, serine, phenylalanine, tyrosine, tryptophan, alanine, valine, asparagine, aspartate, lysine, threonine, isoleucine, methionine, glutamate, GABA, glutamine and NAcGlu) were associated with 72 transcripts. To facilitate the identification of patterns of transcriptional changes in pathways associated with metabolite levels, we determined

the Pearson correlation between the levels of identified metabolites and transcripts of the corresponding metabolic pathway enzymes (Dataset S3). We focussed on the changes in gene expression and metabolite levels associated with carbohydrate, organic acid, and amino acid metabolism (Figure 5). Fifteen transcripts associated with carbohydrate metabolism were positively correlated with the corresponding metabolite levels (Figure 5A). Sucrose levels were best correlated with transcript abundance of *INVERTASE B* (*A/N-INVB*), which encodes an enzyme involved in the hydrolysis of sucrose. Expression of *SUCROSE SYNTHASE 1* (*SUS1*) and *SUCROSE SYNTHASE 2* (*SUS2*) were best correlated with fructose and sucrose, respectively (Figure 5A). Sucrose synthase catalyzes the conversion of glucose and fructose to sucrose and vice versa. Glucose is also associated with a biochemical reaction involving hexokinase, in which hexokinase phosphorylates glucose to glucose-6-phosphate. The expression of two isoforms of hexokinase, *HEXOKINASE 1* (*HXK1*) and *HEXOKINASE 3* (*HXK3*) was best correlated with glucose-6-phosphate levels (Figure 5A). Glucose-6-phosphate, besides the above biochemical reactions, is also involved in an important step of glycolysis, the conversion of glucose-6-phosphate to fructose-6-phosphate. This conversion is mediated by phosphoglucose isomerase and, indeed, *PHOSPHOGLUCOSE ISOMERASE 1* (*PGI1*) was well correlated with glucose-6-phosphate content (Figure 5A). Another important biochemical reaction that incorporates fructose-6-phosphate into glycolysis involves fructokinases which indirectly convert fructose to fructose-6-phosphate. The expression of *PHOSPHOFRUCTOKINASE 1* (*PFK1*) and *PHOSPHOFRUCTOKINASE 5* (*PFK5*) correlated best with fructose content (Figure 5A). Glucose-6-phosphate is also a precursor for the biosynthesis of trehalose. The expression of four genes encoding trehalose-phosphate synthases was best correlated with trehalose and glucose-6-phosphate; the expression of *TREHALOSE-6-PHOSPHATE PHOSPHATASE E* (*TPPE*) and *TREHALOSE-6-PHOSPHATE PHOSPHATASE I* (*TPPI*) had the highest correlation with trehalose. Glucose-6-phosphate, however, was best correlating with *TREHALOSE-6-PHOSPHATASE SYNTHASE S5* (*TPS5*) and *TREHALOSE-6-PHOSPHATE PHOSPHATASE A* (*TPPA*) transcript levels. Xylose and galactinol levels showed good correlation with transcript levels of enzymes associated with their biosynthesis. Xylose level correlated well with the expression of *UDP-D-XYLOSE SYNTHASE 2* (*AXS2*) and *UDP-XYLOSE SYNTHASE 4* (*UXS4*). Galactinol levels correlated well with *GALACTINOL SYNTHASE 1* (*GOLS1*) expression (Figure 5A). The number of transcripts correlating with organic acid metabolism was lower than for carbohydrate metabolism. Four transcripts were positively correlated with the three organic acids, pyruvate, succinate and α-ketoglutarate. Pyruvate is involved in two important pathways, glycolysis and the TCA cycle

(Figure 5) . Levels of pyruvate correlated well with the expression of *PYRUVATE KINASE* (At3g52990) (Figure 5A) and *PYRUVATE DEHYDROGENASE E1A-LIKE SUBUNIT* (*IAR4*) (Figure 5B). Pyruvate kinase transfers the phosphate group from phosphoenolpyruvate (PEP) to ADP, resulting in pyruvate, after which pyruvate dehydrogenase transforms pyruvate into acetyl-CoA, which is incorporated into the TCA cycle. The expression of the gene encoding isocitrate dehydrogenase (IDH) correlated well with α-ketoglutarate levels. This enzyme catalyzes the conversion of isocitrate into α-ketoglutarate. The expression of *CYTOSOLIC NADP*$^{+}$*-DEPENDENT ISOCITRATE DEHYDROGENASE* (*CICDH*) correlates with α-ketoglutarate levels. Succinate correlated well with transcript levels of one isoform of succinate dehydrogenase, which is an enzyme that catalyses the conversion of succinate to fumarate. *SUCCINATE DEHYDROGENASE 1-2* (*SDH1-2*) seems to be the isoform that catalyses the conversion of isocitrate to α-ketoglutarate (Figure 5B).

Abundance of 13 transcripts associated with amino acid metabolism was positively correlated with levels of eight amino acids. Aspartate correlated well with the expression of *ASPARTATE AMINOTRANSFERASE 2* (*AAT2*) (Figure 5B). Since aspartate is a precursor of asparagine through asparagine synthetase, we checked the correlation of isoforms of this enzyme, but none of the transcript levels correlated well with aspartate. However, *ASPARAGINE SYNTHETASE 2* (*ASN2*) and *ASPARAGINE SYNTHETASE 3* (*ASN3*) transcript levels did correlate with asparagine content. Another amino acid, methionine, was found to correlate very well with the transcript level of three isoforms of methionine synthetase, *METHIONINE SYNTHASE 1* (*MS1*), *METHIONINE SYNTHASE 2* (*MS2*) and *HOMOCYSTEINE S-METHYLTRANSFERASE* (*HMT-1*). Glutamate decarboxylase (GAD) is the enzyme responsible for decarboxylation of glutamate to GABA. The expression of two *GAD* isoforms identified during the seed-to-seedling transition correlated well with glutamate and GABA content. *GLUTAMATE DECARBOXYLASE 1* (*GAD1*) correlated with glutamate and GABA, whereas *GLUTAMATE DECARBOXYLASE 2* (*GAD2*) showed correlation only with glutamate. Glutamate is converted to N-acetylglutamate, a reaction catalysed by N-acetylglutamate synthase; however, glutamate did not correlate appreciably with isoforms of this enzyme. N-acetylglutamate levels did correlate with *N-ACETYL-L-GLUTAMATE SYNTHASE 1* (*NAGS1*) (Figure 5B). Glutamate is also precursor in the formation of glutamine, a reaction catalysed by glutamine synthetase. Glutamate, as well as glutamine, showed very good correlation with transcript levels of two isoforms of glutamine synthetase, *GLUTAMINE SYNTHASE 1;1* (*GLN1;1*) and *GLUTAMINE SYNTHETASE 2* (*GS2*), suggesting that these two isoforms play a role in the conversion of glutamate into glutamine during germination and early seedling establishment. Phenylalanine

correlated well with *AROGENATE DEHYDRATASE 1* (*ADT1*), which is involved in phenylalanine biosynthesis. Of the three isoforms of genes involved in serine biosynthesis only the expression of *PHOSPHOSERINE AMINOTRANSFERASE 2* (*PSAT2*) correlated with the serine level (Figure 5B). Taken together, each correlation between gene expression and metabolite content involved in the metabolism of primary metabolites associated with key genes highlights their potential requirement for the seed-to-seedling transition in Arabidopsis.

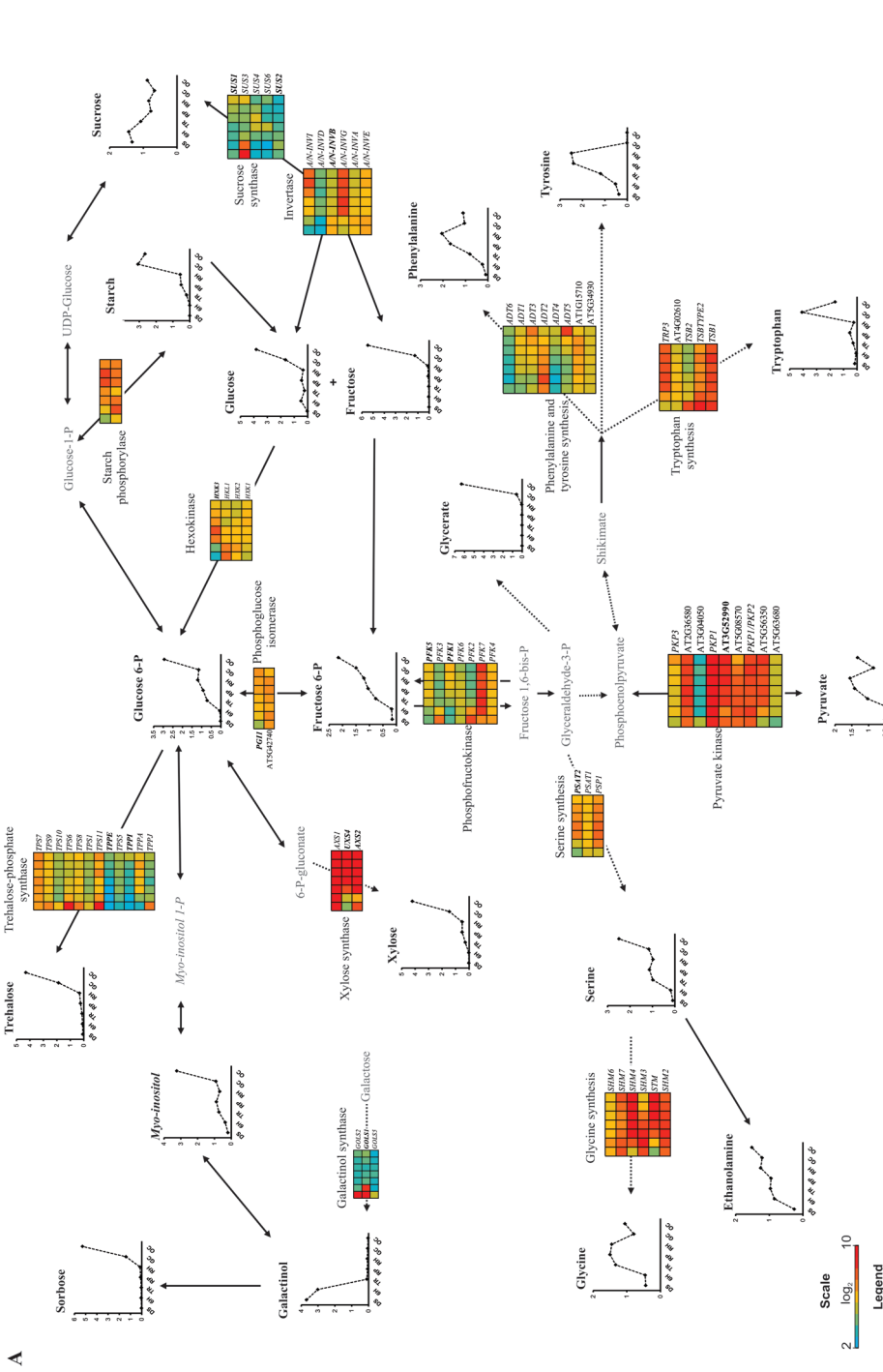


Figure 5. Continued

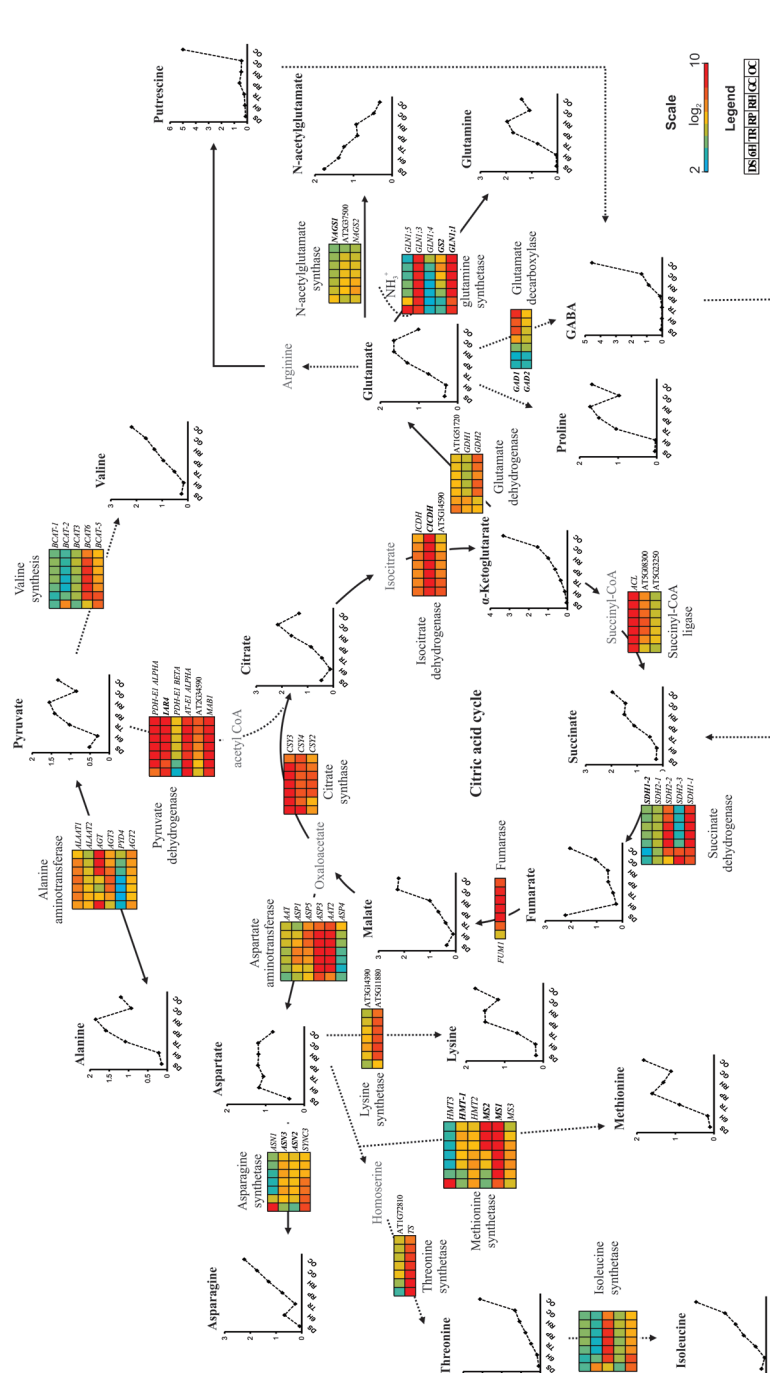


Figure 5. Comprehensive primary metabolic map of the seed-to-seedling transition in Arabidopsis for (**A**) carbohydrate, (**B**) TCA and (**A** and **B**) amino acid metabolism. On the metabolic map 146 transcripts and 36 metabolites are mapped to three main metabolic groups of compounds: carbohydrates, organic acids and amino acids. Metabolite data are means of three replicates using 350 seeds for each replicate. According to previous annotations, expression profiles of transcripts encoding enzymes involved in the metabolite productions are represented. Genes indicated in boldface are highly correlated (r>0.85) to their respective metabolites.

Discussion

This study focused on metabolic changes during the seed-to-seedling transition in *Arabidopsis thaliana*. To determine how the advancement of growth is associated with certain metabolite profiles. This transition involves several developmental stages ranging from DS (mature dry seed) to OC (opening of cotyledons, i.e. the last stage of early seedling establishment). The seed-to-seedling transition is a complex process that has not received much attention in terms of metabolite profiling. In this study we performed a detailed metabolite profiling analysis, using morphological markers of development. We precisely selected individual seeds and seedlings at seven developmental stages.

Radicle protrusion (RP) was considered the morphological marker defining the transition between completion of germination and commencement of seedling development (Bewley et al, 2013). Three stages prior to RP and three stages thereafter were selected for metabolite profiling. A previous study employing metabolite profiling in developing seedlings was targeted at heterosis of two Arabidopsis genotypes (Col-0 and C24) (Meyer *et al.*, 2012). Moreover, in that study samples were selected based on temporal markers rather than morphological markers. For example, samples collected 48h after imbibition were analysed, but these samples consisted of a mixture of three developmental stages: TR, RP, and RH because of biological variation in the germination speed of individual seeds of the population (Meyer *et al.*, 2012). Despite the fact that metabolite profiling has been studied in seed germination (Fait *et al.*, 2006; Shu *et al.*, 2008; Angelovici *et al.*, 2011; Joosen *et al.*, 2013), many questions regarding the regulation of the seed-to-seedling transition remain open, due to the lack of a comprehensive analysis of changes in metabolism by combining expression analysis and metabolite levels. We compared metabolite profiles of seven seed-to-seedling developmental stages and revealed significant variation among them, not only in metabolite levels but also in metabolite-metabolite network correlations. In combination with transcriptomics data, we constructed a metabolic map. Thus, this study presents a broader and more precise metabolic analysis as compared to previous studies (Fait *et al.*, 2006; Meyer *et al.*, 2012).

Metabolic changes during the seed-to-seedling transition were first visualized in a PCA plot. This resulted in the observation of major shifts in metabolite profiles between three phases of this transition: (1) initial stages of germination (DS and 6h); (2) stages of early seedling growth (TR, RP and RH), and (3) the final stages of seedling development (GC and OC) in which the seedling becomes photoautotrophic. Metabolic shifts are represented by the increasing distance between these different

phases, i.e. galactinol, sucrose and NAcGlu, show markedly higher levels at DS and 6H, which, thus, separates them from other stages. In the early-seedling phase tyrosine, allantoin and urea are prominent metabolites, suggesting a role at this phase. The most significant metabolites that contribute to separation of the last phase are amino acids.

The detected carbohydrates and their derivatives such as glucose, fructose, sorbose, xylose, trehalose, fructose-6-phosphate and glucose-6-phosphate displayed high levels at OC, which implies that the carbon status was relatively high at the seedling stage, as compared to the previous stages. Additionally, amino acids with high contents at the GC and OC stages, such as glutamate, glutamine, asparagine and aspartate were variable among the developmental stages. It indicates a high variability of nitrogen status, despite the fact that these amino acids are known to be important for monitoring the C/N balance in plants (Zheng, 2009). Since amino acids are important forms of N storage for seeds (Lohaus and Moellers, 2000; Sanders *et al.*, 2009), our results suggest that, at the metabolomic level, N status in amino acids represented by profile 1 plays an important role in balancing C and N in the transition of the initial phase of germination to the early-seedling stage. Furthermore, these amino acids show outgoing edges to pyruvate within the metabolite-metabolite network. It is known that aminotransferase performs a reversible reaction converting these amino acids into pyruvate and vice versa (Orzechowski *et al.*, 1999; Pinto *et al.*, 2014; McAllister and Good, 2015). Additionally, amino acids present in profile 1 participate substantially in nitrogen transport (Orzechowski *et al.*, 1999; Yoo *et al.*, 2013; Pinto *et al.*, 2014; McAllister and Good, 2015).

Different from profile 1, profile 2 shows a constant increase in metabolites with highest levels at the final stages of seedling establishment. This pattern represents the metabolic shift from the early-seedling stage to the last phase of seedling development. Furthermore, the metabolic network of profile 2 shows a high number of incoming edges for glycolytic and TCA intermediates, which confirms the strong relationship between regulation of glycolysis and the TCA cycle. Amino acids present in profile 2, such as GABA, isoleucine, aspartate, serine, threonine and valine may also influence the C/N balance. For example, GABA is associated with succinate in the TCA cycle, which may provide high levels of succinate during the seed-to-seedling transition. This could be related to the established role of GABA in regulating the C/N ratio when carbon supply is limited (Michaeli *et al.*, 2011). Besides carbohydrates, amino acids and organic acids, starch also displayed a steady increase across the seed-to-seedling transition. This result corroborates a previous study in Arabidopsis (Matsoukas *et al.*, 2013) in which starch metabolism was associated with the juvenile-to-adult plant phase transition during normal growth and

development. It was suggested that plants in the juvenile phase may require starch accumulation to reach a threshold level in order to sustain a steady supply of maltose and/or sucrose during the juvenile-to-adult phase transition (Matsoukas *et al.*, 2013). This observation may, thus, extend to the seed-to-seedling phase transition, where sucrose and starch levels showed an opposite trend indicating that starch may be produced from sucrose degradation.

In addition to providing potential targets for the engineering of seed-to-seedling transition metabolism, this study also allows a general assessment of transcriptional regulation during this transition. Previous studies have suggested that regulation of metabolism occurs at the post-translational level (Jiao and Chollet, 1991; Kolbe *et al.*, 2006; Lea *et al.*, 2006; Bates *et al.*, 2014; Pisithkul *et al.*, 2015). However, we identified several linear correlations between gene expression and metabolite levels. For example, high correlations of carbohydrates and their derivatives with transcripts associated with glycolysis, suggest that metabolism is also regulated transcriptionally. For example, sucrose levels correlated well with transcript levels of sucrose synthase (*SUCROSE SYNTHASE* 2 - *SUS2*) and invertase (*INVERTASE B* - *INVB*). Several studies have shown the importance of these enzymes for normal development (Cheng *et al.*, 1996; Yau and Simon, 2003; Barratt *et al.*, 2009). Our results suggest that *INVB* and *SUS2* may be the key enzymes essential for sucrose metabolism in support of a normal seedling establishment. This corroborates previous reports which demonstrated an important role for *INV1* in root cell development and reproduction in rice (Jia *et al.*, 2008) and for whole plant development in *Lotus japonicus* (Welham *et al.*, 2009). It is not clear whether *INVB* and *SUS2* have regulatory properties that allow flux of C out of sucrose via this pathway in coordination with energy demands by the cells, or whether *SUCROSE SYNTHASE1* (*SUS1*), which is well correlated with fructose content, could compensate for the lack of sucrose to maintain fuelling of energy-demanding processes. Two isoforms of hexokinase (*HEXOKINASE 1* - *HXK1* and *HEXOKINASE 3* - *HXK3*) correlate well with glucose-6-phosphate levels. Hexokinase is an enzyme in the glycolysis pathway and its main substrate is glucose, which also has a function in the control of plant development and expression of different classes of genes (Renz and Stitt, 1993; Dai *et al.*, 1999; Claeyssen and Rivoal, 2007). A previous study has shown that the C flux through hexokinase activity exhibits a high control over glucose during normal root growth (Claeyssen *et al.*, 2013). This suggests that the *HXK1* and *HXK3* encoded enzymes control a step in glycolysis during the seed-to-seedling transition. Besides the hexokinase reaction (phosphorylation of glucose to form glucose-6-phospate), glucose-6-phosphate is precursor for the formation of fructose-6-phosphate through phosphoglucose isomerase (*PGI*). *PGI1* transcript abundance

was well correlated with glucose-6-phosphate levels. It has been shown that *PGI1* plays a role in the transition to flowering in Arabidopsis. The *pgi1-1 mutant* flowers earlier than wild type (Yu *et al.*, 2000). However, the phenotype could be reverted to the wild type by the addition of sugars (glucose, fructose and sucrose), suggesting that it is an important enzyme for a developmental phase transition. Our results show that *PGI1* may be a key enzyme for the supply of C to glycolysis through glucose-6-phosphate during seed-to-seedling transition. The second phosphorylation step in glycolysis is the conversion of fructose-6-phosphate to fructose 1,6-biphosphate via phosphofructokinase (Hellinga and Evans, 1987; Mustroph *et al.*, 2013). This reaction is catalysed by an ATP-dependent phosphofructokinase (PFK). Our results show that *PFK1* and *PFK5* transcript abundance correlates well with fructose-6-phosphate content. This is particularly interesting given the fact that *PFK* isoforms have been characterized for a role in rice during oxygen depletion (Mustroph *et al.*, 2013). It was suggested that *PFK* regulates candidate rice genes for adaptation to anoxic stress. Taken our results into consideration, this suggests that *PKF1* and *PKF5* can be readily utilized as a means of identifying candidate genes for biotechnological modification of the seed-to-seedling transition.

Organic acid levels also showed good correlations with transcripts associated with their biosynthesis. The last step of glycolysis is the conversion of phosphoenolpyruvate to pyruvate via pyruvate kinase (*PKP*) towards the TCA cycle. Our results show that expression of a putative pyruvate kinase (At3g52990) correlates well with pyruvate levels. Several studies have shown the importance of pyruvate kinase (*PKP)* in seed germination (Baud *et al.*, 2007) and in seedling establishment (Andre and Benning, 2007) in Arabidopsis. These studies have shown that *pkp* mutants exhibit delayed germination (Baud *et al.*, 2007) and that *PKP* plays an important role in catabolizing storage compounds in germinating Arabidopsis seeds (Andre and Benning, 2007). The correlation of At3g52990 with pyruvate levels, suggests that also this putative pyruvate kinase isoform plays a role during seedling establishment, although this gene has not been described before in this particular context. Pyruvate is also related to another key reaction via pyruvate dehydrogenase, an enzyme that catalyzes the oxidative decarboxylation of pyruvate, yielding CO_2, acetyl-CoA and NADH (Reed, 1974). Interestingly, our results show pyruvate levels correlate well with expression of the pyruvate dehydrogenase *E1a-like* (*IAR4*), which is 81% identical to a previously characterized Arabidopsis mitochondrial PDH E1α-subunit (*AT-E1 ALPHA* - At1g59900) (Quint *et al.*, 2009). However, *AT-E1 ALPHA* transcript abundance showed a negative correlation with pyruvate levels. The importance of pyruvate dehydrogenase has been demonstrated clearly by a mutation in the E2 subunit, which reduced plant organ size and increased accumulation of

pyruvate (Yu *et al.*, 2012). Our results suggest that *IAR4* is the gene controlling the entry of C, through acetyl-CoA, into the TCA cycle for energy production during the seed-to-seedling transition. In the TCA cycle, *CYTOSOLIC NADP+-DEPENDENT ISOCITRATE DEHYDROGENASE* (*CICDH*) is well correlated with citrate levels and this corroborates the observation that citrate accumulated in three independent knock-out mutants (*icdh-1*, *icdh-2*, and *icdh*-3) of Arabidopsis under normal growth conditions (Mhamdi *et al.*, 2010).

Our results also show that amino acids such as aspartate, methionine, glutamate, NAcGlu, glutamine, phenylalanine, and serine correlate well and positively with transcripts related to their biosynthesis. Amino acids are the major transport form of N in plants (Tegeder and Ward, 2012; Tegeder, 2014) and growth is dependent on N supply, assimilation and utilization (Stitt and Krapp, 1999). Aspartate, glutamate and glutamine are involved in transamination processes and N assimilation (Zheng, 2009). These amino acids show a continuous increase in content during the seed-to-seedling transition and correlated with transcripts associated with their biosynthesis, such as asparagine synthetase (*ASN2* and *ASN3*), glutamate decarboxylase (*GAD1* and *GAD2*) and glutamine synthetase (*GLN1;1* and *GS2*), which suggests that there is N assimilation. N assimilation may lead to the synthesis of other amino acids, such as lysine, phenylalanine, valine and methionine at the final stages of seedling establishment, when growth rate is relatively high.

Taken together, our results enable us to draw important conclusions concerning metabolism during the seed-to-seedling transition in Arabidopsis. We demonstrate that primary metabolism is co-ordinately regulated during the seed-to-seedling transition, displaying two major shifts that separate three groups of developmental stages (Figure 6). The first metabolic shift from the initial stages of germination (DS and 6H) to the early-seedling stages (TR, RP and RH) is represented by profile 1, in which seven amino acids and pyruvate are present. The second metabolic shift occurs between early-seedling and the last stages of seedling establishment and is reflected by profile 2 in which amino acids, organic acids, carbohydrates and their derivatives are present. We also demonstrated that transcriptional changes not always equate to a proportional metabolic response. Detailed correlation analysis, however, enabled us to identify transcripts that do seem to directly influence the flux through metabolic pathways during the seed-to-seedling transition. Our study provides a detailed overview of the transcriptional and metabolic changes during the seed-to-seedling transition in Arabidopsis and opens new perspectives for understanding the complex regulatory mechanism underlying this transition.

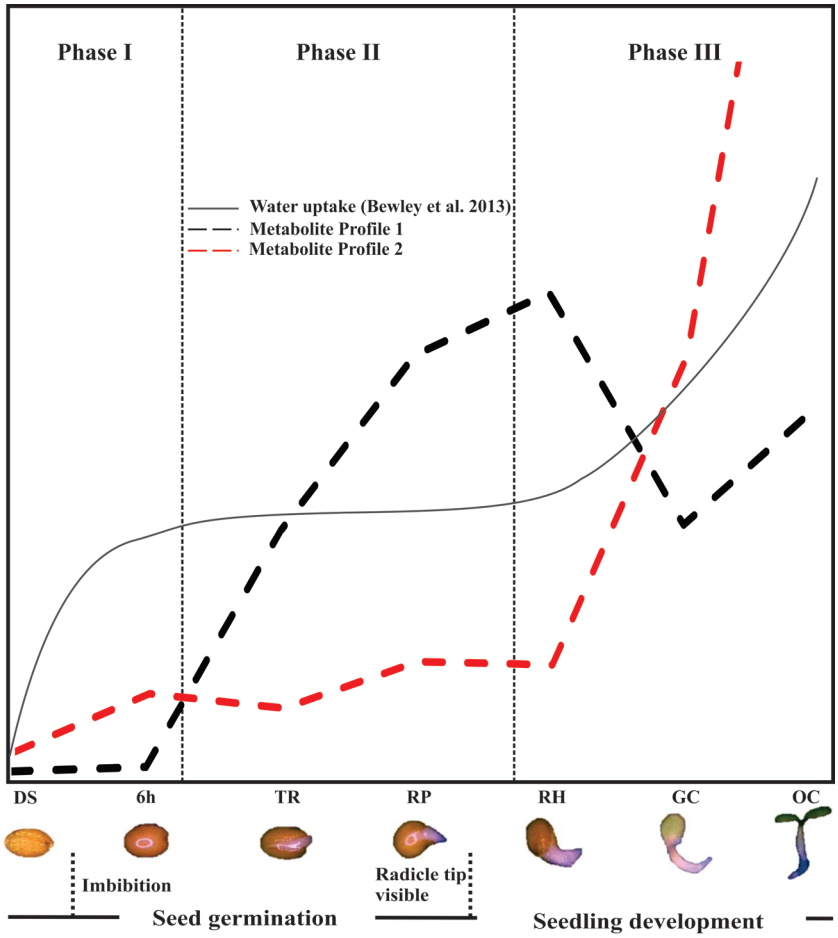


Figure 6. Metabolic shifts during the seed-to-seedling transition in Arabidopsis. The transition from Phase I to Phase II is characterized by levels of metabolites present in profile 1. Immediately preceding Phase III, seedling establishment is prepared with a boost of metabolites present in profile 2. Dashed lines distinguish metabolite profile 1 (black) and metabolite profile 2 (red).

Acknowledgements

This work was funded by "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior" (CAPES, Brazil).

References

Andre C, Benning C.2007 Arabidopsis Seedlings Deficient in a Plastidic Pyruvate Kinase Are Unable to Utilize Seed Storage Compounds for Germination and Establishment. *Plant Physiology* **145:** 1670-1680

Angelovici R, Fait A, Fernie AR, Galili G.2011 A seed high-lysine trait is negatively associated with the TCA cycle and slows down Arabidopsis seed germination. *New Phytologist* **189:** 148-159

Barratt DHP, Derbyshire P, Findlay K, Pike M, Wellner N, Lunn J, Feil R, Simpson C, Maule AJ, Smith AM.2009 Normal growth of Arabidopsis requires cytosolic invertase but not sucrose synthase. *Proceedings of the National Academy of Sciences* **106:** 13124-13129

Bates PD, Johnson SR, Cao X, Li J, Nam JW, Jaworski JG, Ohlrogge JB, Browse J.2014 Fatty acid synthesis is inhibited by inefficient utilization of unusual fatty acids for glycerolipid assembly. *Proceedings of the National Academy of Sciences of the United States of America* **111:** 1204-1209

Baud S, Wuilleme S, Dubreucq B, de Almeida A, Vuagnat C, Lepiniec L, Miquel M, Rochat C.2007 Function of plastidial pyruvate kinases in seeds of Arabidopsis thaliana. *Plant J* **52:** 405-419

Belostotsky DA, Orlando D, Brady S, Koch J, Dinneny J, Benfey P (2009) Manipulating Large-Scale Arabidopsis Microarray Expression Data: Identifying Dominant Expression Patterns and Biological Process Enrichment. *In* Plant Systems Biology, Vol 553. Humana Press, pp 57-77

Bewley JD, Bradford K, Hilhorst HM, Nonogaki H (2013) Germination. *In* Seeds. Springer New York, pp 133-181

Cañas RA, Canales J, Muñoz-Hernández C, Granados JM, Ávila C, García-Martín ML, Cánovas FM.2015 Understanding developmental and adaptive cues in pine through metabolite profiling and co-expression network analysis. *Journal of Experimental Botany*

Cheng WH, Taliercio EW, Chourey PS.1996 The Miniature1 Seed Locus of Maize Encodes a Cell Wall Invertase Required for Normal Development of Endosperm and Maternal Cells in the Pedicel. *The Plant Cell* **8:** 971-983

Claeyssen É, Dorion S, Clendenning A, He JZ, Wally O, Chen J, Auslender EL, Moisan M-C, Jolicoeur M, Rivoal J.2013 The Futile Cycling of Hexose Phosphates Could Account for the Fact That Hexokinase Exerts a High Control on Glucose Phosphorylation but Not on Glycolytic Rate in Transgenic Potato (Solanum tuberosum) Roots. *PLos ONE* **8:** e53898

Claeyssen E, Rivoal J.2007 Isozymes of plant hexokinase: occurrence, properties and functions. *Phytochemistry* **68:** 709-731

Dai N, Schaffer A, Petreikov M, Shahak Y, Giller Y, Ratner K, Levine A, Granot D.1999 Overexpression of Arabidopsis hexokinase in tomato plants inhibits growth, reduces

photosynthesis, and induces rapid senescence. *The Plant Cell* **11:** 1253-1266

Debeaujon I, Koornneef M.2000 Gibberellin Requirement for Arabidopsis Seed Germination Is Determined Both by Testa Characteristics and Embryonic Abscisic Acid. *Plant Physiology* **122:** 415-424

Dekkers BJW, Pearce S, van Bolderen-Veldkamp RP, Marshall A, Widera P, Gilbert J, Drost H-G, Bassel GW, Müller K, King JR, Wood ATA, Grosse I, Quint M, Krasnogor N, Leubner-Metzger G, Holdsworth MJ, Bentsink L.2013 Transcriptional Dynamics of Two Seed Compartments with Opposing Roles in Arabidopsis Seed Germination. *Plant Physiology*

Fait A, Angelovici R, Less H, Ohad I, Urbanczyk-Wochniak E, Fernie AR, Galili G.2006 Arabidopsis Seed Development and Germination Is Associated with Temporally Distinct Metabolic Switches. *Plant Physiology* **142:** 839-854

Focks N, Benning C.1998 wrinkled1: A novel, low-seed-oil mutant of Arabidopsis with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol* **118:** 91-101

Fukushima A, Kusano M, Redestig H, Arita M, Saito K.2011 Metabolomic correlation-network modules in Arabidopsis based on a graph-clustering approach. *BMC Systems Biology* **5:** 1

Fulda M, Schnurr J, Abbadi A, Heinz E, Browse J.2004 Peroxisomal Acyl-CoA Synthetase Activity Is Essential for Seedling Development in Arabidopsis thaliana. *The Plant Cell Online* **16:** 394-405

Gutiérrez RA, Stokes TL, Thum K, Xu X, Obertello M, Katari MS, Tanurdzic M, Dean A, Nero DC, McClung CR, Coruzzi GM.2008 Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. *Proceedings of the National Academy of Sciences* **105:** 4939-4944

Hellinga HW, Evans PR.1987 Mutations in the active site of Escherichia coli phosphofructokinase. **327:** 437-439

Holdsworth MJ, Bentsink L, Soppe WJJ.2008 Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytologist* **179:** 33-54

Jia L, Zhang B, Mao C, Li J, Wu Y, Wu P, Wu Z.2008 OsCYT-INV1 for alkaline/neutral invertase is involved in root cell development and reproductivity in rice (Oryza sativa L.). *Planta* **228:** 51-59

Jiao JA, Chollet R.1991 Posttranslational regulation of phosphoeno/pyruvate carboxylase in C4 and crassulacean acid metabolism plants. *Plant Physiology* **95:** 981-985

Joosen RVL, Arends D, Li Y, Willems LAJ, Keurentjes JJB, Ligterink W, Jansen RC, Hilhorst HWM.2013 Identifying Genotype-by-Environment Interactions in the Metabolism of Germinating Arabidopsis Seeds Using Generalized Genetical Genomics. *Plant Physiology* **162:** 553-566

Kolbe A, Oliver SN, Fernie AR, Stitt M, Van Dongen JT, Geigenberger P.2006 Combined transcript and metabolite profiling of Arabidopsis leaves reveals fundamental effects of

the thiol-disulfide status on plant metabolism. *Plant Physiology* **141:** 412-422

Lea US, Leydecker MT, Quilleré I, Meyer C, Lillo C.2006 Posttranslational regulation of nitrate reductase strongly affects the levels of free amino acids and nitrate, whereas transcriptional regulation has only minor influence. *Plant Physiology* **140:** 1085-1094

Lohaus G, Moellers C.2000 Phloem transport of amino acids in two Brassica napus L. genotypes and one B. carinata genotype in relation to their seed protein content. *Planta* **211:** 833-840

Lommen A.2009 MetAlign: Interface-Driven, Versatile Metabolomics Tool for Hyphenated Full-Scan Mass Spectrometry Data Preprocessing. *Analytical Chemistry* **81:** 3079-3086

Lu C, Hills MJ.2002 Arabidopsis mutants deficient in diacylglycerol acyltransferase display increased sensitivity to abscisic acid, sugars, and osmotic stress during germination and seedling development. *Plant Physiology* **129:** 1352-1358

Lu C, Hills MJ.2002 Arabidopsis mutants deficient in diacylglycerol acyltransferase display increased sensitivity to abscisic acid, sugars, and osmotic stress during germination and seedling development. *Plant Physiol* **129:** 1352-1358

Lv Q, Cheng R, Shi T.2014 Regulatory network rewiring for secondary metabolism in Arabidopsis thaliana under various conditions. *BMC Plant Biology* **14:** 180

Matsoukas IG, Massiah AJ, Thomas B.2013 Starch metabolism and antiflorigenic signals modulate the juvenile-to-adult phase transition in Arabidopsis. *Plant Cell Environ* **36:** 1802-1811

McAllister CH, Good AG.2015 Alanine aminotransferase variants conferring diverse NUE phenotypes in Arabidopsis thaliana. *PLos ONE* **10**

Mercke P, Kappers IF, Verstappen FWA, Vorst O, Dicke M, Bouwmeester HJ.2004 Combined Transcript and Metabolite Analysis Reveals Genes Involved in Spider Mite Induced Volatile Formation in Cucumber Plants. *Plant Physiology* **135:** 2012-2024

Meyer RC, Witucka-Wall H, Becher M, Blacha A, Boudichevskaia A, Dörmann P, Fiehn O, Friedel S, von Korff M, Lisec J, Melzer M, Repsilber D, Schmidt R, Scholz M, Selbig J, Willmitzer L, Altmann T.2012 Heterosis manifestation during early Arabidopsis seedling development is characterized by intermediate gene expression and enhanced metabolic activity in the hybrids. *The Plant Journal* **71:** 669-683

Mhamdi A, Mauve C, Gouia H, Saindrenan P, Hodges M, Noctor G.2010 Cytosolic NADP-dependent isocitrate dehydrogenase contributes to redox homeostasis and the regulation of pathogen responses in Arabidopsis leaves. *Plant, Cell & Environment* **33:** 1112-1123

Michaeli S, Fait A, Lagor K, Nunes-Nesi A, Grillich N, Yellin A, Bar D, Khan M, Fernie AR, Turano FJ, Fromm H.2011 A mitochondrial GABA permease connects the GABA shunt and the TCA cycle, and is essential for normal carbon metabolism. *Plant J* **67:** 485-498

Mustroph A, Stock J, Hess N, Aldous S, Dreilich A, Grimm B.2013 Characterization of the Phosphofructokinase Gene Family in Rice and Its Expression Under Oxygen Deficiency Stress. *Frontiers in Plant Science* **4:** 125

Nonogaki H, Bassel GW, Bewley JD.2010 Germination—Still a mystery. *Plant Science* **179:** 574-581

Orzechowski S, Socha-Hanc J, Paszkowski A.1999 Alanine aminotransferase and glycine aminotransferase from maize (Zea mays L.) leaves. *Acta Biochim Pol* **46:** 447-457

Penfield S, Graham S, Graham IA.2005 Storage reserve mobilization in germinating oilseeds: Arabidopsis as a model system. *Biochem Soc Trans* **33:** 380-383

Pinto JT, Krasnikov BF, Alcutt S, Jones ME, Dorai T, Villar MT, Artigues A, Li J, Cooper AJL.2014 Kynurenine aminotransferase III and glutamine transaminase L are identical enzymes that have cysteine S-conjugate β-lyase activity and can transaminate L-selenomethionine. *Journal of Biological Chemistry* **289:** 30950-30961

Pisithkul T, Patel NM, Amador-Noguez D.2015 Post-translational modifications as key regulators of bacterial metabolic fluxes. *Current Opinion in Microbiology* **24:** 29-37

Quint M, Barkawi LS, Fan K-T, Cohen JD, Gray WM.2009 Arabidopsis IAR4 Modulates Auxin Response by Regulating Auxin Homeostasis. *Plant Physiology* **150:** 748-758

Reed LJ.1974 Multienzyme complexes. *Accounts of Chemical Research* **7:** 40-46

Renz A, Stitt M.1993 Substrate specificity and product inhibition of different forms of fructokinases and hexokinases in developing potato tubers. *Planta* **190:** 166-175

Ribeiro PR, Fernandez LG, de Castro RD, Ligterink W, Hilhorst HWM.2014 Physiological and biochemical responses of Ricinus communis seedlings to different temperatures: A metabolomics approach. **14**

Sanders A, Collier R, Trethewy A, Gould G, Sieker R, Tegeder M.2009 AAP1 regulates import of amino acids into developing Arabidopsis embryos. *Plant J* **59:** 540-552

Schopfer P, Plachy C.1984 Control of Seed Germination by Abscisic Acid : II. Effect on Embryo Water Uptake in Brassica napus L. *Plant Physiology* **76:** 155-160

Shu X-L, Frank T, Shu Q-Y, Engel K-H.2008 Metabolite Profiling of Germinating Rice Seeds. *Journal of Agricultural and Food Chemistry* **56:** 11612-11620

Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T.2011 Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27:** 431-432

Stitt M, Krapp A.1999 The interaction between elevated carbon dioxide and nitrogen nutrition: the physiological and molecular background. *Plant, Cell & Environment* **22:** 583-621

Sumner L, Amberg A, Barrett D, Beale M, Beger R, Daykin C, Fan TM, Fiehn O, Goodacre R, Griffin J, Hankemeier T, Hardy N, Harnly J, Higashi R, Kopka J, Lane A, Lindon J, Marriott P, Nicholls A, Reily M, Thaden J, Viant M.2007 Proposed minimum reporting standards for chemical analysis. **3:** 211-221

Tegeder M.2014 Transporters involved in source to sink partitioning of amino acids and ureides: opportunities for crop improvement. *Journal of Experimental Botany*

Tegeder M, Ward JM.2012 Molecular Evolution of Plant AAP and LHT Amino Acid Transporters. *Front Plant Sci* **3:** 21

Tikunov YM, Laptenok S, Hall RD, Bovy A, de Vos RCH.2012 MSClust: a tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise

aligned data. **8:** 714-718

Toubiana D, Semel Y, Tohge T, Beleggia R, Cattivelli L, Rosental L, Nikoloski Z, Zamir D, Fernie AR, Fait A.**2012** Metabolic Profiling of a Mapping Population Exposes New Insights in the Regulation of Seed Metabolism and Seed, Fruit, and Plant Relations. *PLoS Genet* **8:** e1002612

Verdier J, Lalanne D, Pelletier S, Torres-Jerez I, Righetti K, Bandyopadhyay K, Leprince O, Chatelain E, Vu BL, Gouzy J, Gamas P, Udvardi MK, 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

Welham T, Pike J, Horst I, Flemetakis E, Katinakis P, Kaneko T, Sato S, Tabata S, Perry J, Parniske M, Wang TL.**2009** A cytosolic invertase is required for normal growth and cell development in the model legume, Lotus japonicus. *J Exp Bot* **60:** 3353-3365

Xia J, Psychogios N, Young N, Wishart DS.**2009** MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Research* **37:** W652-W660

Xia J, Sinelnikov IV, Han B, Wishart DS.**2015** MetaboAnalyst 3.0—making metabolomics more meaningful. *Nucleic Acids Research*

Xia J, Wishart DS.**2011** Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. **6:** 743-760

Xue W, Batushansky A, Toubiana D, Botnick I, Szymanski J, Khozin-Goldberg I, Nikoloski Z, Lewinsohn E, Fait A.**2013** The Investment in Scent: Time-Resolved Metabolic Processes in Developing Volatile-Producing Nigella sativa L. Seeds. *PLos ONE* **8**

Yang M, Zhang B, Jia J, Yan C, Habaike A, Han Y.**2013** RRP41L, a putative core subunit of the exosome, plays an important role in seed germination and early seedling growth in Arabidopsis. *Plant Physiology* **161:** 165-178

Yau Y-Y, Simon P.**2003** A 2.5-kb insert eliminates acid soluble invertase isozyme II transcript in carrot (Daucus carota L.) roots, causing high sucrose accumulation. *Plant Molecular Biology* **53:** 151-162

Yoo H, Widhalm JR, Qian Y, Maeda H, Cooper BR, Jannasch AS, Gonda I, Lewinsohn E, Rhodes D, Dudareva N.**2013** An alternative pathway contributes to phenylalanine biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase. *Nature communications* **4**

Yu H, Du X, Zhang F, Zhang F, Hu Y, Liu S, Jiang X, Wang G, Liu D.**2012** A mutation in the E2 subunit of the mitochondrial pyruvate dehydrogenase complex in Arabidopsis reduces plant organ size and enhances the accumulation of amino acids and intermediate products of the TCA cycle. *Planta* **236:** 387-399

Yu T-S, Lue W-L, Wang S-M, Chen J.**2000** Mutation of Arabidopsis Plastid Phosphoglucose Isomerase Affects Leaf Starch Synthesis and Floral Initiation. *Plant Physiology* **123:** 319-326

Zheng ZL.**2009** Carbon and nitrogen nutrient balance signaling in plants. *Plant Signaling and Behavior* **4:** 584-591

Supporting information

Dataset S1: Averaged raw metabolite abundance data with standard errors for the 44 identified metabolites.

Dataset S2: ANOVA post-hoc statistical results.

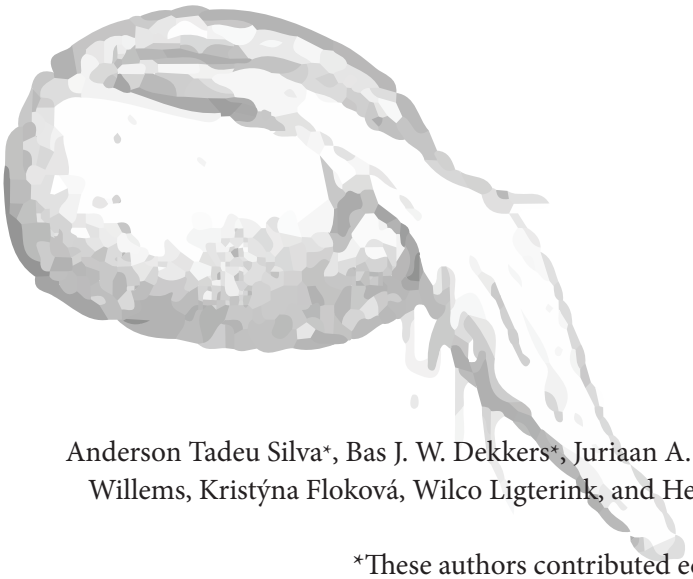
Dataset S3: Transcript-metabolite correlations.

Dataset S4: Assignment of metabolites to the profiles and their pathway enrichment.

Supplementary tables can be found at
http://www.wageningenseedlab.nl/thesis/atsilva/SI/

Chapter 4

Mild air drying treatment (MADT): a novel, efficient and robust protocol for studying desiccation tolerance in germinated seeds



Anderson Tadeu Silva*, Bas J. W. Dekkers*, Juriaan A. Rienstra, Leo A. J. Willems, Kristýna Floková, Wilco Ligterink, and Henk W. M. Hilhorst

*These authors contributed equally to this work

(*in preparation for submission*)

Summary

Desiccation tolerance (DT) is the ability of organisms to survive extreme water losses without accumulation of lethal damage. This trait was, in evolutionary sense, critical to conquer land. DT is present in all taxa including bacteria, fungi, roundworms and plants. In flowering plants, DT only occasionally occurs in vegetative tissues, in so-called resurrection plants, yet the vast majority of plant species produce seeds that tolerate desiccation and long-term dry storage. DT is acquired during seed development and subsequently lost upon germination. However, within a limited developmental window germinated seeds tolerate desiccation provided that they are pre-conditioned by mild osmotic stress before desiccation. The artificial re-induction of DT in germinating seeds, using an osmotic treatment has proven to be a powerful experimental approach to study multiple aspects of DT in seeds, including the genes involved, the role of hormones and the accompanying transcriptome. The re-induction of DT relies on a mild osmotic stress which is often applied using polyethylene glycol (PEG) treatment. However, PEG or any other osmoticum can cause undesirable side-effects. Furthermore, the use of PEG involves thorough cleaning to remove any remnants of PEG off the treated seeds and is therefore labour intensive. For these reasons, we developed an alternative protocol that we coined Mild Air Drying Treatment (MADT), to induce and evaluate DT in Arabidopsis seeds. In a benchmark study using a range of abscisic acid (ABA) deficient and ABA insensitive genotypes we conclude that MADT is easier, faster, robust and less labour intensive than the conventional PEG-based protocol. The new protocol for example allowed us to proof that enhanced ABA accumulation is part of the DT response.

Introduction

Desiccation tolerance (DT) is defined as the ability of organisms to deal with extreme water loss to levels below 0.1g H_2O per gram dry weight as well as with subsequent re-hydration without accumulation of lethal damage (Alpert, 2005; Oliver *et al.*, 2005; Leprince and Buitink, 2010). DT is an intriguing phenomenon since desiccation stress involves a multitude of aspects including cell shrinkage, denaturation of large molecules, loss of enzyme activity, formation of molecular aggregates, and damage caused by reactive oxygen species (Farrant, 2000; Scoffoni *et al.*, 2014). The ability to tolerate desiccation involves the activation of a series of protective measures collectively referred to as the 'desiccome' (Leprince and Buitink, 2010), as well as entering a state of quiescence and metabolic inactivity (Alpert, 2005).

In flowering plants, DT is rare in whole plants and only some 330 resurrection plant species (<0.15% of the total number of vascular plant species) are known that tolerate desiccation of their vegetative tissues (Proctor and Pence, 2002). Despite that DT is rare in vegetative organs of angiosperms, it is present in seeds of most plants species, which are then called orthodox seeds (Bewley *et al.*, 2013; Gaff and Oliver, 2013). DT is usually established just before the drying phase during seed maturation and subsequently lost during germination (Bewley *et al.*, 2013). Several studies have shown that rapid drying of seeds at distinct stages during germination killed seeds already before or quickly after visible germination, due to lack of DT (Vertucci and Farrant, 1995; Lin *et al.*, 1998; Buitink *et al.*, 2003; Daws *et al.*, 2007; Maia *et al.*, 2011). However, even after germinated seeds can survive such drying treatment when pre-conditioned using a mild osmotic stress before desiccation and this has been demonstrated for a number of species including Arabidopsis (Dekkers *et al.*, 2015). Thus, it can be concluded that the application of mild osmotic stress re-induces DT in germinated seeds. Nevertheless, this ability is strictly dependent on the developmental stage and after a certain time interval, germinated seeds completely lose the ability to tolerate extreme drying and become sensitive to desiccation (Leprince *et al.*, 2000; Buitink *et al.*, 2003; Vieira *et al.*, 2010; Maia *et al.*, 2011; Dekkers *et al.*, 2015).

As described above the ability to tolerate this extreme form of water loss is limited to a small developmental time window, the 'DT window' (Buitink *et al.*, 2003; Dekkers *et al.*, 2015). Interestingly, the DT window overlapped with a developmental window of abscisic acid (ABA) sensitivity (Maia *et al.*, 2014), which was identified by Lopez-Molina *et al.* (2001). These authors proposed that during this phase, the young seedlings monitor their environmental osmotic status and can undergo a developmental arrest, protecting them from, potentially lethal, water loss (Lopez-Molina *et al.*, 2001; Lopez-Molina *et al.*, 2002). The re-induction of DT by a mild osmotic treatment depends on the plant hormone ABA. Evidence for this was obtained by the use of the ABA-biosynthesis inhibitor fluridone on *Medicago truncatula* radicles and the use of the Arabidopsis ABA deficient mutant *aba2-1*, which in both cases impaired the re-induction of DT (Buitink *et al.*, 2003; Maia *et al.*, 2014). Interestingly, an ABA treatment alone is enough to re-induce DT, as has been shown in *Cucumis sativus*, Medicago and Arabidopsis (Lin *et al.*, 1998; Buitink *et al.*, 2003; Maia *et al.*, 2014; Costa *et al.*, 2015).

Gene regulatory network analysis identified transcription factors that were highly connected with DT regulated genes (Verdier *et al.*, 2013). These included three well known regulators in ABA signalling, i.e. the B3 class transcription factor ABA INSENSITIVE (ABI) 3 (Koornneef *et al.*, 1984; McCarty *et al.*, 1989; Giraudat *et*

al., 1992), the AP2 domain-containing transcription factor ABI4 (Finkelstein, 1994; Finkelstein *et al.*, 1998) and the bZIP class transcription factor ABI5 (Finkelstein, 1994; Finkelstein and Lynch, 2000; Lopez-Molina and Chua, 2000) which are therefore good candidates for regulators of DT. Indeed, several of these *abi* mutants were compromised for re-induction of DT in germinated seeds, both in Arabidopsis and in Medicago (Terrasson *et al.*, 2013; Maia *et al.*, 2014). However, the genetic interaction of these transcription factors during the induction of DT has not been studied yet.

The re-induction of DT relies on a mild osmotic stress before desiccation. This has been accomplished using a polyethylene glycol (PEG) treatment, for example in *Cedrela fissilis* Vell, *Impatiens walleriana*, Medicago, *Tabebuia impetiginosa* and Arabidopsis (Bruggink and van der Toorn, 1995; Buitink *et al.*, 2003; Vieira *et al.*, 2010; Maia *et al.*, 2011; Maia *et al.*, 2014; Masetto *et al.*, 2014). PEG is used in experiments to apply an osmotic stress and is used in seed treatments like priming (Copeland and McDonald, 1995). In the past, possible toxic effects of PEG have been reported including toxicity related to contaminants, blocking of water movement in plants and a reduced solubility and transport of oxygen (Lagerwerff *et al.*, 1961; Lawlor, 1970; Mexal *et al.*, 1975). Thus PEG treatments can cause undesirable side-effects. Furthermore, the induction of desiccation tolerance in seeds by PEG involves thorough cleaning to remove any remnants of PEG of the treated seeds and is, because of that, labour intensive. Therefore, we developed an easier, faster and less labour intensive protocol to induce and evaluate DT in Arabidopsis seeds and we demonstrate the robustness of this system. In a benchmark study we observed similar phenotypes for a range of ABA-related mutants between our new and the PEG-based protocol. Using the new protocol we could show that enhanced ABA accumulation is an essential part of the DT response. Interestingly, we found that ABA also accumulated in seeds upon drying beyond the DT window indicating that a factor downstream of ABA accumulation is lacking, such that DT cannot be induced in that developmental stage anymore.

Experimental procedures

Plant growth conditions

Seeds of all genotypes were sown on paper. Seedlings were transferred and grown on Rockwool® cubes in a climate room (20ºC day; 18ºC night) with a photoperiod of 16h light and 8h dark. Each Rockwool® cube was watered with Hyponex® solution (1g/l) twice a week. Matured seeds were harvested in four replicates from at least three plants.

Induction of DT in germinated Arabidopsis seeds by MADT and PEG-based protocols

Arabidopsis thaliana seeds, accession Col-0 (N60000), were cold-stratified for 72 hours at 4 ºC in Petri dishes with two layers of wet filter paper (Anchor paper Co.). After stratification, germination was performed under continuous light at 22 ºC. During germination seeds were selected at four different developmental stages, using a stereo microscope, to assess the ability to re-induce DT. The four developmental stages include seeds at testa rupture (TR), radicle protrusion (RP), appearance of the first root hairs (RH) and when cotyledons are turned green (GC) which occur in Col-0 around 20, 26, 33 and 44 hours after stratification, respectively. For each developmental stage three or four biological replicates were sampled each containing between 25 to 200 seeds. The samples were transferred to a 47 mm cellulose black membrane (Whatman®) which was placed on a blue filter paper moistened with 25 ml of demi water in a plastic tray (15 cm x 21 cm) and placed in an incubator (Van Den Berg Klimaattechniek, Reeuwijk, Netherlands) with air flow, set at 22 ºC and relative humidity of 32%. Under these conditions seeds were fully dehydrated over a period of six hours. After the dehydration treatment the seed samples were rehydrated on a Copenhagen table with a photoperiod of 12/12 hours at 22 ºC. Seeds that resumed their developmental program and turned into seedlings were scored as being DT. The induction of DT over these four stages was also performed using a PEG treatment as described by Maia *et al.* (2011).

Dehydration curves

The dehydration curves were made for all four developmental stages (TR, RP, RH and GC) in four biological replicate of 150 seeds each. The seed water content was determined at nine time points over a period of three days according the method described by Maia *et al.* (2011).

Construction of *ABI* double and triple mutant combinations

Crosses were made between *abi3*, *abi4* and *abi5* mutants to obtain the double and triple mutant combinations. The mutant alleles used were *abi3-9*, *abi3-10*, *abi4-3* and *abi5-7* which were originally reported by Nambara *et al.* (2002). To analyse the crosses, the F2 plants were genotyped using (d)CAPS markers which were developed using dCAPS Finder 2.0 (http://helix.wustl.edu/dcaps/dcaps.html) (Neff *et al.*, 2002). The following markers were used:

Abi3-9 dCAPS marker: For: CAACTAAAGAGGCAAGAAAGCAA Rev: CCATCACTGGCGGTAATTG, restriction with BsrDI digests the mutant fragment.

Abi3-10 CAPS marker: For: ACTCCGTTTGTGTTCCTCAG Rev: CCATCACTGGCGGTAATTG, restriction with FokI digests the WT fragment.

Abi4-3 CAPS marker: For: CAACTCCAAGTTCCGTTACC Rev: GGAGGTGGAAGGAGAAGAAG, restriction with DdeI digests the mutant fragment.

Abi5-7 dCAPS marker: For: CGTCAGAGCGAGAAGTAGAG Rev: GCGGGGCGGGGGCACGGGGGGGATTGTTATTATTCTCCTCTGCGAT, restriction with DpnII digests the WT fragment.

DNA was isolated using a protocol described by Cheung *et al.* (1993) and fragments were amplified by PCR using a 15 µl reaction volume using the Firepol DNA polymerase enzyme according to manufacturer's instruction. The PCR programme consisted of an initial heating to 95 ºC of 4 minutes followed by 40 cycles of 15 seconds at 95 ºC, 20 seconds at 57 ºC and 30 seconds at 68 ºC. To discriminate the WT and mutant alleles 5 µl of the PCR product was digested with the indicated restrictions enzymes (New England Biolabs) according to manufacturer's instructions and the fragment sizes were analysed on a 2.5% agarose gel.

Germination assay to screen for ABA sensitivity

Germination assays were performed in plastic (15 cm x 21 cm) trays containing two layers of blue filter paper (Anchor paper Co) supplemented with 50 ml of 10mM 2-N-morpholino ethanesulfonic acid buffer or MES containing ABA solutions. Per genotype four biological replicates of 100–150 seeds were sown. The germination was scored using the Germinator package as described previously (Joosen *et al.*, 2010).

ABA extraction and measurement

For ABA analysis, samples of 5–7 mg of dry weight of germinated seeds at RP were frozen in liquid nitrogen and ground in a dismembrator (Plate Shaker MM 400, MO BIO Laboratories Inc.) for 2 min with stainless steel balls. ABA was extracted with 1.5 ml of ice cold 10% methanol/H_2O (v/v) containing 5 pmol of [$^2H_6$]-(+)-cis,trans-ABA as an internal standard. Tubes were vortexed and sonicated for 3 min in an ultrasonic bath (Branson 3510, Danbury, CT, USA). Samples were incubated while shaking for 25 min at 4 ºC and subsequently centrifuged (5 min, 14000 rpm, 4 ºC). The supernatant was carefully transferred to a 4 ml glass vial. The pellets were re-extracted once more with 1 ml 10% methanol without the internal standard, after which tubes were vortexed and centrifuged again for 5 min. Both supernatants

were combined and further purified using Strata™-X polymeric reversed phase columns (30mg/3ml, Phenomenex Inc., Torrance, CA) according to Flokova *et al.* (2014) with modifications. Sorbent of column was equilibrated with 1ml of 100% MeOH and 1ml of MQ water. After samples were loaded onto column, highly polar compounds were removed by wash with 1ml of extraction solvent and pre-concentrated analyte was eluted with 3ml of 80% MeOH. The samples were dried in a SpeedVac centrifuge (SPD121P; Thermo Scientific, Tewksbury, MA, USA) and pellet was dissolved in 50 µl 15% acetonitrile: 85% 15 mM HCOOH (v/v). Purified samples were further analyzed by the ultra-high performace liquid chromatography – tandem mass spectrometry with electrospray interphase, consisting of the ACQUITY UPLC System (Waters, Milford, MA, USA) coupled to Xevo™ TQ-S (Waters MS Technologies, Manchester, UK) triple quadrupole mass spectrometer. The mixture of 15mM HCOOH (A) and acetnitrile (B) was used as a mobile phase. The samples were injected onto ACQUITY UPLC CSH™ C_{18} column (100x2.1 mm, 1.7 µm; Waters, Milford, MA, USA) and eluted with following binary gradient at flow rate 0.4ml/min and 40°C: isocratic elution at 15% B (from 0 to 1 min), linear increase to 60% B (1-7 min), linear increase to 80% B (7-9 min), logarithmic increase to 100% B (9-10 min). Finally the column was equilibrated for initial conditions for 2 min. The effluent was introduced into the electrospray ion source of mass spectrometer. ABA and [$^2H_6$]-(+)-cis,trans-ABA were quantified by multiple ion monitoring mode (MRM). MS/MS conditions were optimized to increase sensitivity of diagnostic transitions (ABA: 263.15 > 153.10; [$^2H_6$]-ABA: 269.15 > 159.10): the source block/ desolvation temperature 120°C/ 550°C, capillary voltage 3kV, cone voltage 25V and collision energy 10eV. Nitrogen was used as both desolvation (650 $l.h^{-1}$) and cone gas (150 $l.h^{-1}$). Quantification was achieved by standard isotope dilution method. The MassLynx™ software (version 4.1, Waters, Milford, MA, USA) was used to operate the instrument, acquire and process the MS data.

Results

Mild Air Drying Treatment (MADT): a novel protocol to induce DT in germinated Arabidopsis seeds.

Germinating Arabidopsis seeds are sensitive to a fast drying treatment, for example when seeds are placed on dry filter paper and dried under a relative humidity of 32% (Maia *et al.*, 2011). Drying under these conditions is very fast and seeds may not have enough time to deploy their protective mechanisms to cope with desiccation. Therefore we developed a method with a slower seed dehydration that would hopefully allow induction of DT. This was achieved by placing seeds on a wetted (instead of a dry) filter paper before drying at a low relative humidity. We named this *Mild Air Drying*

Treatment, or short, the MADT protocol. For this study we germinated Arabidopsis seeds and isolated seeds at four developmental stages along the germination time curve to assess the ability to re-induce DT. The four developmental stages were seeds at testa rupture (TR), radicle protrusion (RP), appearance of the first root hairs (RH) and when cotyledons turned green (GC) (Figure 1A).

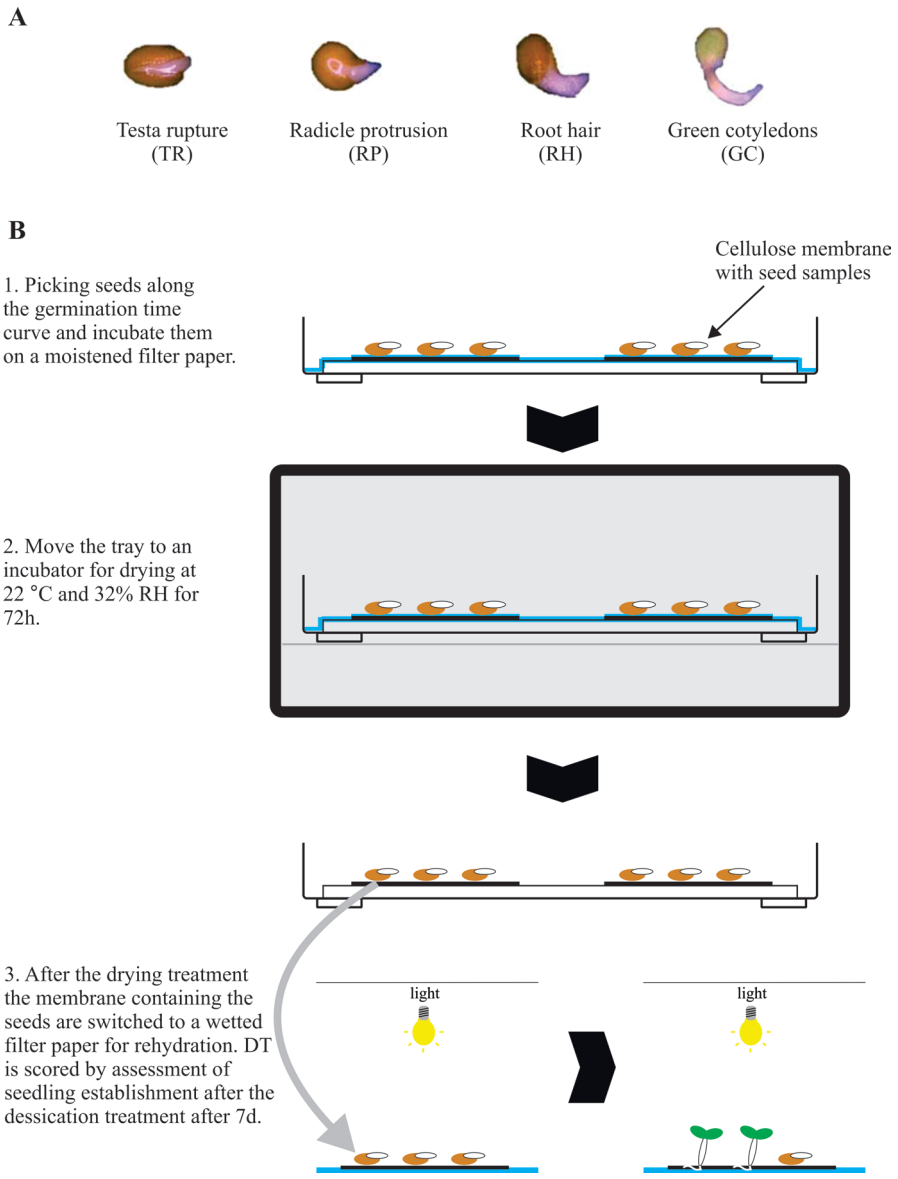


Figure 1. Re-establishment of DT using mild air drying treatment (MADT) **A.** Arabidopsis seeds at different developmental stages. **B.** Summarizing scheme of the MADT for the re-establishment of desiccation tolerance (DT) in germinated Arabidopsis seeds.

The MADT is initiated by transferring the germinated seeds to a cellulose membrane placed on a 14 x 19.5 cm blue filter paper that is moistened with 25 ml of demi water in a plastic tray (15 x 21 cm) (Figure 1B). To dry the seeds the tray is placed in an incubator, set at 22ºC and relative humidity of 32% with a forced air flow. After the dehydration treatment of three days the desiccated seeds are rehydrated in H_2O at 22ºC on a Copenhagen Table under a 12/12 h dark/light regime, and seeds that resume their developmental program and turn into seedlings are scored as being DT.

To obtain an impression of the seed water loss under these conditions we determined the seed water contents at nine time points over a period of three days for all four developmental stages. The dehydration curves revealed that the seed water content decreased steadily over the first six hours and seeds were desiccated (seed water content <0.1gr H_2O gr dry weight) after 10-24 hours depending on the stage (Figure 2A). Interestingly, seeds submitted to MADT had a temporal water content which is comparable to the water content obtained in Arabidopsis seeds using a -2.5 MPa PEG treatment (Maia *et al.*, 2011) (Figure 2B). Compared to fast drying, which results in a very quick drop in seed water content over the first two hours (Maia *et al.*, 2011)(Figure 2B), the dehydration by the MADT protocol is delayed by approximately 4 hours. This delay is enough to re-induce DT in germinated seeds as we observed that seeds at TR and RP stages were able to survive desiccation with this protocol in contrast to the fast drying protocol (Figure 3) (Maia *et al.*, 2011).

The ability to induce DT drops to ~20% at the RH stage while at the GC stage no DT was induced at all and these results are in agreement with the results obtained using the PEG-based protocol by Maia *et al.* (2011) (Figure 3A). The new protocol was used to induced DT in germinated Col-0 WT seeds in five independent experiments for the RP and RH stages and the obtained results were very similar showing the robustness of the MADT protocol (Figure 3B).

Thus, our novel MADT protocol shows that a reduced rate of water loss, compared with the fast drying treatment, is enough to allow induction of DT in germinated seeds just as with the PEG-based protocol. The results obtained with the new protocol are robust and reproducible and the protocol is less labour intensive as the PEG-based protocol. We employed our new protocol to examine the role of ABA in the re-establishment of DT and to study the genetic interactions between three transcription factors that are involved in the induction of DT.

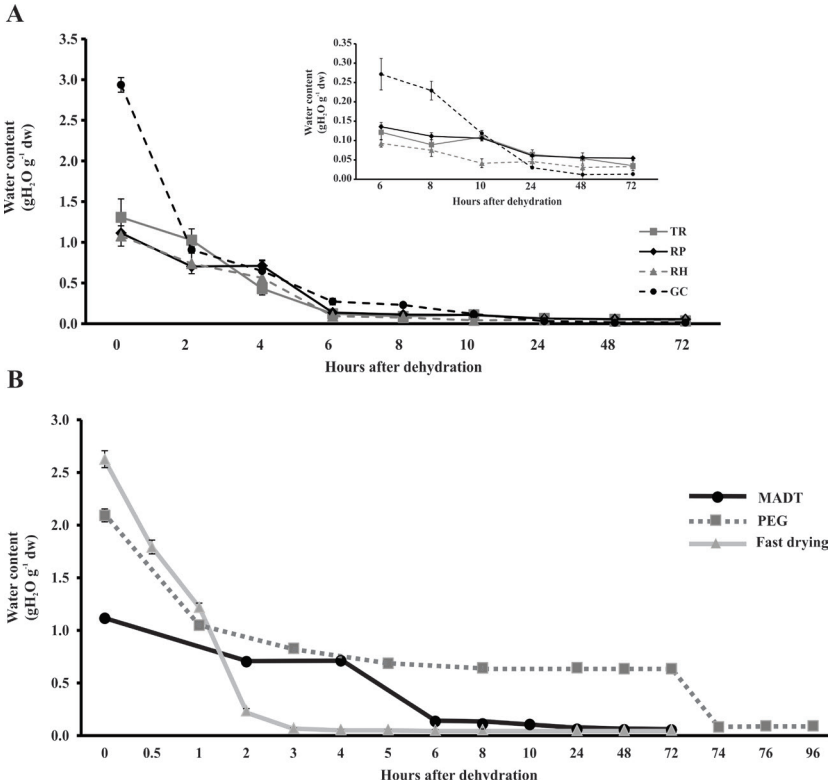


Figure 2. Arabidopsis seed water content. **A.** Dehydration using the mild air drying treatment (MADT) protocol. **B.** Comparison of the MADT protocol with -2.5 MPa polyethylene glycol (PEG) and Fast drying of Maia et al. (2011), samples were compared at radicle protrusion.

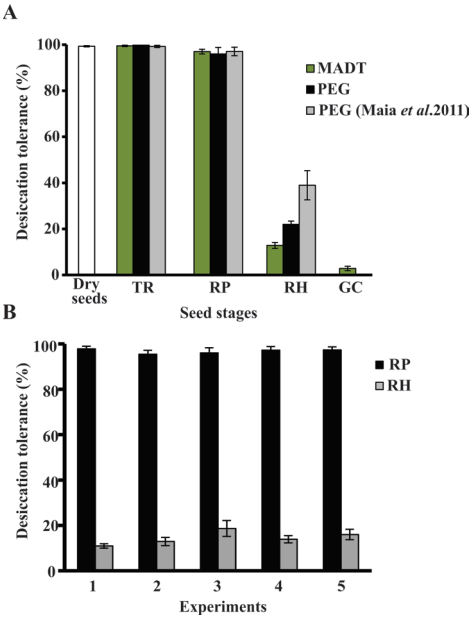


Figure 3. Re-establishment of DT **A.** Seedling survival rate of Arabidopsis seed developmental scored after mild air drying treatment (MADT) and-2.5 MPa polyethylene glycol (PEG) protocol followed by 5 days of rehydration in comparison with PEG (Maia et al. 2011). **B.** Five independent experiments. TR – testa rupture; RP – radicle protrusion; RH – root hair and GC – greening cotyledon stage.

ABA CONTENT INCREASES IN DESICCATING GERMINATED SEEDS

Earlier studies already indicated a role for ABA in desiccation tolerance. For example, DT induction in germinated seeds of the Arabidopsis ABA deficient mutant *aba2-1* was less efficient (Maia *et al.*, 2014). Here we compared the phenotypes of germinating seeds upon desiccation and rehydration of two ABA deficient mutants, *aba2-1* and *aba3-1*, using the MADT and PEG-based protocols. Seeds were germinated and RP seeds were used for a DT induction experiment. After the drying treatment the seeds were rehydrated and DT was assessed by looking at the percentage of normal developing seedlings. With both protocols the ABA deficient mutants showed a reduced induction of DT showing that DT induction by the MADT protocol also depends on ABA and, again, that the MADT protocol gives very similar results as the more labour intensive PEG-based protocol (Figure 4).

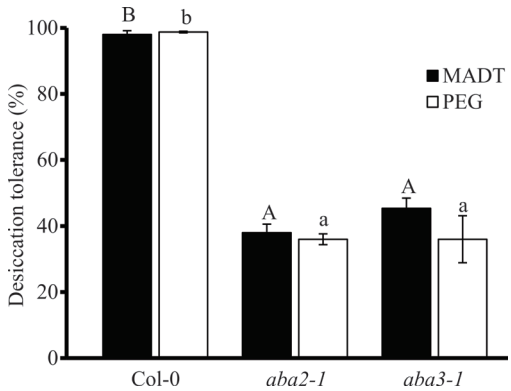


Figure 4. Comparison between mild air drying treatment (MADT) and polyethylene glycol (PEG) protocol with the re-establishment of ABA-deficient mutants. Different upper and lower case letters above bars represent differences between samples by Tukey's HSD (p<0.05). Upper case: MADT protocol and lower case: PEG protocol. Colours of bars refer to the protocol used.

Next, we investigated whether ABA accumulation takes place during the induction of desiccation tolerance using the MADT protocol. ABA content was determined at five time points; before drying, at three time points during drying (2, 6 and 72 hours) and 24 hours after rehydration. At 6 and 72 hours after drying higher seed ABA content was observed which declined upon rehydration to a similar level as observed before drying (Figure 5). The seed ABA content was also measured at the RH stage. Seeds in this stage show a severely reduced capacity to induce DT. We questioned whether this reduced ability is caused by the lack of ABA accumulation at this stage upon drying. Surprisingly, we observed similar ABA contents in the RP

and RH samples over the five time points (Figure 5). Thus, although many seeds at the RH stage do not display DT, these seeds do accumulate ABA to a similar level as compared to the RP stage. This indicates that the difference in the ability to induce DT is not linked to a reduced ability to accumulate ABA at the RH stage.

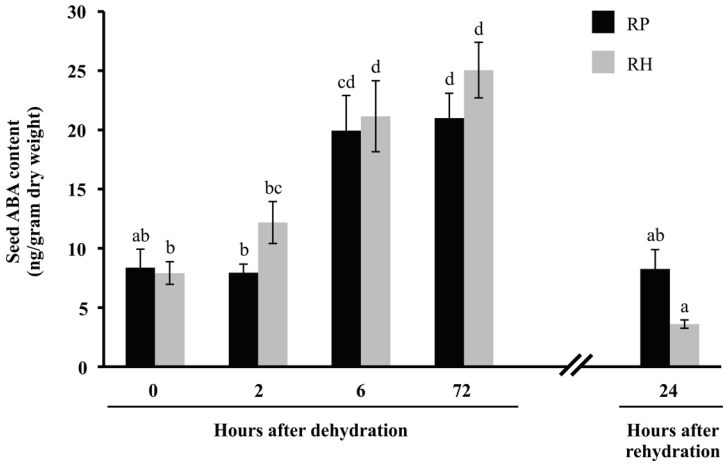


Figure 5. ABA content of Arabidopsis seeds at RP and RH developmental stages, and in different time points during mild air drying treatment (MADT) protocol. Bars represent the average of four independent replicates. Different lower case letters above bars represent differences between samples by Tukey's HSD ($p<0.05$). Colours of bars refer to developmental stages. RP – radicle protrusion and RH – root hair stage.

Genetic interactions between *ABI3*, *ABI4* and *ABI5* during the establishment of DT in germinated seeds.

Earlier work showed that the transcription factors *ABI3*, *ABI4* and *ABI5* are involved in the acquisition of DT in germinated seeds (Terrasson *et al.*, 2013; Maia *et al.*, 2014). However, the genetic interactions between the three transcription factors during the induction of DT have not been studied. Here we used *abi3-9*, *abi3-10*, *abi4-3* and *abi5-7* single mutants and constructed double and triple mutants between them to investigate their genetic relationships. Including the Col-0 WT, this resulted in twelve different genotypes which were confirmed by genotyping using (derived) Cleaved Amplified Polymorphic Sequence ((d)CAPS) markers (Figure 6).

First, we investigated the response of the different genotypes to ABA in a germination-based screen using a dose response curve ranging from 0 to 30 μM of ABA. The *abi3-9* and *abi3-10* alleles showed a slightly higher ABA insensitivity than *abi4-3* and *abi5-7* (Figure 7A). The double mutants *abi3-9 abi4-3*, *abi3-9 abi5-*

7, *abi3-10 abi4-3*, *abi3-10 abi5-7* and *abi4-3 abi5-7* all showed a strongly increased insensitivity to ABA compared to the single mutants and thus reveal genetic interaction effects between all *ABI* genes which are of a synergistic nature (Figure 7). Moreover, the triple mutants *abi3-9 abi4-3 abi5-7* and *abi3-10 abi4-3 abi5-7* again perform better on ABA compared to the double mutants and are highly insensitive even to the higher concentrations of 20-30 µM of ABA used in our assay (Figure 7).

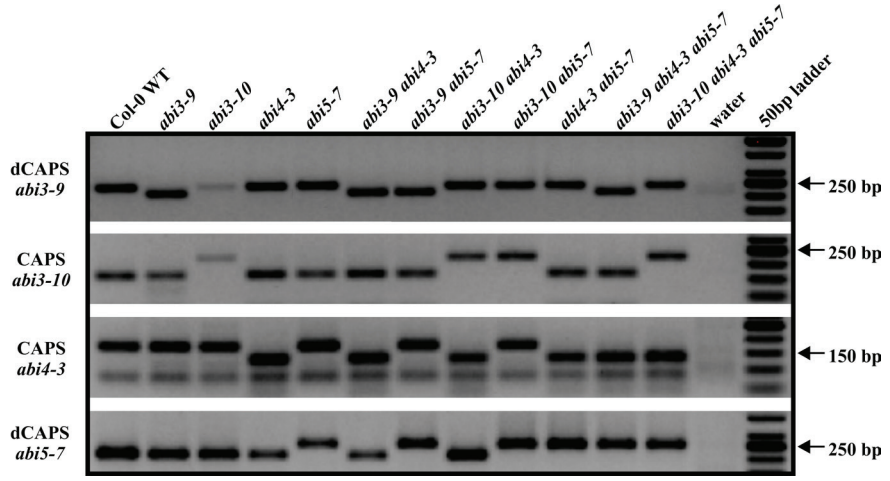


Figure 6. Genotyping of the double and triple mutants among *abi3*, *abi4* and *abi5*.

Next, we investigated the responses of these genotypes upon DT induction, both using the PEG-based and our new MADT protocol. This enables us to see whether the responses and the genetic interactions that they reveal are affected by the protocol used. We found that the response of the different mutants is similar for the PEG-based and the MADT protocol confirming our observations described above (Figure 8). In this experiment we used the RP stage for all genotypes. As expected the *abi* mutants were affected in the induction of DT, except *abi5-7* which did not show a phenotype at the RP stage but only at one stage later when the radicle has a length between 0.3 and 0.5 mm (Maia *et al.*, 2014). The fact that *abi5-7* does not show a phenotype while the *abi3-9* single mutant has a strong phenotype may impede the analysis of the genetic interactions. Nevertheless, the double mutants *abi4-3 abi5-7*, *abi3-9 abi4-3* and *abi3-10 abi4-3* showed a significantly reduced acquisition of DT compared with the single mutants. The results of this experiment suggest synergistic interactions between *ABI3* and *ABI4* and between *ABI4* and *ABI5* similarly to what we found for the germination response on ABA (Figure 7).

Interestingly, the re-establishment of DT in the double mutants of *abi3-9 abi5-7* and *abi3-10 abi5-7* resembles that of *abi3-9* and *abi3-10* single mutants, respectively

(Figure 8). This may either indicate that *ABI3* acts epistatically over *ABI5* in the DT response or that *ABI3* and *ABI5* act additively. However, since *abi5-7* lacks a DT phenotype at the RP stage we cannot differentiate between both possibilities. Moreover, no difference in the re-establishment DT was observed between the double mutant *abi3-9 abi4-3* and the *abi3-9 abi4-3 abi5-7* triple mutant and between the double mutant *abi3-10 abi4-3* and the triple mutant *abi3-10 abi4-3 abi5-7*.

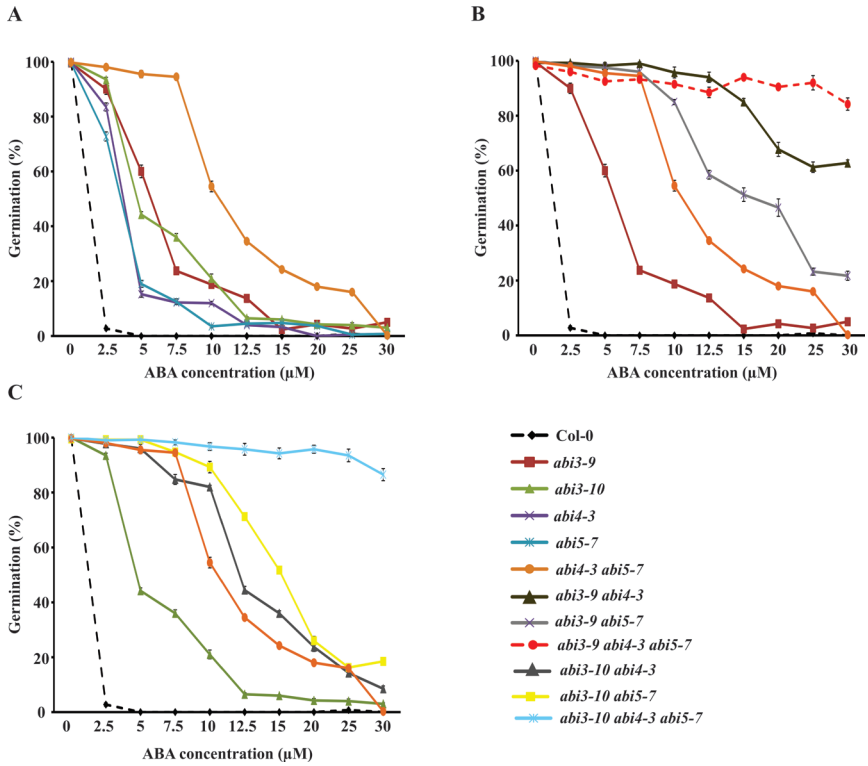


Figure 7. Dose response curve for ABA sensitivity of seed germination of single, double and triple *abi* mutant. ABA concentration from 0 µM to 30 µM.

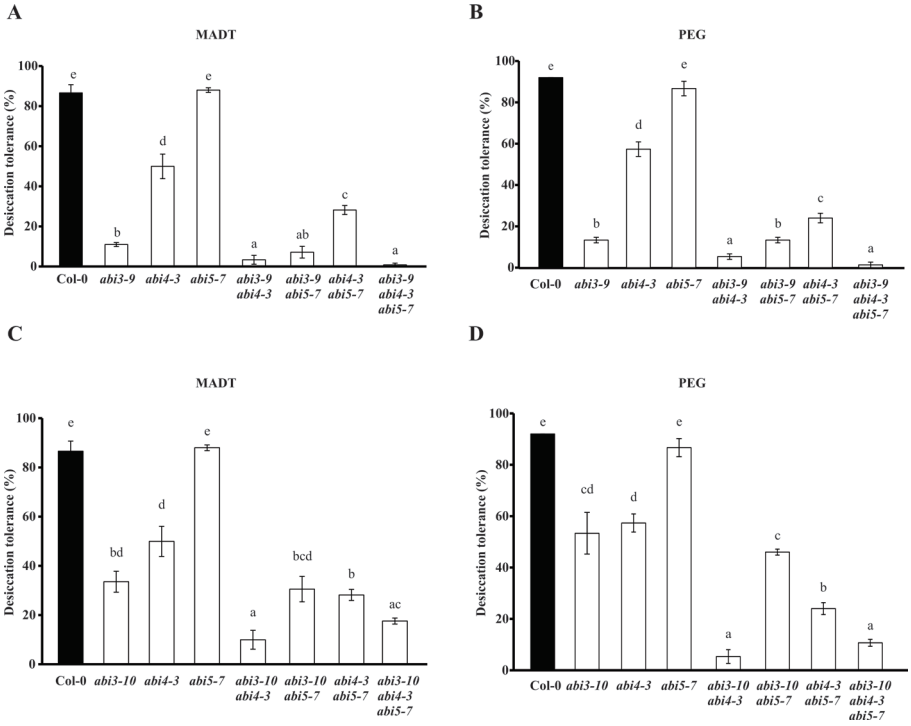


Figure 8. Comparison between mild air drying treatment (MADT) and 2.5 MPa polyethylene glycol (PEG) of the DT re-establishment in Arabidopsis seeds at radicle protrusion (RP) stage in several *abi* mutants. Bars represent the average of four independent replicates. Different lower case letters above bars represent differences between samples by Tukey's HSD (P<0.05)

Discussion

The induction of desiccation tolerance in germinated seeds provides an excellent system to understand different aspects of DT at the physiological and molecular level. It, for example, provides a system to evaluate the acquisition of DT in mutants to investigate gene function. Here we report a new, efficient and robust method, MADT, to induce and study DT in the model plant Arabidopsis. Using Col-0 WT and a range of mutant genotypes, our experiments demonstrate that the results obtained with our method are similar to results obtained with the conventional PEG-based DT re-induction method. Our new protocol is less labour intensive. Using PEG as an osmoticum involves thorough cleaning of the treated seeds to remove any remnants of PEG which makes this method more labour intensive. In our new protocol the seeds are sown on a cellulose membrane for the drying treatment which

is simply switched to a wetted filter paper for rehydration and subsequent scoring of the seedling phenotypes. Furthermore, the protocol is faster. The total time for the PEG-based protocol is 17 days, (three days stratification, about one day germination, three days PEG treatment, three days desiccation treatment followed by 7 days of rehydration and seedling scoring). Since the three day PEG treatment and three days of desiccation are replaced by a three day desiccation treatment on moistened filter paper our new protocol is three days faster. Thus, compared to the PEG-based protocol our new method delivers similar results although it is easier, less labour intensive and faster and therefore provides a more efficient alternative.

Despite the fact that the results are similar the two methods differ in two ways. The first concerns the accumulation of ABA in drying germinated seeds. The two ABA-deficient mutants tested in this study (*aba2-1* and *aba3-1*) displayed reduced re-establishment of DT, in both methods, confirming that ABA plays a critical role in the re-induction of DT (Buitink *et al.*, 2003; Maia *et al.*, 2014). For this reason, Maia *et al.* (2014) studied gene expression upon PEG treatment of genes related to ABA biosynthesis and ABA catabolism and observed that ABA biosynthesis genes were enhanced in expression while ABA catabolism genes were down-regulated. Thus, according to this gene expression analysis it was expected that the ABA level would increase upon PEG treatment. Surprisingly, two independent measurements of ABA content in PEG-treated seeds revealed no differences in ABA content between non-treated and treated seeds (Maia *et al.*, 2014). In contrast, we observed a clear increase in seed ABA content when seeds were desiccated using our new MADT protocol. ABA leakage from premature embryos of wheat imbibed in water appeared to be one of the important factors causing their premature germination (Suzuki *et al.*, 2000). Therefore, the striking difference in seed ABA content between both protocols, may be caused by the leaching or washing out of ABA during the 3-day PEG treatment.

Secondly, both methods differ in the time that is needed to induce DT. The induction of DT in a -2.5MPa PEG solution is relatively slow and an incubation period of three days is needed for maximal effect (Maia *et al.*, 2011). Fast drying of seeds (on a dry filter paper at 32% relative humidity) results in a drop of water content from 2.6 to 0.2 gram of H_2O per gram dry weight within 2 hours, and such harsh treatment is detrimental for germinated seeds. Using our new protocol, seeds were dried on moistened filter paper and a similar drop in water content now required six hours and resulted in DT seeds. This means that this delay is enough to employ the ‘desiccome‘ in germinated seeds (and hence induce DT) and much faster than the three days that are needed when using a PEG treatment. This relatively slow response of DT induction in PEG may indicate that PEG has some negative effects on the treated seeds. Perhaps the PEG solution reduces the availability of oxygen

as has been suggested before by Mexal *et al.* (1975), in the seeds slowing down their metabolism and hence the induction of DT. Alternatively this difference in time needed to induce DT might be attributed to the lower levels of ABA accumulating upon drying, thus decreasing the speed of DT induction of the PEG-based protocol. Since there is such a difference in the speed in which DT is induced, it may suggest that our new protocol mimics 'natural' drying in the soil better, as compared to the PEG-based protocol.

Genetic interactions between *ABI3*, *ABI4* and *ABI5* uncovered by double and triple mutants that differ in ABA sensitivity and DT re-establishment

Mutants in *abi3*, *abi4* and *abi5* were identified using ABA-based germination assays (Koornneef *et al.*, 1984; Finkelstein, 1994; Lopez-Molina and Chua, 2000). *ABI3*, *ABI4* and *ABI5* encode transcription factors of the B3, AP2 and bZIP class, respectively (McCarty *et al.*, 1989; Giraudat *et al.*, 1992; Finkelstein *et al.*, 1998; Finkelstein and Lynch, 2000; Lopez-Molina and Chua, 2000). They play a pivotal role in ABA, nutrient and stress signalling during seed maturation and seedling establishment (Finkelstein, 2013). The relationships between the ABI genes in ABA signalling are complex and not fully understood. Transcriptome analysis indicated that target genes of ABI3, ABI4 and ABI5 partly overlap (Reeves *et al.*, 2011; Finkelstein, 2013). Furthermore, there is cross regulation among the three *ABI* genes (Soderman *et al.*, 2000; Suzuki *et al.*, 2003). Here we investigated the genetic interactions between these three transcription factors during ABA signalling as well as the induction of DT by analysis of double and triple mutants.

ABI3 is an important regulator of seed maturation and severe mutations in this gene such as *abi3-5* and *abi3-6*, show a severely disturbed seed development (Santos-Mendoza *et al.*, 2008). These mutants produce seeds that are green, non-dormant, impaired in storage protein accumulation, sensitive to desiccation and displaying strongly reduced longevity (Ooms *et al.*, 1993; Nambara *et al.*, 1995; Delmas *et al.*, 2013). These severe mutants are also very insensitive to ABA and due to the strong nature of the ABA insensitivity the use of such mutants may hide genetic interactions present between two genes. Therefore we used weaker alleles of *abi3* that are ABA insensitive but show a relatively normal seed development as demonstrated by their brown seeds, which are desiccation tolerant and storable for a considerable period of time(Ooms *et al.*, 1993; Nambara *et al.*, 2002; Clerkx *et al.*, 2004).

Comparisons between the single mutants and double mutants showed that *ABI4* interacts genetically in a synergistic fashion with *ABI3* and *ABI5* both

during ABA signalling (based on ABA inhibition of seed germination) as well as the induction of DT, indicating that *ABI4* acts in part redundantly with *ABI3* and *ABI5* to regulate both responses. The strong ABA insensitive response of the *abi4-3 abi5-7* double mutant compared to the single mutants are in agreement with those observed for the *abi4-1 abi5-1* double mutant reported by Reeves *et al.* (2011). Also *ABI5* showed synergistic interaction with *ABI3* alleles during ABA signalling. However, such synergistic interaction was not detected during the re-induction of DT. Instead our results either indicate that *ABI3* acts epistatically over *ABI5* in the DT response or that *ABI3* and *ABI5* act genetically in an additive manner. However, since *abi5-7* lacks a DT phenotype at the RP stage we cannot differentiate between both possibilities. We could investigate the genetic relationship between *ABI3* and *ABI5* during the re-establishment of DT at stage three (when the radicle has a length between 0.3 and 0.5 mm, Maia *et al*. 2011). At this stage *abi5-7* shows a DT phenotype and would allow to obtain a more conclusive answer regarding this genetic interaction. Interestingly in this respect, is the fact that ABI3 and ABI5 have been reported to interact physically (Nakamura *et al.*, 2001). Furthermore, Lopez-Molina *et al.* (2002) studied the genetic relationship between *ABI3* and *ABI5* and their results suggested that ABI5 acts down-stream of ABI3 to arrest early seedling development. These observations suggest that ABI3 and ABI5 act, at least in part, in a similar signalling pathway to regulate ABA responses which is in agreement with the genetic interaction that we uncovered during the re-establishment of DT. The different genetic interactions between *ABI3* and *ABI5* that we observed between ABA signalling and the induction of DT could be explained by the fact that we scored ABA signalling by germination (radicle protrusion) which is a different response from the induction of DT, suggesting that multiple different genetic interactions may exist among the *ABI* genes during various ABA-regulated processes. The genotypes used here, could also be screened in other processes like germination at high temperature, or under osmotic or salt stress as well as several sugar responses which are all known to be ABA-regulated. This would aid to increase our insights concerning the genetic interactions among the *ABI* genes in a variety of ABA-regulated responses.

Acknowledgements

This work was funded by “Coordenação de Aperfeiçoamento de Pessoal de Nível Superior” (CAPES, Brazil). We thank Wim Soppe (Max Planck Institute for Plant Breeding Research, Cologne, Germany) for providing the *aba3-1* seeds.

References

Alpert P.2005 The Limits and Frontiers of Desiccation-Tolerant Life. *Integrative and Comparative Biology* **45:** 685-695

Bewley JD, Bradford K, Hilhorst HM, Nonogaki H (2013) Germination. *In* Seeds. Springer New York, pp 133-181

Bruggink T, van der Toorn P.1995 Induction of desiccation tolerance in germinated seeds. *Seed Science Research* **5:** 1-4

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

Cheung WY, Hubert N, Landry BS.1993 A simple and rapid DNA microextraction method for plant, animal, and insect suitable for RAPD and other PCR analyses. *PCR Methods Appl* **3:** 69-70

Clerkx EJM, Blankestijn-De Vries H, Ruys GJ, Groot SPC, Koornneef M.2004 Genetic differences in seed longevity of various Arabidopsis mutants. *Physiologia Plantarum* **121:** 448-461

Copeland LO, McDonald MB (1995) Principles of seed science and technology. Chapman & Hall, New York, USA

Costa M, Righetti K, Nijveen H, Yazdanpanah F, Ligterink W, Buitink J, Hilhorst HM.2015 A gene co-expression network predicts functional genes controlling the re-establishment of desiccation tolerance in germinated Arabidopsis thaliana seeds. *Planta***:** 1-15

Daws MI, Bolton S, Burslem DFRP, Garwood NC, Mullins CE.2007 Loss of desiccation tolerance during germination in neo-tropical pioneer seeds: Implications for seed mortality and germination characteristics. *Seed Science Research* **17:** 273-281

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

Delmas F, Sankaranarayanan S, Deb S, Widdup E, Bournonville C, Bollier N, Northey JG, McCourt P, Samuel MA.2013 ABI3 controls embryo degreening through Mendel's I locus. *Proc Natl Acad Sci U S A* **110:** E3888-3894

Farrant J.2000 A comparison of mechanisms of desiccation tolerance among three angiosperm resurrection plant species. *Plant Ecology* **151:** 29-39

Finkelstein R.2013 Abscisic Acid Synthesis and Response. *The Arabidopsis Book / American Society of Plant Biologists* **11:** e0166

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

Finkelstein RR, Li Wang M, Lynch TJ, Rao S, Goodman HM.1998 The Arabidopsis Abscisic Acid Response Locus ABI4 Encodes an APETALA2 Domain Protein. *The Plant Cell* **10:** 1043-1054

Finkelstein RR, Lynch TJ.2000 The Arabidopsis Abscisic Acid Response Gene ABI5 Encodes a Basic Leucine Zipper Transcription Factor. *The Plant Cell* **12:** 599-609

Finkelstein RR, Wang ML, Lynch TJ, Rao S, Goodman HM.1998 The Arabidopsis abscisic acid response locus ABI4 encodes an APETALA 2 domain protein. *The Plant Cell* **10:** 1043-1054

Flokova K, Tarkowska D, Miersch O, Strnad M, Wasternack C, Novak O.2014 UHPLC-MS/MS based target profiling of stress-induced phytohormones. *Phytochemistry* **105:** 147-157

Gaff DF, Oliver M.2013 The evolution of desiccation tolerance in angiosperm plants: A rare yet common phenomenon. *Functional Plant Biology* **40:** 315-328

Giraudat J, Hauge BM, Valon C, Smalle J, Parcy F, Goodman HM.1992 Isolation of the Arabidopsis ABI3 gene by positional cloning. *Plant Cell* **4:** 1251-1261

Joosen RVL, Kodde J, Willems LAJ, Ligterink W, van der Plas LHW, Hilhorst HWM.2010 germinator: a software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. *The Plant Journal* **62:** 148-159

Koornneef M, Reuling G, Karssen CM.1984 The isolation and characterization of abscisic acid-insensitive mutants of Arabidopsis thaliana. *Physiologia Plantarum* **61:** 377-383

Lagerwerff JV, Ogata G, Eagle HE.1961 Control of osmotic pressure of culture solutions with polyethylene glycol. *Science* **133:** 1486-1487

Lawlor DW.1970 Absorption of Polyethylene Glycols by Plants and their Effects on Plant Growth. *New Phytologist* **69:** 501-513

Leprince O, Buitink J.2010 Desiccation tolerance: From genomics to the field. *Plant Science* **179:** 554-564

Leprince O, Harren FJM, Buitink J, Alberda M, Hoekstra FA.2000 Metabolic Dysfunction and Unabated Respiration Precede the Loss of Membrane Integrity during Dehydration of Germinating Radicles. *Plant Physiology* **122:** 597-608

Lin T-P, Yen W-L, Chien C-T.1998 Disappearance of desiccation tolerance of imbibed crop seeds is not associated with the decline of oligosaccharides. *Journal of Experimental Botany* **49:** 1203-1212

Lopez-Molina L, Chua NH.2000 A null mutation in a bZIP factor confers ABA-insensitivity in Arabidopsis thaliana. *Plant Cell Physiol* **41:** 541-547

Lopez-Molina L, Mongrand S, Chua N-H.2001 A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factor in Arabidopsis. *Proceedings of the National Academy of Sciences* **98:** 4782-4787

Lopez-Molina L, Mongrand S, McLachlin DT, Chait BT, Chua NH.2002 ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *Plant J* **32:** 317-328

Maia J, Dekkers BJW, Dolle MJ, Ligterink W, Hilhorst HWM.2014 Abscisic acid (ABA) sensitivity regulates desiccation tolerance in germinated Arabidopsis seeds. *New Phytologist***:** n/a-n/a

Maia J, Dekkers BJW, Provart NJ, Ligterink W, Hilhorst HWM.2011 The Re-

Establishment of Desiccation Tolerance in Germinated *Arabidopsis thaliana* Seeds and Its Associated Transcriptome. *PLos ONE* **6**

Masetto TE, Faria JM, Fraiz AC.2014 Re-induction of desiccation tolerance after germination of Cedrela fissilis Vell. seeds. *Anais da Academia Brasileira de Ciências* **86:** 1273-1286

McCarty DR, Carson CB, Stinard PS, Robertson DS.1989 Molecular Analysis of viviparous-1: An Abscisic Acid-Insensitive Mutant of Maize. *The Plant Cell* **1:** 523-532

Mexal J, Fisher JT, Osteryoung J, Reid CP.1975 Oxygen availability in polyethylene glycol solutions and its implications in plant-water relations. *Plant Physiol* **55:** 20-24

Nakamura S, Lynch TJ, Finkelstein RR.2001 Physical interactions between ABA response loci of Arabidopsis. *Plant J* **26:** 627-635

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

Nambara E, Suzuki M, Abrams S, McCarty DR, Kamiya Y, McCourt P.2002 A Screen for Genes That Function in Abscisic Acid Signaling in Arabidopsis thaliana. *Genetics* **161:** 1247-1255

Neff MM, Turk E, Kalishman M.2002 Web-based primer design for single nucleotide polymorphism analysis. *Trends Genet* **18:** 613-615

Oliver MJ, Velten J, Mishler BD.2005 Desiccation tolerance in bryophytes: a reflection of the primitive strategy for plant survival in dehydrating habitats? *Integr Comp Biol* **45:** 788-799

Ooms J, Leon-Kloosterziel KM, Bartels D, Koornneef M, Karssen CM.1993 Acquisition of Desiccation Tolerance and Longevity in Seeds of Arabidopsis thaliana (A Comparative Study Using Abscisic Acid-Insensitive abi3 Mutants). *Plant Physiology* **102:** 1185-1191

Proctor MCF, Pence VC (2002) Vegetative tissues: Bryophytes vascular resurrection plants and vegetative propagules. *In* M Black, HW Pritchard, eds, Desiccation and survival in plants: drying without dying. CAB International, Wallingford, pp 207–238

Reeves W, Lynch T, Mobin R, Finkelstein R.2011 Direct targets of the transcription factors ABA-Insensitive(ABI)4 and ABI5 reveal synergistic action by ABI4 and several bZIP ABA response factors. *Plant Molecular Biology* **75:** 347-363

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

Scoffoni C, Vuong C, Diep S, Cochard H, Sack L.2014 Leaf Shrinkage with Dehydration: Coordination with Hydraulic Vulnerability and Drought Tolerance. *Plant Physiology* **164:** 1772-1788

Soderman EM, Brocard IM, Lynch TJ, Finkelstein RR.2000 Regulation and function of the Arabidopsis ABA-insensitive4 gene in seed and abscisic acid response signaling networks. *Plant Physiol* **124:** 1752-1765

Suzuki M, Ketterling MG, Li Q-B, McCarty DR.2003 Viviparous1 Alters Global Gene Expression Patterns through Regulation of Abscisic Acid Signaling. *Plant Physiology* **132:** 1664-1677

Suzuki T, Matsuura T, Kawakami N, Noda K.2000 Accumulation and leakage of abscisic acid during embryo development and seed dormancy in wheat. *Plant Growth Regulation* **30:** 253-260

Terrasson E, Buitink J, Righetti K, Ly Vu B, Pelletier S, Zinsmeister J, Lalanne D, Leprince O.2013 An emerging picture of the seed desiccome: confirmed regulators and newcomers identified using transcriptome comparison. *Front Plant Sci* **4:** 497

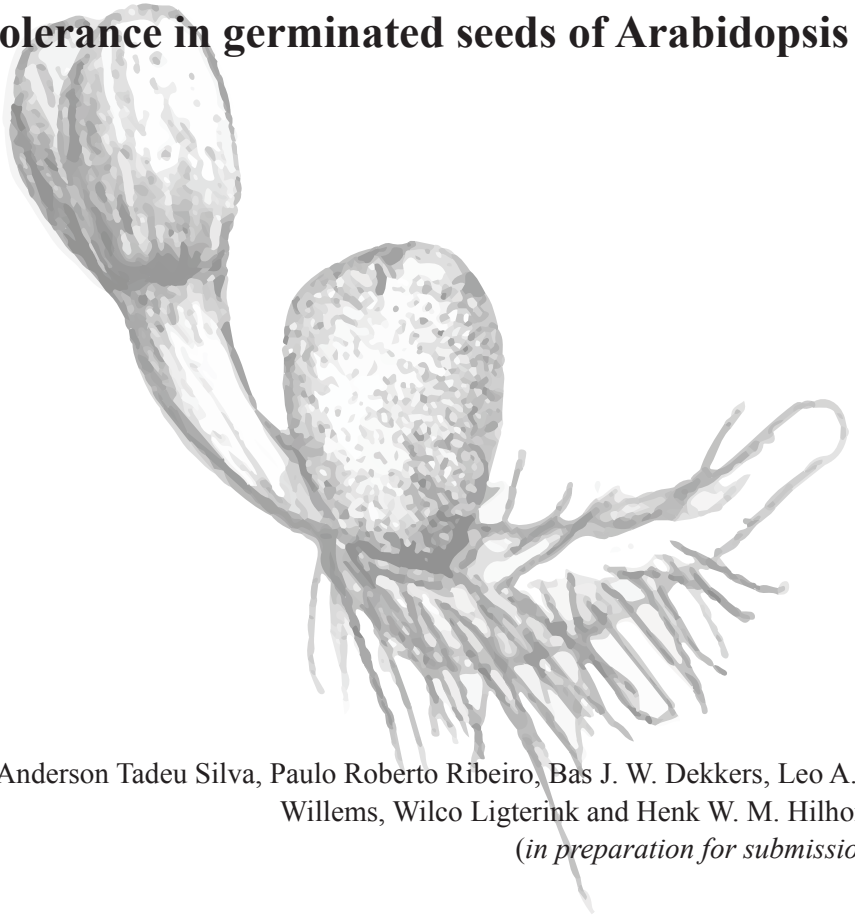
Verdier J, Lalanne D, Pelletier S, Torres-Jerez I, Righetti K, Bandyopadhyay K, Leprince O, Chatelain E, Vu BL, Gouzy J, Gamas P, Udvardi MK, Buitink J.2013 A Regulatory Network-Based Approach Dissects Late Maturation Processes Related to the Acquisition of Desiccation Tolerance and Longevity of Medicago truncatula Seeds. *Plant Physiology* **163:** 757-774

Vertucci C, Farrant J (1995) Acquisition and loss of desiccation tolerance. *In* J Kigel, G Galili, eds, Seed development and germination. Marcel Dekker, New York, pp 237–271

Vieira CV, da Silva EAA, de Alvarenga AA, de Castro EM, Toorop PE.2010 Stress-associated factors increase after desiccation of germinated seeds of Tabebuia impetiginosa Mart. *Plant Growth Regulation* **62:** 257-263

Chapter 5

ABA-dependent and -independent transcription factors control the re-establishment of desiccation tolerance in germinated seeds of Arabidopsis



Anderson Tadeu Silva, Paulo Roberto Ribeiro, Bas J. W. Dekkers, Leo A. J. Willems, Wilco Ligterink and Henk W. M. Hilhorst

(*in preparation for submission*)

Summary

Seeds are capable of surviving extremely low cellular water levels for remarkable periods of time. It is the result of an adaptation that limits cellular damage and is commonly known as 'desiccation tolerance' (DT). DT is gradually lost during seed germination but can be fully re-established at the stage of radicle protrusion (RP) but not at the root hair (RH) developmental stage. Re-establishment of DT requires the plant hormone abscisic acid (ABA). To delineate this role, we explored five ABA INSENSITIVE (*ABI)* mutant lines, *abi3-8*, *abi3-9*, *abi3-10*, *abi4-3* and *abi5-7*, for their ability to re-establish DT by a Mild Air Drying Treatment (MADT). The *abi3-9* mutant appeared to be more sensitive to dehydration than Col-0 and the other *abi* mutants at the RP stage. Thus, to decipher the molecular mechanisms involved in the re-establishment of DT upon dehydration, we executed a comprehensive transcriptomic analysis of the response of desiccation tolerant (RP of Col-0) and desiccation sensitive (RP of *abi3-9* and RH of Col-0) germinated seeds. The results indicate that a common set of genes is coordinately up-regulated (DP1 and DP2) or down-regulated (DP6), when DT is acquired, independently of developmental stage or *ABI3*. These DT-specific gene sets seem to promote adaptation to stress of Col-0 at RP and may be partly responsible for survival in the dry state. Finally, we identified a sub-network in the DT gene co-expression network, that connects DT-specific gene sets to the ABI3 regulon. A predictive transcriptional module of this sub-network suggests that the re-establishment of DT in Arabidopsis is controlled by crosstalk between ABA-dependent and ABA-independent pathways.

Introduction

Seed is an important dispersal unit in the plant's life cycle and for its survival as a species. Seeds are capable of surviving extremely low cellular water levels for remarkable periods of time (Meurs *et al.*, 1992; Potts, 2001). The ability to survive such a desiccated state, at water contents below 0.1 gram H_2O per gram dry weight, is the result of an adaptation that limits cellular damage during this period (Koster, 1991), which is commonly known as 'desiccation tolerance' (DT). Although the DT mechanism is present in most seeds, there are approximately 120 higher plants that can also survive desiccation of their vegetative tissues; these are called resurrection plants (Bartels and Hussain, 2011). In a recently study of the resurrection plant *Boea hygrometrica* (Xiao *et al.*, 2015), it was shown that the changes in the regulation of gene expression in this species under dehydration involve alternative splicing of transcripts and the plant hormone abscisic acid (ABA). The DT mechanism of

resurrection plants most likely resembles the pre-existing DT mechanism of seeds (Gaff and Oliver, 2013; Xiao *et al.*, 2015). DT has been studied widely in seed development and in relation with its re-establishment during germination, and has been shown to be controlled by ABA in seeds (Meurs *et al.*, 1992; Buitink *et al.*, 2003; Maia *et al.*, 2011; Terrasson *et al.*, 2013; Verdier *et al.*, 2013; Dekkers *et al.*, 2015).

DT is gradually lost during seed germination, but it can be fully re-established within a developmental time window between radicle protrusion (RP) and the root hair (RH) stage in *Arabidopsis thaliana* (Maia *et al.*, 2011). Re-establishment of DT in germinated seeds has also been observed in other species such as *Medicago truncatula* (Buitink *et al.*, 2003), *Tabebuia impetiginosa* (Vieira *et al.*, 2010) and *Cedrela fissilis* Vell (Masetto *et al.*, 2014). Several *abi* mutants of Arabidopsis (*abi3-8*, *abi3-9*, *abi4-3* and *abi5-7*) have been shown to fail to fully re-establish DT in germinated seeds at the RP stage, except for *abi5-7* (Maia *et al.*, 2014). Thus although ABA is clearly important in the re-establishment of DT there are indications that ABA-signalling is not the sole signal-transduction pathway involved. A desiccation-induced homeodomain-leucine zipper protein CPHB-1, of which expression is induced early in the drying process applied to the resurrection plant *Craterostigma plantagineum*, is not induced by ABA (Frank *et al.*, 1998).

The ability of germinated seeds to tolerate low water levels is limited to a short developmental time window (Buitink *et al.*, 2003; Maia *et al.*, 2011), during which plants monitor the environmental osmotic status before initiating vegetative growth (Lopez-Molina *et al.*, 2001). Considering that an osmotic- or ABA treatment can re-establish DT in germinated seeds (Buitink *et al.*, 2003; Maia *et al.*, 2011; Maia *et al.*, 2014), together with the observation that ABA is commonly synthesized upon dehydration stress (Nakashima and Yamaguchi-Shinozaki, 2013; Virlouvet *et al.*, 2014), we investigated whether there is cross-talk between the ABA-signalling pathway and pathways induced by dehydration, independently of ABA, to re-establish DT.

To study the re-establishment of DT in germinated Arabidopsis seeds, we developed a novel method called 'Mild Air Drying Treatment (MADT)', which is able to re-establish DT in germinated seeds at the RP stage without the need for an osmotic treatment or application of ABA (Chapter 4). This new method better mimics the 'natural' drying conditions that can be experienced in the soil, as compared to treatment with an osmoticum, such as the often used polyethylene glycol (PEG). To assess the molecular mechanisms involved in re-establishment of DT upon dehydration, we used transcriptome analysis on desiccation tolerant and desiccation sensitive seeds upon dehydration in time. The gene expression patterns confirm that

ABA is essential for the re-establishment of DT in germinated Arabidopsis seeds. However, also non-ABA-dependent genes seem to play a role in the acquisition of DT. We identified a putative regulatory circuit involving DT-responsive transcription factors, some of which are ABA-dependent while others are not.

Experimental procedures

Phenotypic analysis

Five *ABA INSENSITIVE (ABI*) mutant lines, *abi3-8*, *abi3-9*, *abi3-10*, *abi4-3* and *abi5-7* (Nambara *et al.*, 2002), were used to determine their ability to re-establish DT. Seeds of the mutants together with corresponding wild-type seeds (Col-0 [N60000]) were sown until germinated and seedlings were grown on Rockwool® cubes in a climate room (20ºC day; 18ºC night) with a photoperiod of 16h light and 8h dark. Each Rockwool® cube was watered with Hyponex® solution (1g/L) twice a week. Matured seeds were harvested in four replicates from at least three plants. For germination, seeds were cold stratified for 72h at 4ºC in a 9-cm Petri dishes with two moistened blue filter paper sheets (Anchor paper Co.) in 10 mL of demi-water to remove residual dormancy. Thereafter, seeds were transferred to a germination cabinet at 22ºC with constant white light. Re-establishment of DT in germinated seeds was performed as described before (Chapter 4).

Plant material, sample preparation and RNA extraction

Each sample used for RNA extraction was determined according to rate of seedling survival during the dehydration procedure (Chapter 4). Four time-points were selected during dehydration of germinated seeds at the RP stage: zero hours (0HAD) as a control and two hours (2HAD), six hours (6HAD) and 72 hours (72HAD). Additionally, a sample was taken of seeds at the RP stage that had been dried for 72 hours in an incubator under controlled conditions and that were rehydrated for 24 hours on a Copenhagen table (24HAR).To determine the transcriptomic characteristics of failing re-establishment of DT in the *abi3-9* mutant at RP and in Col-0 at RH, three intervals were selected: 0HAD, 2HAD and 6HAD. Only these three time-points were selected to avoid sampling of dead tissues, because seedling survival rate after eight hours of dehydration is already down to 75% for these seeds. RNA was extracted, using a slightly modified hot-borate protocol (Maia *et al.*, 2011).

Microarray hybridization and data analysis

Quality control, RNA labeling, hybridization and data extraction were performed at ServiceXS B.V. (Leiden, The Netherlands). Labelled ss-cDNA was synthesized using the Affymetrix NuGEN Ovation PicoSL WTA v2 kit and Biotin Module using 50 ng of total RNA as template. The fragmented ss-cDNA was utilized for hybridization on an Affymetrix ARAGene 1.1ST array plate. The Affymetrix HWS Kit was used for the hybridization, washing and staining of the plate. Scanning of the array plates was performed using the Affymetrix GeneTitan scanner. All procedures were performed according to the instructions of the manufacturers (nugen.com and affymetrix.com). The resulting data were analyzed using the R statistical programming environment and the Bioconductor packages (Gentleman *et al.*, 2005). The data was normalized using the RMA algorithm (Irizarry *et al.*, 2003) with the TAIRG v17 cdf file (http://brainarray.mbni.med.umich.edu). Expression data are in Supplemental Table S1. Validation of the microarray data set was performed by comparison with a gene expression analysis using RT-qPCR of ten genes (Supplemental Table S1). RT-qPCR was performed as described by Dekkers *et al.* (2013).

Identification of co-expression gene sets

Dominant expression patterns. Dominant expression patterns (DPs) were identified as previously described by Belostotsky *et al.* (2009). DPs were identified of the 50% most variant genes, corresponding to 13.918 mRNAs. The R function FANNY (http://cran.r-project.org/web/packages/cluster/cluster.pdf) with a minimum Pearson correlation of 0.85 was used to evaluate the number of cluster (K) choices from 1 to 50 with a cut-off for cluster membership of 0.4. The K choice that yielded the greatest number of transcriptional modules was seven.

Dehydration-inducible gene set. We used the Limma package (Gentleman *et al.*, 2005) for selection of differentially expressed transcripts induced upon dehydration for 6HAD and 72HAD relative to all others in a given stage and time-point. Differentially expressed transcripts specific to 6HAD and 72HAD in Col-0 at the RP stage were selected based on differential expression at a significance with $P < 0.01$ and an expression two-fold or higher compared to the other samples.

Co-expression network and transcriptional module analysis

Co-expression network. Pearson correlation coefficients were calculated for each pair of genes. The threshold for correlation coefficient values was determined according

to Freeman *et al.* (2007). Edge adjacency threshold was set at 0.94, resulting in a network consisting of 2.332 nodes with 33.476 edges. Correlation coefficient values were imported into Cytoscape v.2.8.2 (Smoot *et al.*, 2011) and the network visualization was done with the organic layout method.

Transcriptional modules. Transcriptional modules were produced by the ChipEnrich software package developed by Brady *et al.* (2007) and modified by Belmonte et al. (2013). ChipeEnrich determines the significance of GO terms, metabolic processes, DNA motifs and transcription factors (TFs) using *P* values calculated from their hypergeometric distribution (Belostotsky *et al.*, 2009; Belmonte *et al.*, 2013). For the hypergeometric distribution, lists of GO terms, metabolic processes, DNA motifs and TFs were used based on the Arabidopsis Gene Regulatory Information Server – AGRIS (http://arabidopsis.med.ohio-state.edu/AtTFDB/). An optimized ChipEnrich by Belmonte *et al.* (2013) was used to identify significantly enriched DNA motifs associated with TFs and GO terms. Tables generated by ChipEnrich were imported into Cytoscape (version 2.8.2) and the transcriptional module networks were constructed using the yFiles Organic layout.

Results

A mutation in *ABI3* markedly reduces the ability to re-establish DT

All *abi* mutant plants grew similarly to the Col-0 plants and produced DT seeds. To determine whether the mutations in the *ABI* genes allowed re-establishment of DT in germinated seeds at the RP stage, we compared the differences in seedling formation rate after 72h of dehydration followed by 5 days of rehydration (Figure 1). Seedling formation for *abi3-8*, *abi3-9*, *abi3-10* and *abi4-3*, but not for *abi5-7*, was significantly compromised compared with Col-0 seeds (Figure 1) with the most severe phenotype for *abi3-9*. The seedling formation rate of *abi3-9* at RP after 72h of dehydration and rehydration was comparable with seedling formation rate of Col-0 at RH.

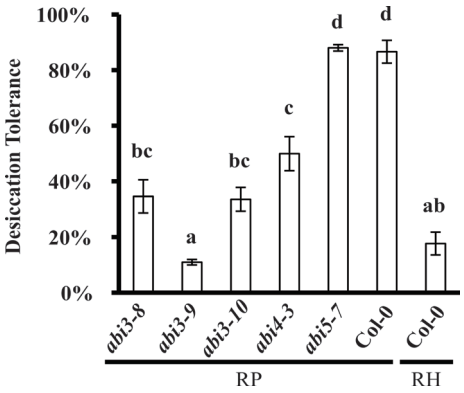


Figure 1: Re-establishment of desiccation tolerance in germinated seeds of wild-type (Col-0) and *abi* mutants of *Arabidopsis thaliana*. Different lower case letters above bars represent differences between samples by Tukey's HSD ($P<0.05$). Developmental stages : RP- radicle protrusion and RH-root hair stage.

Genes differentially regulated by early- and late-responses sustain the re-establishment of desiccation tolerance in germinated seeds of Col-0

To examine the role of *ABI3* in the transcriptional network in response to dehydration and re-establishment of DT in germinated seeds, we used a microarray approach. Considering that in Arabidopsis Col-0, DT at RH stage is only re-established in about 20% of the seeds and at RP in about 100% (Figure 1), these seeds were used in an across dehydration intervals in a further attempt to unravel the mechanism of re-establishment of DT. However, when employing only Col-0 seeds at RP and RH stages, only genes associated with the germination program rather than the program of re-establishment of DT are showing up. To exclude that, the *abi3-9* mutant was used at the RP stage, in which DT is re-established only in about 15% of the seeds (Figure 1). Also, to determine which mechanism are re-activated after dehydration Col-0 at RP stages after rehydration were used. The complete results derived from the microarray analysis are represented in Supporting Information Table S1. Principal component analysis (PCA) revealed the distribution among the different samples (Figure 2), and demonstrated that the two most important factors separating the different samples were developmental stage (RP and RH) and the treatment condition (duration of dehydration and rehydration).

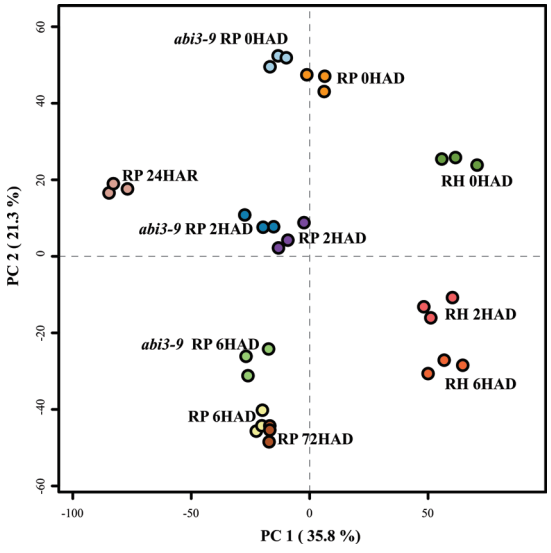


Figure 2: PCA plot of transcript abundance of dehydration time-points in wild-type (Col-0) seeds at the stages RP and RH and in *abi3-9* mutant seeds at the RP stage. Developmental stages: RP - radicle protrusion and RH -root hair. Time points: 0HAD- 0 hours after dehydration; 2HAD- 2 hours after dehydration; 6HAD-6 hours after dehydration; 72HAD-72 hours after dehydration and 24HAR-24 hours after rehydration.

When considering the time component, considerable variation could be observed (Figure 3). At RP without dehydration stress (0HAD) 39 genes were up-regulated and 11 down-regulated in Col-0 seeds as compared to *abi3-9* seeds, showing that at the stage of radicle protrusion there is little difference between the transcriptomes of *abi3-9* and Col-0. The number of transcripts changing between the RP and RH stages in Col-0, demonstrates the differences among these developmental stages, and shows that for the RP stage 202 genes were up-regulated and 714 down-regulated as compared to seeds at RH. To better understand the role of *ABI3* and other genes in the re-establishment of DT, we hereafter focused on differences in gene expression induced by dehydration between *abi3-9* (RP) and Col-0 (RP and RH).

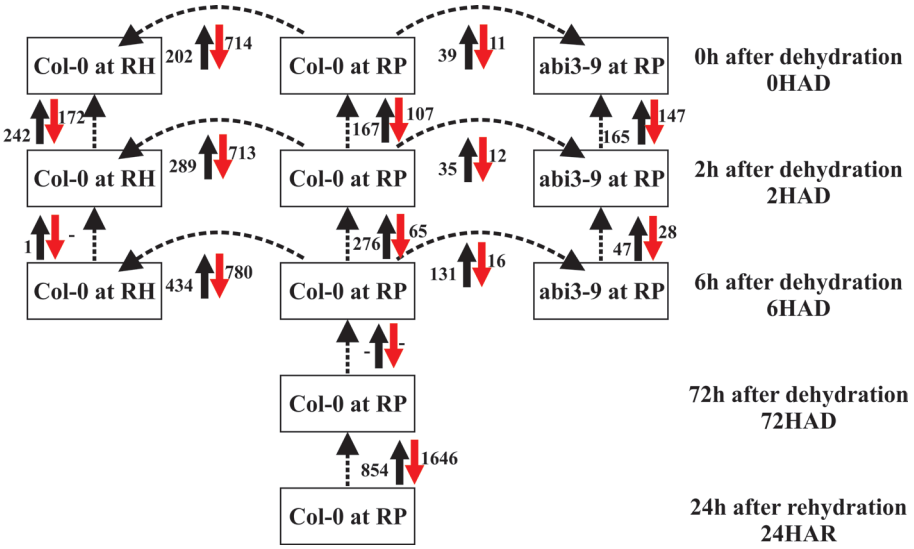


Figure 3: Overview of the comparisons made among the different samples. Arrows in black represent the number of up-regulated genes, whereas red represent the number of down-regulated genes. Dashed arrows indicated how the comparisons were performed. Developmental stages: RP - radicle protrusion and RH -root hair. Time points: 0HAD- 0 hours after dehydration; 2HAD-2 hours after dehydration; 6HAD-6 hours after dehydration; 72HAD-72 hours after dehydration and 24HAR-24 hours after rehydration.

Early response. After 2h of dehydration five comparisons were made, of which three were between two time-points (0HAD and 2HAD) in the same genetic background and developmental stage (Figure 3). At 2HAD of the RP stage in Col-0 167 genes were up-regulated. The number of transcripts accumulating in *abi3-9* after 2HAD from the RP and Col-0 at the RH stage was 165 and 242, respectively. Besides

the comparison between 0HAD and 2HAD in the same genetic background and in the same developmental stage, we also made comparisons between samples at the same early interval (2HAD). The number of differentially expressed genes upon 2h of dehydration between Col-0 and *abi3-9* at RP was 35 up-regulated in Col-0 and 12 up-regulated in *abi3-9* seeds. When the comparison was made between the RP and RH stage in Col-0, the number of genes changing in expression was 713 up-regulated at RH and 289 up-regulated for RP. To better understand the similarity in early responses with respect to the transcriptomes, we examined the overlap of transcripts which accumulated upon 2HAD in Col-0 at RP and RH, and in *abi3-9* at RP (Figure 4A). Of the genes up-regulated upon 2HAD, 107 were up-regulated both in Col-0 and *abi3-9* at RP, and 67 were up-regulated in Col-0 at both RP and RH. Of these, 57 genes were shared between all three samples; Col-0 at RP and RH and *abi3-9* at RP. Taken together, these results suggest that the 50 genes up-regulated exclusively in Col-0 at RP play a crucial role in the re-establishment of DT in response to dehydration in germinated Arabidopsis seeds.

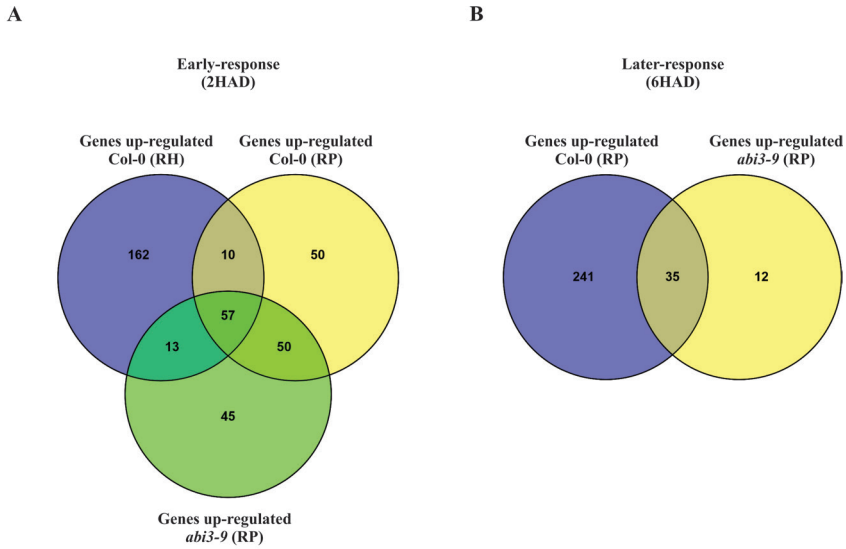


Figure 4: Venn diagram of genes in early-response (**A**) and late-response data sets (**B**). Developmental stages: RP - radicle protrusion and RH -root hair. Time points: 2HAD-2 hours after dehydration and 6HAD-6 hours after dehydration.

Late response. Also for later responses to dehydration (6HAD) five comparisons were made in a similar way as for 2HAD; three were between two intervals (2HAD and 6HAD) in the same genetic background and developmental stage. Among these comparisons, the number of up-regulated genes (276) upon 6h

of dehydration was considerably higher than the number of down-regulated genes (65) in Col-0 at RP (Figure 3). The number of transcripts accumulating for *abi3-9* at RP after 6h of dehydration was 47, whereas in Col-0 at RH, expression of only one gene had changed. Comparing Col-0 with *abi3-9* after 6h of dehydration, we observed 131 genes up-regulated in Col-0 and 16 up-regulated in *abi3-9* (Figure 3). Comparing the RP and RH stages in Col-0, the number of transcripts changing in abundance was 780 up-regulated genes at RH and 434 at RP. We checked for overlap between transcripts which were up-regulated upon 6HAD in Col-0 and in *abi3-9* at RP (Figure 4B). Col-0 at RH stage was not included in this analysis since it only had one differentially expressed gene. Different from what was found for the early response, the majority of genes up-regulated in *abi3-9* during this late response overlapped with up-regulated genes found in Col-0. But the number of genes differentially expressed in Col-0 at RP (241) was substantially higher than in *abi3-9* (12). The 241 up-regulated genes that are unique for Col-0 (Figure 4B) may be the ones responsible for the late response to dehydration that is required for the re-establishment of DT.

To assess whether prolonging the dehydration time would result in additional differences in gene expression, we measured gene expression in Col-0 at RP after 72h of dehydration. Interestingly, no differences in transcript abundance were observed at 72HAD, as compared with 6HAD (Figure 3). However, when comparing 24h of rehydration (24HAR) with 72HAD, a large number of transcripts was differentially expressed: 1.646 up-regulated in 72HAD and 854 up-regulated in 24HAR (Figure 3). These results are in agreement with the PCA (Figure 2), in which 6HAD and 72HAD were close to each other and at considerable distance from 24HAR.

Sets of co-expressed genes underline principal mechanisms of DT

To determine how transcript abundance changed over dehydration time, we clustered genes from all intervals and stages to identify seven dominant expression patterns (DPs) (Figure 5A; Supplemental Table S2), using the *Fuzzy K-Means* clustering method (Belostotsky *et al.*, 2009). Three out of seven of these DPs consisted of genes commonly up- or downregulated by dehydration (DP1, DP2, and DP6) showing similar trends, either for stages or genetic background. DP1 and DP2 displayed an increase and DP6 a decrease in transcript abundance upon dehydration. Gene expression levels of respective stages in Col-0 (RP and RH), and genetic background (Col-0 and *abi3-9* at RP) were rather similar. In addition to the trend of decreasing gene expression upon dehydration, DP6 showed a slight increase in gene expression after rehydration, suggestive of a re-activation of the (seed germination)

developmental genetic program. In contrast, four DPs consisted of genes with high expression at each time-point and stage (DP3, DP4, DP5, and DP7) with no difference among intervals across the dehydration period. However, these four DPs showed differences at 24HAR; DP3 and DP7 displayed increased transcript abundance whereas DP4 and DP5 showed a decrease. These variations in transcript abundance across dehydration time-points (DP1, DP2 and DP6), and between stages (DP3 and DP5) illustrate the complexity of the gene regulatory networks that operate during dehydration in germinated seeds.

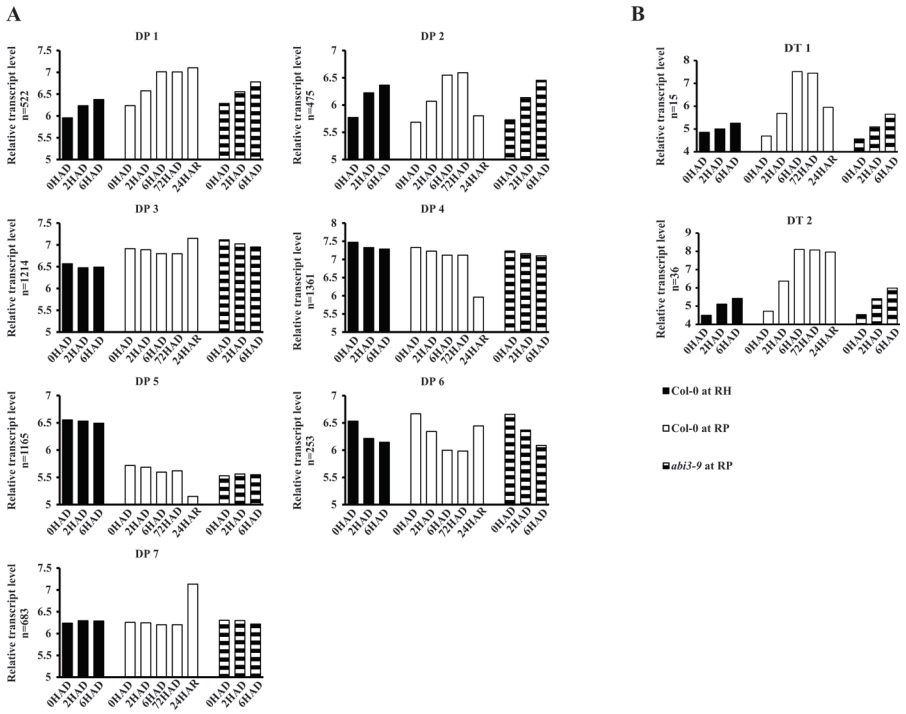


Figure 5: Dominant patterns of gene expression across dehydration time-points. Seven DPs were found using Fuzzy K-means clustering of the 50% most variant transcripts. Bar graphs represent averages of expression levels of transcripts (left to right, black – Col-0 at RH, white – Col-0 at RP, and dashed – *abi3-9* at RP).Developmental stages: RP - radicle protrusion and RH -root hair. Time points: 0HAD- 0 hours after dehydration; 2HAD-2 hours after dehydration; 6HAD-6 hours after dehydration; 72HAD-72 hours after dehydration and 24HAR-24 hours after rehydration.

Because in *abi3-9* DT is hardly re-established in germinated seeds just as for Col-0 at RH (Figure 1), we identified transcripts that were specifically highly expressed upon dehydration in Col-0 at RP (Figure 5B) since these seeds show

full re-establishment of DT. We identified two sets of DT-specific transcripts with statistically significant (P <0.01), two-fold or higher levels of expression at 6HAD, 72HAD, and/or 24HAR relative to all others at the same time-point (Figure 5B). These data sets consist of relatively few genes with DT-specific expression, later during dehydration. Only 15 genes were highly expressed at 6HAD and 72 HAD (DT1), and 36 were highly expressed at 6HAD, 72HAD and 24HAR (DT2) compared with seeds in which DT could not be re-established. Many of these DT-specific genes were present in DP1 or DP2 (Figure S1). This suggests that a common set of genes is coordinately up-regulated (DP1 and DP2) or down-regulated (DP6) during dehydration and that these promote adaptation to the desiccation stress and are responsible for survival in the dry state.

Functional groups of desiccation-inducible genes

To understand the functional implication of the desiccation responsive genes, we identified significantly enriched (P < 0.001, hypergeometric distribution) GO terms, DNA binding motifs, and (P < 0.05) metabolic pathways for early and late responses, DPs that showed patterns of gene expression induced by dehydration (DP1, DP2, and DP6), as well as the DT-specific gene sets (Supplemental Table S2). This analysis revealed that many genes with associated GO terms for *abscisic acid* and *water deprivation response* were among the commonly up-regulated genes in DP1 and DP2 (Figure 6A). In fact, all enriched GO terms for all analysed gene sets were somehow associated with dehydration stress. Although ten DNA binding motifs were identified in these data sets (Supplemental Table S2), we could not separate data sets based on specific DNA binding motifs but we observed that all DNA binding motifs enriched in these data sets have been associated previously with abiotic stress inducible genes such as the ABRE-, ABF-, and CBF- binding motifs (Fujita et al., 2013; Vysotskii et al., 2013). To investigate whether these data sets would possibly differ in other aspects, a metabolic pathway analysis was performed (Figure 6B). Interestingly, metabolite pathways such as *trehalose* and *sucrose biosynthesis* were enriched in DP1, whereas *abscisic acid* and *carotenoid biosynthesis*, and *methionine degradation* were enriched in DP2. Furthermore, *sucrose biosynthesis* was also enriched in the early response data set. The late-response data set was highly enriched for *abscisic acid biosynthesis*. These results suggest that ABA starts to plays a role only at later stages of the re-establishment of DT. It corroborates with high peak of ABA accumulation at 6HAD (Chapter 4).

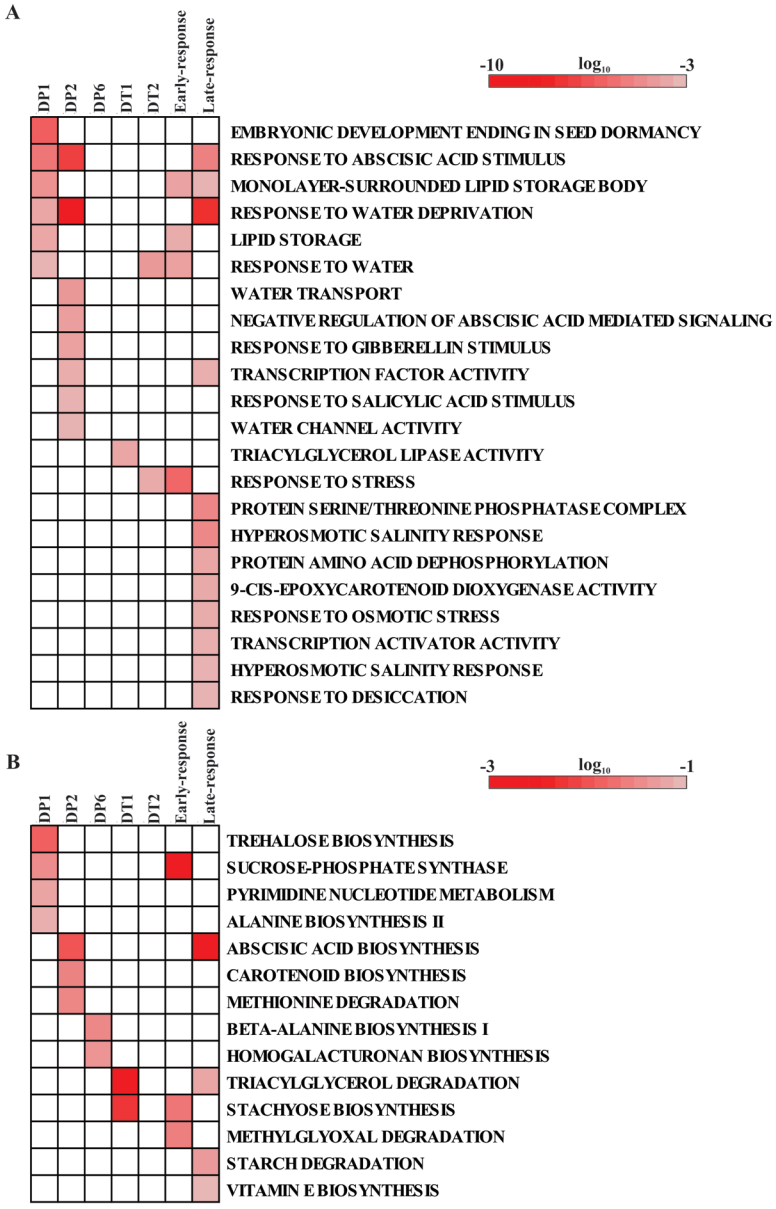


Figure 6: Heat map showing the *P* value significance of enrichment of (**A)** GO terms and (**B**) metabolite pathways for early-term and late-term responses for DPs that showed pattern of genes induced by dehydration (DP1, DP2, and DP6), and DT-specific DPs (Supplemental Table S2).

Gene co-expression networks and transcriptional modules reveal regulatory genes possibly controlling the re-establishment of DT

Gene co-expression networks have been used to identify gene interactions relevant to specific biological processes, such as the acquisition of DT in *Medicago truncatula* during seed development (Verdier *et al.*, 2013) and the transition from dormancy to germination in Arabidopsis (Bassel *et al.*, 2011). Thus, to identify regulatory processes that are controlling the re-establishment of DT, we built a gene co-expression network for this process in germinated Arabidopsis seeds. An unweighted gene co-expression network was constructed using Cytoscape® 2.8.2. Edge adjacency threshold was set at 0.94 as described by Villa-Vialaneix *et al.* (2013), resulting in a network consisting of 2.332 genes represented as nodes, which are connected by 33.476 edges. Each of the data sets induced (either up- or down-regulated genes) by dehydration in germinated seeds of Col-0 and *abi3-9*, could be identified on the DT co-expression network (Figure 7).

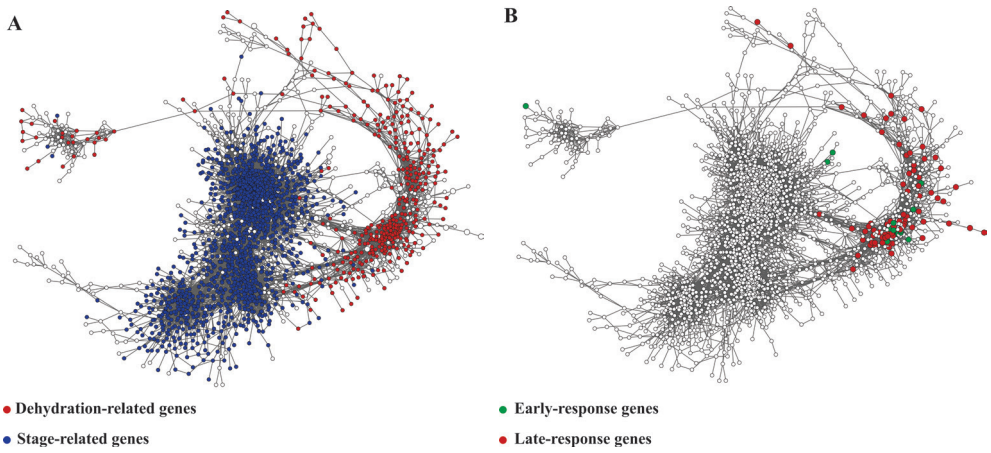


Figure 7: DT gene co-expression network. Nodes are coloured according to (**A**) genes associated with dehydration (red) and developmental stages (blue) and (**B**) according to genes associated with early-response (green) and late-response transcripts (red).

Two distinct regions of the network were found to correspond with genes responding to dehydration and developmental stages, respectively (Figure 7A). These two regions are also related to the seven observed DPs (Figure 5A and Figure S2). Genes related to dehydration stress are present in DP1 and DP2 (Figure S2), whereas genes which are associated with developmental stage are present in DP3, DP4, DP5, and DP7. Genes of DP1 clustered clearly separately from the developmental stage cluster, whereas genes in DP2 were adjacent to it (Figure S2). Besides DPs, we also highlighted early- and late-response genes on the network (Figure 7B). The few

genes present in the early-response data set were clustered close to each other in the DP2 region (Figure 7B and Figure S2), but the late-response genes were spread over the whole network region of the dehydration-related genes (DP1 and DP2).

ABI3 is an essential component of the regulatory network controlling the maturation of Arabidopsis seeds (Mönke *et al.*, 2012), which is the developmental stage at which DT is acquired. We projected on the network the members of the previously identified ABI3 regulon activated by VP1, FUS3 and LEC2 (Mönke *et al.*, 2012). The ABI3 regulon members are present in the DP2 cluster region rather than in the DP1 cluster region within the network region of the dehydration-related genes (Figure 8A). To better visualize the regulatory network of the ABI3 regulon on our network, we separated the ABI3 regulon and genes connected with its members, creating a DT sub-network with 238 genes (nodes) connected by 2.932 edges (Figure 8B). Since we were interested in identifying regulators that control the re-establishment of DT in germinated seeds, we combined the physiological process with the transcriptome data. Considering that the ability of germinated Arabidopsis seeds to survive the dry state was restricted to Col-0 at the RP stage, as compared with Col-0 at RH and *abi3-9* at RP (Figure 1), we projected the DT-specific data sets on the DT gene co-expression network, as well as on the sub-network. We also plotted genes up-regulated in *abi3-9* at RP upon dehydration on the network and on the DT sub-network. DT-specific genes, as well as *abi3-9* up-regulated genes were present on the dehydration-related gene region on the network (Figure S3). However, when these genes were plotted on the DT sub-network, none of the *abi3-9* up-regulated genes were found (Figure 8B) implying that these genes are not playing a role in the re-establishment of DT.

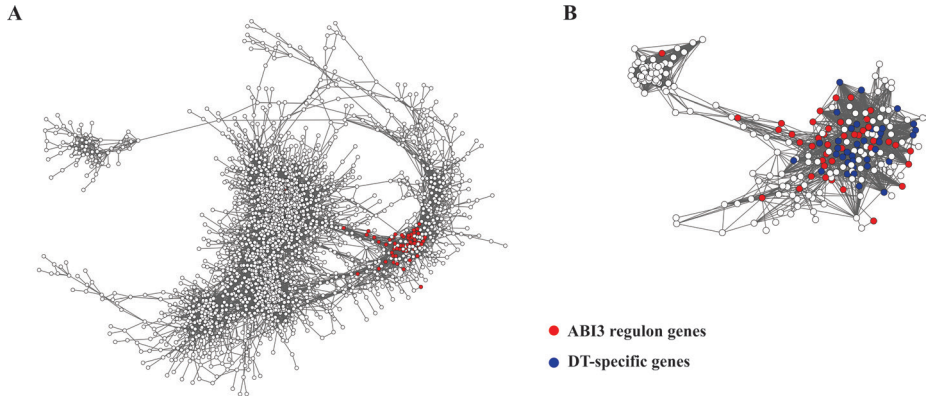


Figure 8: (**A**) DT gene co-expression network and (**B**) sub-network created from the ABI3 regulon presented in the gene co-expression network and genes connected with its members. Nodes are coloured in red for ABI3 regulon members in the DT gene co-expression network and in the sub-network, and in blue for genes of DT-specific data sets in the sub-network.

To identify the most important regulators for re-establishment of DT, we built transcriptional modules according to Belmonte *et al.* (2013) for genes present in the DT sub-network. These transcriptional modules consist of desiccation-inducible genes containing ABRE, ABF, DPBF1&2, G-box, and CBF sequence motifs and putative transcription factors involved in DT, such as ABI5 (Verdier *et al.*, 2013) and the ABA-gene regulator INDUCER OF CBF EXPRESSION 1 (A) (Liang and Yang, 2015) (Figure 9). Besides these motifs, we found 14 genes with an RY-repeat promoter motif (Supplemental Table S2). The RY element is crucial for transactivation through ABI3 (Ezcurra *et al.*, 2000). Taken together, this approach identified important regulators of the establishment of DT.

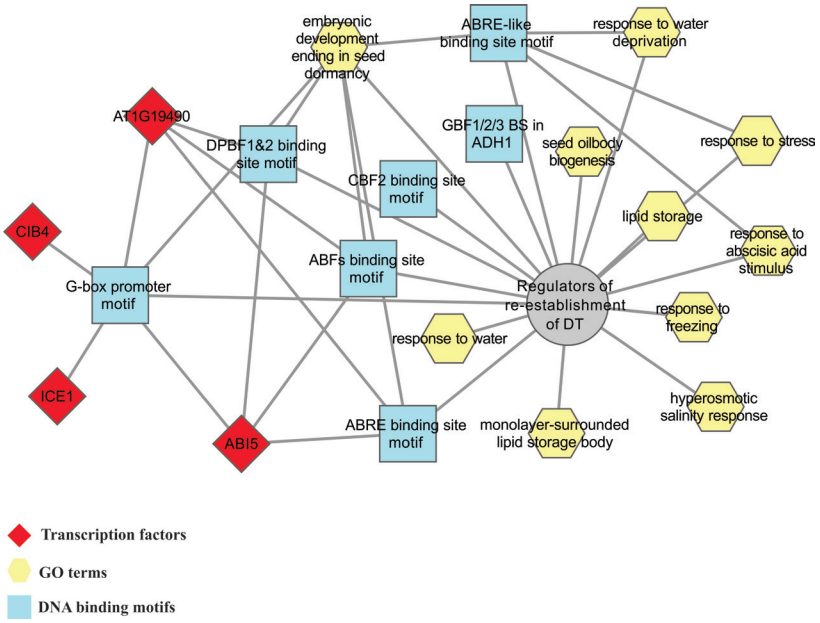


Figure 9: Transcriptional module predicted to regulate the re-establishment of DT in Arabidopsis germinated seeds. DNA motifs (light-blue rectangles) and GO terms (light-yellow octagons) that are significantly overrepresented ($P<0.001$) within the sub-network gene sets (ellipse) together with co-expressed transcription factors (red diamonds). Transcriptional modules were predicted for early-response, late-response, DP1 and DP2 and DT-specific (Figure S4).

Discussion

Desiccation tolerance (DT) is defined as the ability of an organisms to survive water levels as low as 0.1 gram of H_2O per gram of dry weight and subsequent re-hydration without lethal damage (Hoekstra *et al.*, 2001; Dekkers *et al.*, 2015). To cope with environmental stresses, such as (severe) dehydration, plants employ a signal transduction network in which diverse transcription factors control the expression of large numbers of genes (Yamaguchi-Shinozaki and Shinozaki, 2006; Verdier *et al.*, 2013). Moreover, stress responses include transcriptional regulation of gene expression and this depends on the interaction of transcription factors with *cis*-regulatory sequences (Maruyama *et al.*, 2012). The *cis*-elements of dehydration- and ABA-responsive genes have been widely studied (Narusaka *et al.*, 2003; Kim *et al.*, 2004; Nakashima *et al.*, 2006). ABA-dependent pathways are mediated by gene expression through ABA-responsive elements (ABREs) and b-zip transcription factors (Busk and Pages, 1998; Kagaya *et al.*, 2002). One of these b-zip transcription factors is ABA INSENSITIVE 5 (ABI5), which is involved in ABA signalling during seed maturation and germination and in the regulation of stress responses (Bensmihen *et al.*, 2002). ABI5 is a transcriptional activator that binds to G-box and ABREs and the gene encoding ABI5 is strongly expressed in desiccation tolerant tissues in Arabidopsis (Drea *et al.*, 2006). Additionally, in a gene co-expression network during seed maturation of *Medicago truncatula,* also *ABI3* and *ABI4* are highly connected to DT-associated genes (Verdier *et al.*, 2013). Given that these three transcription factors are important for DT, but not sufficient, we predicted that other transcription factors may be involved in the re-establishment of DT in Col-0 at the RP stage.

We found that a mutation in *ABI3* (*abi3-9*) (Nambara *et al.*, 2002) clearly reduced the re-establishment of DT in germinated seeds at RP. This result corroborates with a previous study of re-establishment of DT in germinated seeds governed by osmotic stress (Maia *et al.*, 2014). Our results with the *abi* mutants show that re-establishment of DT in germinated seeds is largely mediated by ABA signalling pathways. However, the role of *ABI3* in the re-establishment of DT remains elusive. Using a genome-wide transcriptome data set for Col-0 at RP and RH, and *abi3-9* at RP across a range of dehydration levels, we provide the most comprehensive description of gene activity, so far, during the re-establishment of DT in Arabidopsis.

Although DT cannot be re-established in the *abi3-9* mutant at RP and Col-0 at RH, our data show that they share many of the transcripts that are induced in Col-0 at RP by dehydration, and which does result in DT (DP1, DP2 and DP6). However, the functional differentiation among samples occurs through at least four distinct sets

of processes (early- and late- responses, and two DT-specific data sets). First, genes expressed specifically in Col-0 at RP (early-response) appear to play a role in events which may promote the initial protection to dehydration stress. The transcriptional module for the early-response shows the transcription factor ABSCISIC ACID RESPONSIVE ELEMENT-BINDING FACTOR 1 (*ABF1*) to be connected to ABRE, ABF and G-box binding motifs. This hypothesis of initial protection is supported by studies on drought, in which the addition of a mutation in *ABF1* to the *cis*-regulatory elements of AREB/ABF triple mutant *areb1 areb2 abf3,* resulted in a quadruple mutant (*areb1 areb2 abf3 abf1*) which displayed increased sensitivity to drought and decreased sensitivity to ABA, as compared with the triple mutant (Yoshida *et al.*, 2015). Second, genes expressed specifically in Col-0 at RP (late-term) are associated with GO terms such as *desiccation response* and DNA binding motifs such as ABFs, ABREs, DREs and GBFs. Besides the DNA binding motifs, this set of genes also included TFs such as CIRCADIAN 1 (*CIR1*), *A*LTERED SEED GERMINATION 4 (*ASG4*), POPEYE (*PYE*) and ABA-RESPONSIVE KINASE SUBSTRATE 1 (*ASK1*). These TFs are known to function in growth regulation (*ASG4* and *PYE*) (Martínez-Trujillo *et al.*, 2014), light signalling pathways (*ASK1*) (Takahashi *et al.*, 2013), and the circadian clock (*CIR1*). Taken together, it suggests that the regulators of the late-response may play a role in the ability of germinated seeds to adapt to changes in water dynamics in a coordinated response with the circadian clock (Takase *et al.*, 2011). It corroborates a study which showed that clock genes may control the expression of stress-responsive genes (Habte *et al.* 2014). Moreover, the late-response data set shows genes such as RESPONSIVE TO DESSICATION 29A (*RD29A*) and LATE EMBRYOGENESIS ABUNDANT 14 (*LEA14*) which are present in the GO term *response to desiccation*. A transcriptome data analysis of the *areb1 areb2 abf3* triple mutant demonstrated that LEAs and group-Ab PP2Cs genes were down regulated in response to water stress (Yoshida *et al.*, 2010). Additionally, the role of LEAs during further dehydration suggested that they may provide a layer of hydroxylated residues to interact with the surface groups of other proteins, acting as replacement of water (Shewry *et al.*, 1999) or even contribute to the formation of cytoplasmic glass to protect the cells (Hoekstra *et al.*, 2001). Thus, we conclude that the late-response genes promote adaptation to the desiccated state.

The two dominant patterns, DP1 and DP2, include genes commonly induced by dehydration. However, we identified within these DPs 15 genes up-regulated at 6HAD and 72HAD and 36 up-regulated at 6HAD, 72HAD and 24HAR, suggesting that these gene sets are specific for the re-establishment of DT in germinated Col-0 seeds. For example, consistent with this is that the transcriptional module for DT-specific genes is enriched with DNA binding motifs known to interact with genes

induced by water stress such as ABF, ABRE and CBF binding motifs (Narusaka *et al.*, 2003; Yoshida *et al.*, 2010; Fujita *et al.*, 2013; Vysotskii *et al.*, 2013).

A central argument that emerges from this comprehensive profiling of transcript populations is that there is overlap in the gene expression programs that characterize drought-tolerance and acquisition of desiccation tolerance during seed development. Although Col-0 at RP and RH, and *abi3-9* at RP express transcripts that accumulate in each sample and interval, a global comparison demonstrated unexpected similarities between *abi3-9* and Col-0. These similarities result, at least in part, from the large number of dehydration-inducible genes in *abi3-9* that are co-expressed with Col-0. Contradictory with these common responses in gene activity, is that the DT-specific set is exclusively expressed in Col-0 at RP. This suggests substantial coordination of re-establishment of DT in Col-0 at RP through this DT-specific gene set.

The re-establishment of DT in germinated seeds seems to be comparable with the natural acquisition of desiccation tolerance during seed development (Maia *et al.*, 2011; Terrasson *et al.*, 2013). Therefore, to address the control of re-establishment of DT we used the ABI3 regulon data set derived from seed developmental studies (Mönke *et al.*, 2012). ABI3, FUS3 and LEC2 are B3 transcription factors which interact with each other to control seed developmental processes such as storage reserve accumulation, as well as the acquisition of dormancy and DT (To *et al.*, 2006; Santos-Mendoza *et al.*, 2008). Mutations in these TFs, such as *fus3-3* (Tiedemann *et al.*, 2008) and *abi3* (Ooms *et al.*, 1993) result in desiccation sensitive seeds. A recent study on the re-establishment of DT in germinated seeds of Arabidopsis has shown that germinated *abi3-9* seeds lack the possibility to re-establish DT and that they remain desiccation sensitive after an osmotic treatment (Maia *et al.*, 2014). Our results show that on the DT gene co-expression network, genes of the ABI3 regulon are highly connected with our DT-specific data sets. Nevertheless, the transcriptional module for the DT sub-network genes includes TFs that are both ABA-dependent (*ABI5* and AT1G19490) and ABA–independent (CRY2-INTERACTING BHLH 4 (*CIB4*) and *ICE1*). Taken together with the DT-phenotype of *abi5-7* (Figure 1), it suggests that the other TFs present in the regulatory system for the re-establishment of DT act redundantly to *ABI5*. The ABA-independent ICE1 and CIB4, together with the ABA-dependent At1g19490 are good candidates to interact with ABI5. ICE1 was initially identified as a transcription factor that directly controls the transcription of *CBF3* (Chinnusamy *et al.*, 2007). The constitutive expression of CBF3/DREB1A in rice enhanced tolerance to drought (Oh *et al.*, 2005). Moreover, the expression of many CRY-regulated genes, to which the ABA-independent CIB4 belongs, are regulated by other signaling pathways. These pathways are those related to phytochromes and phytohormones, which suggests that the CRY-dependent genes

are intimately integrated with the general regulatory networks that control plant development (Liu *et al.*, 2011). This result may extend to *CIB4* induction, which could increase the signals of phytochromes and phytohormones, thus promoting the arrest of seed germination at the RP stage.

Furthermore, *ABI3* and *ABI4* were highly expressed in the *ice1-2* mutant as compared with wild type, in response to ABA, suggesting that *ICE1* acts as a negative regulator of ABA-dependent responses (Liang and Yang, 2015).

Taken together, this has led to a model with the possible role of these TFs in the ABA-dependent and -independent pathways, controlling the re-establishment of DT in germinated seeds (Figure 10). We demonstrate that the data set can be used to identify sets of genes involved in the re-establishment of DT in germinated seeds. ABI3, ABI4 and ABI5 play an important role in the establishment of DT during seed development (Verdier *et al.*, 2013). We demonstrate that, to a significant extent, they may also play similar important roles in the re-establishment of DT. Under dehydration stress, the transcriptional profiles of *abi3-9*, were useful to filter out the genes that are not the main regulators of DT. A transcriptional module with ABA-dependent and –independent transcription factors was identified and may serve as a convergence point in the crosstalk between these two pathways during the re-establishment of DT in Arabidopsis. It is clear that ABA exerts a profound influence on the re-establishment of DT *via ABI3*, *ABI4* and *ABI5*, but our data also suggests the existence of an ABA-independent pathway – with extensive cross-talk with the ABA-dependent pathway - *via ICE1* and *CIB4*.

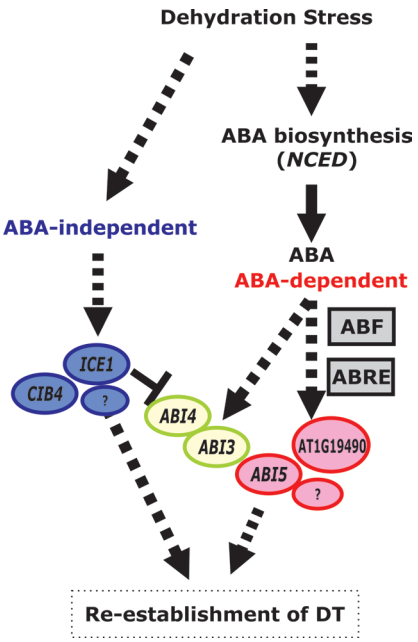


Figure 10: A proposed model for the regulators associated with re-establishment of DT, showing a putative cross-talk between ABA-dependent (ellipses in red) and independent (ellipses in blue) transcription factors. DNA binding motifs (rectangle in grey). Ellipses in green shows two ABA-dependent transcription factors (ABI3 and ABI4), these are not present on the transcriptional module (Figure 9). However, it is assumed that *ICE1* may act as an inhibitor of *ABI3* and *ABI4* (Liang and Yang, 2015).

Acknowledgements

This work was funded by "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior" (CAPES, Brazil).

References

Bartels D, Hussain SS (2011) Resurrection plants: physiology and molecular biology. *In* U Lüttge, E Beck, D Bartels, eds, Desiccation tolerance in plants, Ecological Studies. Springer, Heidelberg, pp 339-364

Bassel GW, Lan H, Glaab E, Gibbs DJ, Gerjets T, Krasnogor N, Bonner AJ, Holdsworth MJ, Provart NJ.2011 Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. *Proceedings of the National Academy of Sciences* **108:** 9709-9714

Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, Hashimoto M, Fei J, Harada CM, Munoz MD, Le BH, Drews GN, Brady SM, Goldberg RB, Harada JJ.2013 Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. *Proceedings of the National Academy of Sciences* **110:** E435-E444

Belostotsky DA, Gadbury G, Garrett K, Allison D (2009) Challenges and Approaches to Statistical Design and Inference in High-Dimensional Investigations. *In* Plant Systems Biology, Vol 553. Humana Press, pp 181-206

Belostotsky DA, Orlando D, Brady S, Koch J, Dinneny J, Benfey P (2009) Manipulating Large-Scale Arabidopsis Microarray Expression Data: Identifying Dominant Expression Patterns and Biological Process Enrichment. *In* Plant Systems Biology, Vol 553. Humana Press, pp 57-77

Bensmihen S, Rippa S, Lambert G, Jublot D, Pautot V, Granier F, Giraudat J, Parcy F.2002 The homologous ABI5 and EEL transcription factors function antagonistically to fine-tune gene expression during late embryogenesis. *Plant Cell* **14:** 1391-1403

Brady SM, Orlando DA, Lee J-Y, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN.2007 A High-Resolution Root Spatiotemporal Map Reveals Dominant Expression Patterns. *Science* **318:** 801-806

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

Busk PK, Pages M.1998 Regulation of abscisic acid-induced transcription. *Plant Mol Biol* **37:** 425-435

Chinnusamy V, Zhu J, Zhu J-K.2007 Cold stress regulation of gene expression in plants. *Trends in Plant Science* **12:** 444-451

Dekkers BJW, Pearce S, van Bolderen-Veldkamp RP, Marshall A, Widera P, Gilbert J, Drost H-G, Bassel GW, Müller K, King JR, Wood ATA, Grosse I, Quint M, Krasnogor N, Leubner-Metzger G, Holdsworth MJ, Bentsink L.2013 Transcriptional

Dynamics of Two Seed Compartments with Opposing Roles in Arabidopsis Seed Germination. *Plant Physiology*

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

Drea SC, Lao NT, Wolfe KH, Kavanagh TA.2006 Gene duplication, exon gain and neofunctionalization of OEP16-related genes in land plants. *Plant J* **46:** 723-735

Ezcurra I, Wycliffe P, Nehlin L, Ellerstrom M, Rask L.2000 Transactivation of the Brassica napus napin promoter by ABI3 requires interaction of the conserved B2 and B3 domains of ABI3 with different cis-elements: B2 mediates activation through an ABRE, whereas B3 interacts with an RY/G-box. *Plant J* **24:** 57-66

Frank W, Phillips J, Salamini F, Bartels D.1998 Two dehydration-inducible transcripts from the resurrection plant Craterostigma plantagineum encode interacting homeodomain-leucine zipper proteins. *Plant J* **15:** 413-421

Freeman TC, Goldovsky L, Brosch M, van Dongen S, Mazière P, Grocock RJ, Freilich S, Thornton J, Enright AJ.2007 Construction, Visualisation, and Clustering of Transcription Networks from Microarray Expression Data. *PLoS Computational Biology* **3:** e206

Fujita Y, Yoshida T, Yamaguchi-Shinozaki K.2013 Pivotal role of the AREB/ABF-SnRK2 pathway in ABRE-mediated transcription in response to osmotic stress in plants. *Physiologia Plantarum* **147:** 15-27

Gaff DF, Oliver M.2013 The evolution of desiccation tolerance in angiosperm plants: A rare yet common phenomenon. *Functional Plant Biology* **40:** 315-328

Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S, Smyth GK (2005) limma: Linear Models for Microarray Data. *In* Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer New York, pp 397-420

Habte E, Müller LM, Shtaya M, Davis SJ, Von Korff M.2014 Osmotic stress at the barley root affects expression of circadian clock genes in the shoot. *Plant, Cell and Environment* **37:** 1321-1337

Hoekstra FA, Golovina EA, Buitink J.2001 Mechanisms of plant desiccation tolerance. *Trends Plant Sci* **6:** 431-438

Irizarry RA, Hobbs B, Collin F, Beazer, Barclay YD, Antonellis KJ, Scherf U, Speed TP.2003 Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4:** 249-264

Kagaya Y, Hobo T, Murata M, Ban A, Hattori T.2002 Abscisic Acid–Induced Transcription Is Mediated by Phosphorylation of an Abscisic Acid Response Element Binding Factor, TRAB1. *The Plant Cell* **14:** 3177-3189

Kim S, Kang JY, Cho DI, Park JH, Soo YK.2004 ABF2, an ABRE-binding bZIP factor, is an essential component of glucose signaling and its overexpression affects multiple stress tolerance. *Plant Journal* **40:** 75-87

Koster KL.1991 Glass Formation and Desiccation Tolerance in Seeds. *Plant Physiology* **96:**

302-304

Liang CH, Yang CC.2015 Identification of ICE1 as a negative regulator of ABA-dependent pathways in seeds and seedlings of Arabidopsis. *Plant Mol Biol* **88:** 459-470

Liu H, Liu B, Zhao C, Pepper M, Lin C.2011 The action mechanisms of plant cryptochromes. *Trends in Plant Science* **16:** 684-691

Lopez-Molina L, Mongrand S, Chua N-H.2001 A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factor in Arabidopsis. *Proceedings of the National Academy of Sciences* **98:** 4782-4787

Maia J, Dekkers BJW, Dolle MJ, Ligterink W, Hilhorst HWM.2014 Abscisic acid (ABA) sensitivity regulates desiccation tolerance in germinated Arabidopsis seeds. *New Phytologist***:** n/a-n/a

Maia J, Dekkers BJW, Provart NJ, Ligterink W, Hilhorst HWM.2011 The Re-Establishment of Desiccation Tolerance in Germinated *Arabidopsis thaliana* Seeds and Its Associated Transcriptome. *PLos ONE* **6**

Martínez-Trujillo M, Méndez-Bravo A, Ortiz-Castro R, Hernández-Madrigal F, Ibarra-Laclette E, Ruiz-Herrera LF, Long TA, Cervantes C, Herrera-Estrella L, López-Bucio J.2014 Chromate alters root system architecture and activates expression of genes involved in iron homeostasis and signaling in Arabidopsis thaliana. *Plant Molecular Biology* **86:** 35-50

Maruyama K, Todaka D, Mizoi J, Yoshida T, Kidokoro S, Matsukura S, Takasaki H, Sakurai T, Yamamoto YY, Yoshiwara K, Kojima M, Sakakibara H, Shinozaki K, Yamaguchi-Shinozaki K.2012 Identification of cis-acting promoter elements in cold- and dehydration-induced transcriptional pathways in arabidopsis, rice, and soybean. *DNA Research* **19:** 37-49

Masetto TE, Faria JM, Fraiz AC.2014 Re-induction of desiccation tolerance after germination of Cedrela fissilis Vell. seeds. *Anais da Academia Brasileira de Ciências* **86:** 1273-1286

Meurs C, Basra AS, Karssen CM, Van Loon LC.1992 Role of abscisic acid in the induction of desiccation tolerance in developing seeds of Arabidopsis thaliana. *Plant Physiology* **98:** 1484-1493

Mönke G, Seifert M, Keilwagen J, Mohr M, Grosse I, Hähnel U, Junker A, Weisshaar B, Conrad U, Bäumlein H, Altschmied L.2012 Toward the identification and regulation of the Arabidopsis thaliana ABI3 regulon. *Nucleic Acids Research* **40:** 8240-8254

Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K.2006 Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of Arabidopsis. *Plant Molecular Biology* **60:** 51-68

Nakashima K, Yamaguchi-Shinozaki K.2013 ABA signaling in stress-response and seed development. *Plant Cell Reports* **32:** 959-970

Nambara E, Suzuki M, Abrams S, McCarty DR, Kamiya Y, McCourt P.2002 A Screen for Genes That Function in Abscisic Acid Signaling in Arabidopsis thaliana. *Genetics*

161: 1247-1255

Narusaka Y, Nakashima K, Shinwari ZK, Sakuma Y, Furihata T, Abe H, Narusaka M, Shinozaki K, Yamaguchi-Shinozaki K.2003 Interaction between two cis-acting elements, ABRE and DRE, in ABA-dependent expression of Arabidopsis rd29A gene in response to dehydration and high-salinity stresses. *Plant J* **34:** 137-148

Narusaka Y, Nakashima K, Shinwari ZK, Sakuma Y, Furihata T, Abe H, Narusaka M, Shinozaki K, Yamaguchi-Shinozaki K.2003 Interaction between two cis-acting elements, ABRE and DRE, in ABA-dependent expression of Arabidopsis rd29A gene in response to dehydration and high-salinity stresses. *Plant Journal* **34:** 137-148

Oh SJ, Sang IS, Youn SK, Jang HJ, Soo YK, Kim M, Kim YK, Baek HN, Kim JK.2005 Arabidopsis CBF3/DREB1A and ABF3 in transgenic rice increased tolerance to abiotic stress without stunting growth. *Plant Physiology* **138:** 341-351

Ooms J, Leon-Kloosterziel KM, Bartels D, Koornneef M, Karssen CM.1993 Acquisition of Desiccation Tolerance and Longevity in Seeds of Arabidopsis thaliana (A Comparative Study Using Abscisic Acid-Insensitive abi3 Mutants). *Plant Physiology* **102:** 1185-1191

Potts M.2001 Desiccation tolerance: a simple process? *Trends Microbiol* **9:** 553-559

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

Shewry P, Casey R, Cuming A (1999) LEA Proteins. *In* Seed Proteins. Springer Netherlands, pp 753-780

Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T.2011 Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27:** 431-432

Takahashi Y, Ebisu Y, Kinoshita T, Doi M, Okuma E, Murata Y, Shimazaki KI.2013 BHLH transcription factors that facilitate k+ uptake during stomatal opening are repressed by abscisic acid through phosphorylation. *Sci Signal* **6**

Takase T, Ishikawa H, Murakami H, Kikuchi J, Sato-Nara K, Suzuki H.2011 The circadian clock modulates water dynamics and aquaporin expression in arabidopsis roots. *Plant and Cell Physiology* **52:** 373-383

Terrasson E, Buitink J, Righetti K, Ly Vu B, Pelletier S, Zinsmeister J, Lalanne D, Leprince O.2013 An emerging picture of the seed desiccome: confirmed regulators and newcomers identified using transcriptome comparison. *Front Plant Sci* **4:** 497

Tiedemann J, Rutten T, Monke G, Vorwieger A, Rolletschek H, Meissner D, Milkowski C, Petereck S, Mock HP, Zank T, Baumlein H.2008 Dissection of a complex seed phenotype: novel insights of FUSCA3 regulated developmental processes. *Dev Biol* **317:** 1-12

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

Verdier J, Lalanne D, Pelletier S, Torres-Jerez I, Righetti K, Bandyopadhyay K,

Leprince O, Chatelain E, Vu BL, Gouzy J, Gamas P, Udvardi MK, 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

Vieira CV, da Silva EAA, de Alvarenga AA, de Castro EM, Toorop PE.2010 Stress-associated factors increase after desiccation of germinated seeds of Tabebuia impetiginosa Mart. *Plant Growth Regulation* **62:** 257-263

Villa-Vialaneix N, Liaubet L, Laurent T, Cherel P, Gamot A, SanCristobal M.2013 The structure of a gene co-expression network reveals biological functions underlying eQTLs. *PLos ONE* **8:** e60045

Virlouvet L, Ding Y, Fujii H, Avramova Z, Fromm M.2014 ABA signaling is necessary but not sufficient for RD29B transcriptional memory during successive dehydration stresses in Arabidopsis thaliana. *Plant Journal* **79:** 150-161

Vysotskii DA, de Vries-van Leeuwen IJ, Souer E, Babakov AV, de Boer AH.2013 ABF transcription factors of Thellungiella salsuginea: Structure, expression profiles and interaction with 14-3-3 regulatory proteins. *Plant Signaling and Behavior* **8:** 64-70

Xiao L, Yang G, Zhang L, Yang X, Zhao S, Ji Z, Zhou Q, Hu M, Wang Y, Chen M, Xu Y, Jin H, Xiao X, Hu G, Bao F, Hu Y, Wan P, Li L, Deng X, Kuang T, Xiang C, Zhu JK, Oliver MJ, He Y.2015 The resurrection genome of Boea hygrometrica: A blueprint for survival of dehydration. *Proc Natl Acad Sci U S A* **112:** 5833-5837

Yamaguchi-Shinozaki K, Shinozaki K.2006 Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu Rev Plant Biol* **57:** 781-803

Yoshida T, Fujita Y, Maruyama K, Mogami J, Todaka D, Shinozaki K, Yamaguchi-Shinozaki K.2015 Four Arabidopsis AREB/ABF transcription factors function predominantly in gene expression downstream of SnRK2 kinases in abscisic acid signalling in response to osmotic stress. *Plant, Cell and Environment* **38:** 35-49

Yoshida T, Fujita Y, Sayama H, Kidokoro S, Maruyama K, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K.2010 AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. *Plant Journal* **61:** 672-685

Supporting information

If not displayed bellow, supplementary information can be downloaded from : http://www.wageningenseedlab.nl/thesis/atsilva/SI/

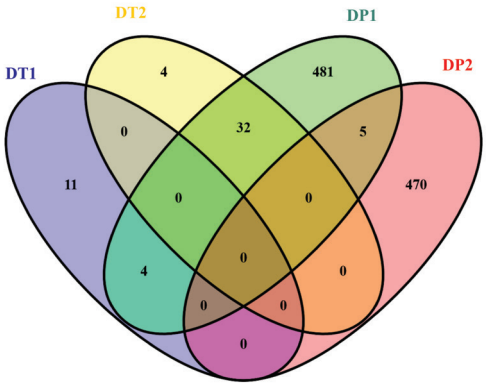


Figure S1: Venn diagram of common genes among DT-specific, DP1 and DP2 data sets.

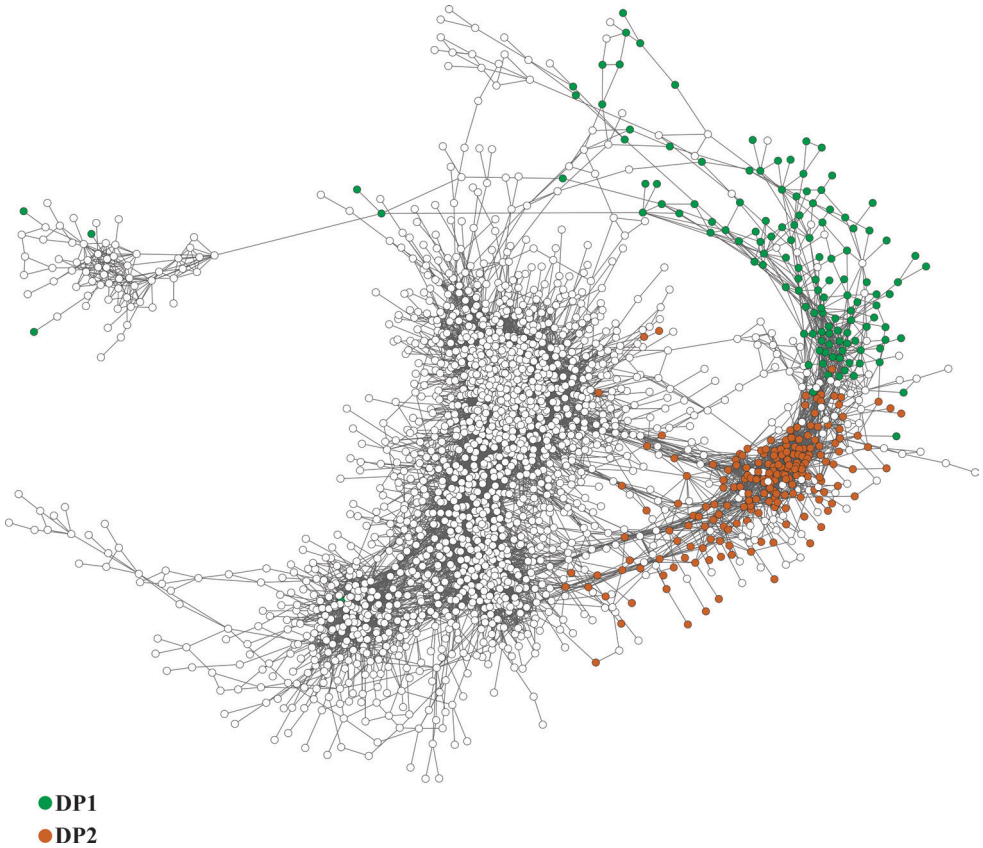


Figure S2: DT gene co-expression network. Nodes are coloured according to genes present in DP1 (green) and DP2 (orange).

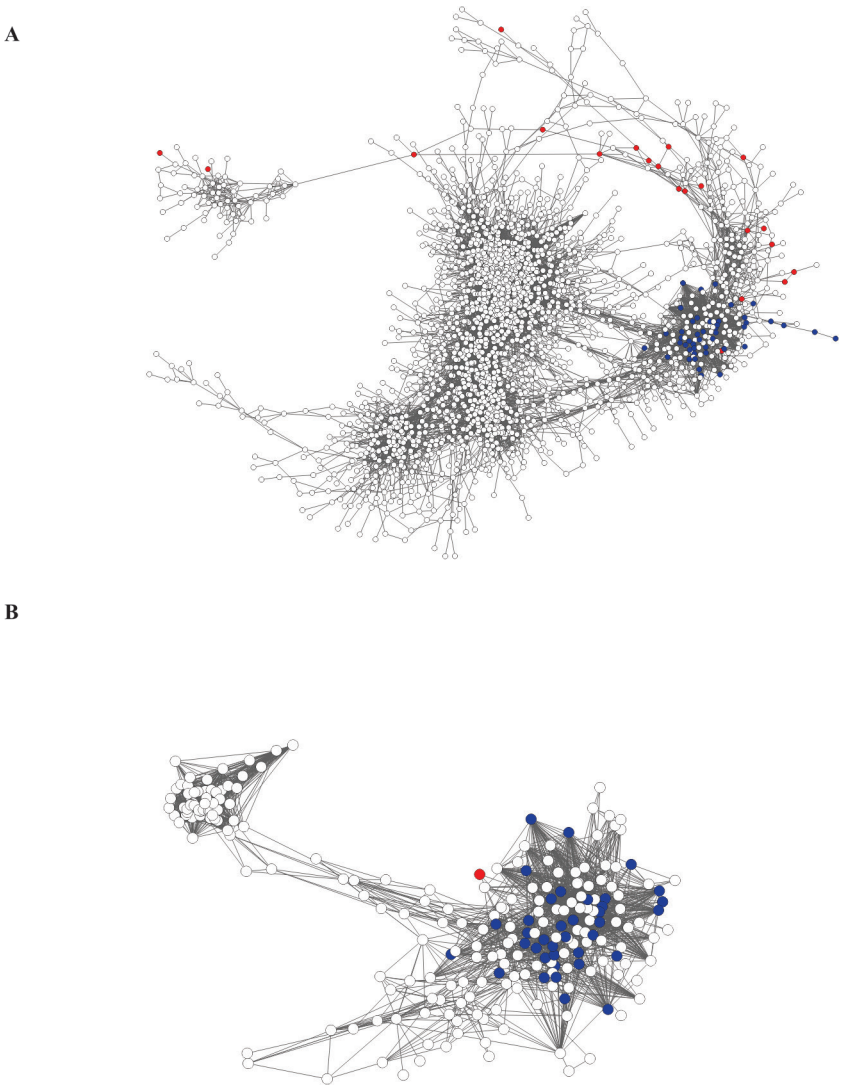


Figure S3: (**A**) DT gene co-expression network and (**B**) sub-network created from the ABI3 regulon presented in the gene co-expression network and genes connected with its members. Nodes are coloured in red for genes up-regulated in *abi3-9* (6HAD) and in blue for DT-specific data sets.

Supplementary material

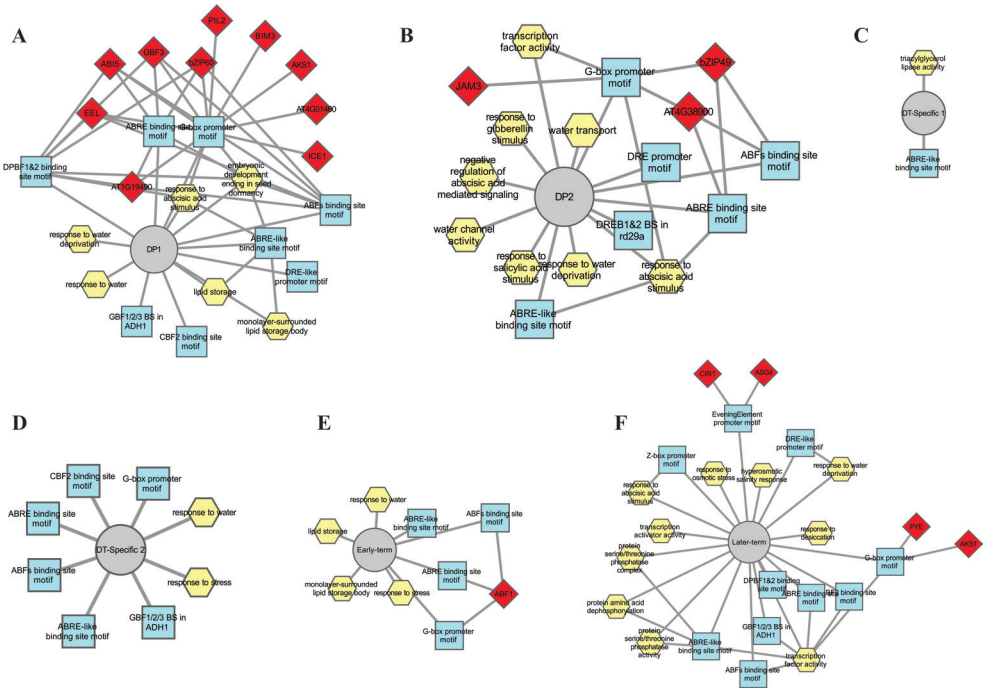


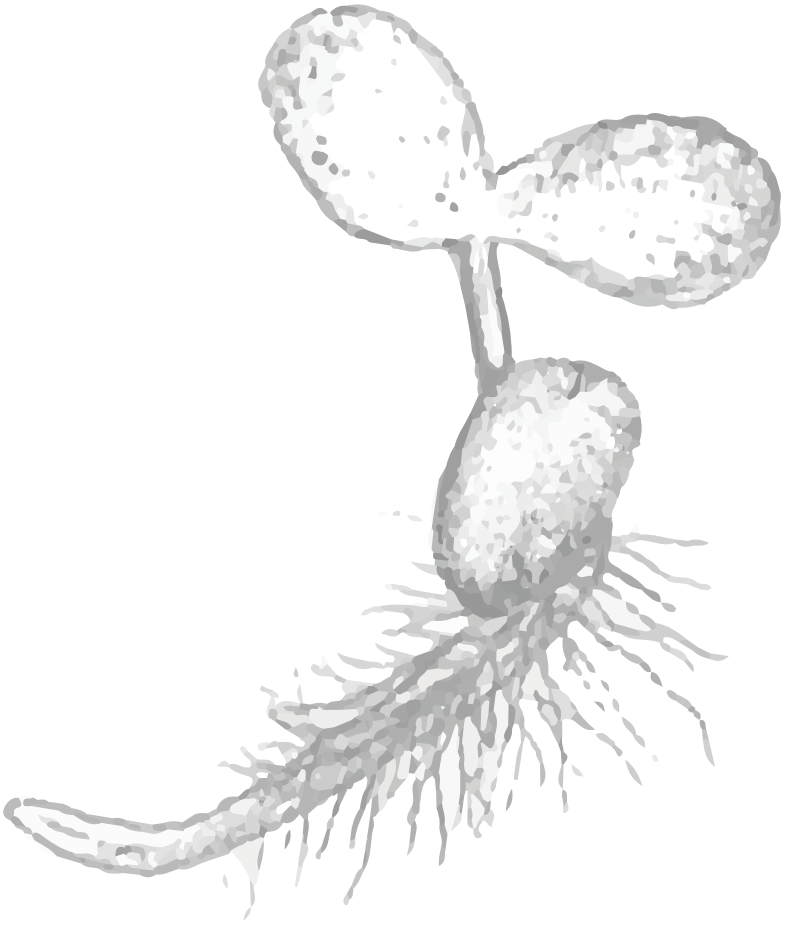
Figure S4: Transcriptional modules predicted for (**A**) DP1, (**B**) DP2, (**C**) DT-specific 1, (**D**) DT-specific 2, (**E**) early-response and (**F**) late-response gene data sets. DNA motifs (light-blue rectangles) and GO terms (light-yellow octagons) that are significantly overrepresented (P<0.001) within the sub-network gene sets (ellipse) together with co-expressed transcription factors (red diamonds).

Table S1: Expression data are and validation of the microarray data

Table S2: Dominant patterns, early-response, late-response and DT-specific data sets, and their associated enriched GO terms, DNA motifs, and metabolic pathways.

Chapter 6

General Discussion



Introduction

Plants undergo a series of phase transitions during their life cycle, transforming from the juvenile to the adult and from the vegetative to the reproductive phase. During these phase transitions plants pass through several changes that are often discernible by morphological characteristics. Arguably, the most critical phase change of the plant's life cycle is the transition from a seed to an established seedling that is dependent on developmental genetic programs that are triggered and modulated by environmental and endogenous stimuli. This transition is controlled by the endogenous balance and the signalling pathways of plant hormones, such as ethylene, auxin, abscisic acid (ABA), and gibberellins (GAs). Progression through stages across this transition is accompanied by reserve mobilization, chlorophyll synthesis and by the gradual loss of desiccation tolerance (DT). With the advent of postgenomics technologies, high-throughput analysis on a molecular basis, such as at RNA and metabolite profiling, have been widely implemented to reveal specific aspects of the seed germination process. A study encompassing seed germination of *Arabidopsis thaliana* has resulted in a genome-wide network (SeedNet) describing its transcriptional interactions (Bassel *et al.*, 2011). This network has shown to be accurate in defining regulators that behave as activators or inhibitors of germination.

In spite of the importance of the seed-to-seedling transition for crop establishment and yield, similar studies as those on germination are largely lacking for this process. In an attempt to make up for that, the research presented in my thesis focuses on four aspects of the seed-to-seedling phase transition: in **Chapter 2** I analysed the transcriptomic regulation of this phase transition; in **Chapter 3** I describe the metabolic shifts that occur during this transition; **Chapter 4** addresses the response of several distinct seed-to-seedling developmental stages to a novel mild air drying treatment (MADT); **Chapter 5** describes the transcriptomic control of the re-establishment of DT in germinated seeds.

The seed-to-seedling developmental phase transition in *Arabidopsis thaliana*

The phase transition from seed (heterotrophic) to seedling (photo-autotrophic) is characterized by organ development and maturation of the chloroplasts with the onset of photosynthesis, to support the new life cycle of the plant (Lourens Poorter, 2007). I divided the seed-to-seedling transition into several (morphological) stages: mature dry seed (DS); seeds imbibed for 6h with germination-related initiation of metabolic activity (6H); embryo swelling and testa rupture (TR); protrusion of the

radicle through the endosperm (RP), primarily through cell elongation, followed by further embryonic root extension and beginning of root hair formation (RH) and succeeded by the appearance of greening cotyledons (GC) and cotyledons that are fully opened (OC). Progression through these stages is accompanied by reserve mobilization and by the loss of desiccation tolerance (DT) (Bewley *et al.*, 2013).

The regulation of the seed-to-seedling phase transition occurs through modification of gene expression, including of transcription factors, and chromatin remodelling or (other) epigenetic modifications (Bouyer *et al.*, 2011; Jia *et al.*, 2014; Van Zanten *et al.*, 2014). For example, it was shown that the polycomb repressive complex 2 (PRC2), which catalyses the trimethylation of Lysine 27 on histone H3 (H3K27me3) (Schwartz and Pirrotta, 2008), is vital for the reprogramming of developmental fates mediating the switch from the embryonic state to the growing seedling (Bouyer *et al.*, 2011). Moreover, the seed-to-seedling phase transition is regulated by the activation of many genes controlling the stage changes and, simultaneously, the repression of genes controlling the preceding stage (Van Zanten *et al.*, 2014). The transition is highly regulated by hormones such as ABA, GA, auxin, cytokinin, ethylene, and brassinosteroids *via* their biosynthetic- and signalling pathways. For example, LEAFY COTYLEDON 1 (LEC1) was found to be involved in the regulation of hypocotyl elongation-related functions by targeting genes in auxin-, brassinosteroid- and light signalling networks in Arabidopsis (Junker *et al.*, 2012).

The control of seedling establishment is governed by many pathways and their crosstalk. Thus, to answer the question of what are the molecular mechanisms necessary for a vigorous seedling establishment, the molecular basis of the developmental stages encompassing the seed-to-seedling transition needs to be assessed at the highest resolution, using strategies that minimize experimental noise. The highly curated genome of *Arabidopsis thaliana* and the vast information on gene expression, metabolic- and protein profiles available in public databases such as 'The Arabidopsis Information Recourse' (TAIR) has enabled large-scale comparative studies. Therefore, I combined these advantages of the model plant Arabidopsis with the accurate morphological and physiological selection of developmental stages across the seed-to-seedling transition to try to unravel the mechanisms underlying the seed-to-seedling transition and the loss and re-acquisition of DT.

The molecular 'signatures' of the transition from a seed to a photo-autotrophic seedling, expressed in seven developmental stages across this phase transition were identified in **Chapters 2** and **3.** In **Chapter 2** I propose a signal transduction pathway that regulates seedling establishment. In this chapter I identified seven stage-specific- and ten dominant gene expression patterns (DPs), differentially expressed across

these seven stages. In the same chapter I built a gene co-expression network for the seed-to-seedling transition. With this network I showed that particularly DP7 harbours important regulators that control the seed-to-seedling transition. With a list of approximately 1.700 genes in DP7 I built a putative transcriptional module, and with this transcriptional module I could narrow down to a list of candidate regulators controlling the seed-to-seedling transition. One of the identified key regulators, the homeodomain leucine zipper I transcription factor *ATHB13,* is expressed during germination, but affects late seedling development. A recent study with *ATHB13* revealed a crucial role for this transcription factor in the juvenile-to-adult, and in the reproductive phase transitions of Arabidopsis (Ribone *et al.*, 2015).

Homeodomain leucine zipper I and plant phase transitions

The homeodomain leucine zipper (HD-Zip) group is unique to plants and is divided into four subgroups, named I-IV of which HD-Zip I is the largest subgroup with seventeen members (Schena and Davis, 1992; Ariel *et al.*, 2007). The first characterised member of this subgroup, ATHB1, acts as a mediator of leaf cell fate determination (Aoyama *et al.*, 1995), whereas another member, ATHB5, is involved in ABA-responses of developing seedlings (Johannesson *et al.*, 2003). It underlines the wide range of signalling pathways in which members of this subgroup act. In a few cases, members of HD-Zip I have been related to morphological and phase transition events (Saddic *et al.*, 2006; Son *et al.*, 2010). ATHB12 has been suggested to play a role in stem development through the control of GA-related genes in Arabidopsis (Son *et al.*, 2010). These authors demonstrated that a high level of transcripts of gibberellin 20-oxidase 1 (GA20ox1), a key enzyme in the synthesis of gibberellins, was detected in *athb12* mutant stems, while transgenic lines overexpressing *ATHB12* had a reduced level of GA20ox1. LMI1, also a member of the HD-Zip I subgroup, has been related to flowering phase transitions (Saddic *et al.*, 2006). This gene acts in regulating meristem identity by targeting *LEAFY* (*LFY*), affecting leaf shape and bract formation. Although this member of HD-Zip I has been described with a role in phase transitions, none of the members of the HD-Zip I family have been related to the seed-to-seedling transition up to now. *ATHB13*, which was found to be associated with seedling establishment (**Chapter 2),** appears to be a very interesting candidate. Evaluation of *athb13-1* mutant plants during seedling establishment showed that the primary root was significantly longer than wild type roots. This faster growth of the primary root may be associated with the rate of cell division. This corroborates a recent study in which the *athb13-1* mutant plants showed longer stems than the corresponding wild type (Ribone *et al.*, 2015). These authors suggested that this is

due to cell division rate, as the cell sizes of mutant and wild type stems did not differ. In the same study, a double mutant was generated: *athb13 athb23*. *ATHB13* and *ATHB23* have been described as paralogues (Arce *et al.*, 2011). Indeed, *ATHB23* appears to play a role, together with *ATHB13*, in stem elongation, whereas *ATHB13* alone plays a role in the reproductive phase transition (Ribone et al., 2015). The transcriptional module for the seed-to-seedling transition presented in **Chapter 2**, included not only *ATHB13* and *ATHB23*, but also *ATHB20*. Although the *ATHB13* and *ATHB23* genes act redundantly in stem elongation in Arabidopsis, I hypothesize that this may not be the case for root growth during seedling establishment, because in the seed-to-seedling co-expression network *ATHB23* is not highly correlated with *ATHB13*. However, *ATHB20* could be a good candidate to interact with *ATHB13* in seedling establishment as they are highly correlated. Furthermore, I showed that in the seed-to-seedling network, *ATHB13* and *ATHB20* are connected to five genes related with root development, seven with photosynthesis and one with auxin signalling. With this data follow-up research may be proposed combining the *athb13-1* mutant and auxin-responsive reporters such as DII-VENUS and/or DR5 (Brunoud *et al.*, 2012). A cross between *athb13-1* and plants with these reporters could reveal how the auxin response in the root development of *athb13-1* differs from wild type root development and highlight important aspects of a role of auxin in seedling establishment, as the hormone regulates cell division, cell elongation, cell differentiation and patterning in roots (Ljung *et al.*, 2005). The auxin signalling gene that correlated with *ATHB13* in the seed-to-seedling network is *INDOLE-3-ACETIC ACID 17* (*IAA17*), an auxin response repressor (Ding and Friml, 2010). It has been demonstrated that *IAA17* is essential for the maintenance of auxin gradients and cell differentiation in the root tip (Kim *et al.*, 2006; Tian *et al.*, 2014). Another follow-up to the data of **Chapter 2** could be an in-depth study of the *athb13-1 athb20-1* double mutant since these transcription factors correlated not only with root development, but also with genes in photosynthetic pathways. Interestingly, *ATHB13* was shown to be involved in the regulation of cotyledon and leaf development in response to carbon availability during seedling establishment (Hanson *et al.*, 2001). Building on this notion that the balance of carbon (C) and nitrogen (N) is very important for seedling establishment (Ribeiro *et al.*, 2014), I investigated the primary metabolite profiles during the same seven developmental stages across the seed-to-seedling transition in **Chapter 3**.

Metabolic shifts of the seed-to-seedling phase transition

The initial step of C acquisition by non-photosynthetic tissues occurs through carbohydrate partitioning, whereas amino acids are used for transport of N to dependent tissues (Yadav *et al.*, 2015). Metabolites found at high levels during the seed-to-seedling transition in **Chapter 3** are in agreement with the metabolites found to increase during germination (Fait *et al.* 2006). However, the metabolic analysis of **Chapter 3** identified methionine as an important amino acid for seedling establishment but this amino acid was not found by (Fait *et al.*, 2006) in their germination study. This suggests that such other metabolites may be fundamental for vigorous seedling establishment in Arabidopsis and much less for germination. The majority of the metabolites identified in **Chapter 3** were highly accumulated in the seedling stages GC and OC. These metabolites are likely involved in the tight coordination of the C/N balance. For example, CO_2 is assimilated through photosynthesis (at seedling stages) and, through conversion of sucrose and glucose in glycolysis and the tricarboxylic acid cycle (TCA) to α-ketoglutarate. This metabolite serves as a C skeleton for the synthesis of glutamate by incorporating photorespiratory $NH4^+$, which results in the production of glutamate and glutamine to donate $NH4^+$ for the synthesis of all other amino acids (Zheng, 2009). Taken together, it explains the high abundance of metabolites such as glucose, α-ketoglutarate and amino acids at seedling stages. In **Chapter 3** I used correlation analysis on the metabolites to investigate the interaction of metabolites that are more or less connected during early seedling establishment. The metabolite levels were shown to fall in two metabolite profiles. Profile 1 included seven amino acids (alanine, glutamine, glycine, lysine, phenylalanine, proline, and methionine) and pyruvate, whereas profile 2 was comprised of six amino acids (asparagine, GABA, isoleucine, serine, threonine, and valine), five carbohydrates (fructose, glucose, sorbose, xylose and trehalose), five organic acids (malate, α-ketoglutarate, succinate, ascorbate and glycerate), three carbohydrate derivatives (myo-inositol, fructose-6-phosphate and glucose-6-phosphate), adenosine, putrescine and phosphate. These profiles reflect the timing of metabolite changes, principally in profile 1, in which metabolites show a transient increase from the RP to the RH stage. Methionine a fundamental metabolite in its function of building blocks for protein synthesis (Ravanel *et al.*, 1998), is present in profile 1. Methionine is an important precursor of glucosinolates which are associated with pathways that control plant growth and development (Grubb and Abel, 2006). RNAi-inhibited expression of *CYTOCHROME P450 79F1*(*CYP79F1*) and *CYP79F2* in Arabidopsis resulted in a significant accumulation of methionine during plant development (Chen *et al.*, 2012). This is because CYP79F1 and CYP79F2 catalyze the biosynthesis of short-chain

and long-chain aliphatic glucosinolates using chain-elongated methionine substrates (Grubb and Abel, 2006). Another study with a *cyp79f1*mutant, and co-suppression of *CYP79F1* and *CYP79F2* yielded plants with strong phenotypes such as dwarf, bushy, and semi-sterile (Reintanz *et al.*, 2001). Targeting different pathways, a study using a proteomic approach showed the importance of methionine for seed germination and, consequently, for seedling development (Gallardo *et al.*, 2002). These authors demonstrated that methionine synthase increased strongly during the first 24 hours after imbibition (HAI), prior to radicle protrusion, but the level of this enzyme did not increase further at 48 HAI, which coincided with radicle protrusion. However, another enzyme (S-adenosylmethionine (AdoMet)) accumulated at the radicle protrusion stage. Consistent with an important role of methionine during the seed-to-seedling transition, a specific inhibitor of methionine (DL-propargylglycine) delayed seed germination and blocked seedling establishment; however, the phenotypes were recuperated after transfer to a medium supplemented with methionine (Gallardo *et al.*, 2002). In conclusion, methionine is an important amino acid for seedling development, likely as a key precursor of glucosinolate pathways, influencing this process. It is also interesting to speculate that the metabolites present in profile 2 represent a shift in the nutritional state preceding seedling establishment while the amino acids in profile 1 participate substantially in nitrogen transport (Orzechowski *et al.*, 1999; Yoo *et al.*, 2013; Pinto *et al.*, 2014; McAllister and Good, 2015).

Likely, metabolic regulation of the heterotrophic to autotrophic transition extends beyond primary metabolism. However, it is possible to identify primary metabolites involved in gene expression events, by examining their pattern in relation with the expression of specific genes (Gibon *et al.*, 2006). Therefore, in order to identify putative regulators of metabolite levels during the seed-to-seedling transition, I determined linear correlations between the gene expression data of **Chapter 2** and metabolite contents of **Chapter 3**. These correlations may indicate potential genes controlling primary metabolite levels during the seed-to-seedling transition. For example, the expression of genes encoding three isoforms of methionine synthetase, *METHIONINE SYNTHASE 1* (*MS1*), *METHIONINE SYNTHASE 2* (*MS2*) and *HOMOCYSTEINE S-METHYLTRANSFERASE* (*HMT-1*) was highly correlated with methionine levels. Although the expression of many genes correlated strongly with metabolite contents, it does not reveal how genes affect metabolite levels or *vice versa*. To determine whether the observed correlation between gene and metabolite plays a role during seedling establishment, further experiments are required, including, for example by gene expression measurements in a controlled metabolite feeding experiment, or by metabolite profiling of the relevant mutants. Although my results provide a view of the correlations of transcript

abundance and metabolite contents, I have only considered correlations between single entities (metabolite *vs*. transcript) due to inherent difficulties in the complexity of data-sets integration in larger networks (Reed *et al.*, 2014). Looking at **Chapter 2** and **3** together, it is interesting to speculate that a shift in the state of metabolites important for seedling establishment such as methionine (Gallardo *et al.*, 2002), precedes the major regulatory pattern in transcriptional state (DP7) from germination to seedling establishment. This may indicate that an appropriated metabolic state prior to seedling establishment is necessary for the onset of the transcriptional program (DP7) for a vigorous seedling establishment. This new hypothesis on the regulation of the seed-to-seedling transition in Arabidopsis opens new perspectives for understanding the complex regulatory mechanism underlying this transition.

After addressing the molecular basis of the seed-to-seedling phase transition in **Chapters 2** and **3**, I focused on an important physiological feature occurring during this transition, namely desiccation tolerance (DT) in **Chapters 4** and **5**. DT is an extremely important trait that plants, such as Arabidopsis, acquire naturally during seed development (Meurs *et al.*, 1992; Gutierrez *et al.*, 2007) and which is lost during the seed-to-seedling transition (Maia *et al.*, 2011).

The seed-to-seedling transition and desiccation tolerance

DT is acquired during the phase of seed development when embryo growth ceases and storage products accumulate (Gutierrez *et al.*, 2007). The seed-to-seedling phase transition is a mirror image of seed development when DT is lost and storage products mobilized. It is interesting to note that a gene set associated with the re-establishment of DT in **Chapter 5** (DP1) included genes with a role in lipid storage, suggesting the re-activation of the seed maturation program. Seed lipid bodies are organelles for the long term storage of neutral lipids accumulated during seed development and maturation, and eventually mobilised after the completion of germination (Bewley *et al.*, 2013). The accumulation of lipids during seed development (Siloto *et al.*, 2006) and during dehydration (**Chapter 5**), appears to be an alternative carbon source for the synthesis of sucrose or trehalose, providing osmoprotection to the cells because DP1 was also enriched for metabolic processes such as trehalose and sucrose biosynthesis. Several studies on anhydrobiosis have emphasized the importance of the non-reducing disaccharides trehalose and sucrose, one or other of these sugars being present at high concentrations during desiccation of anhydrobiotic nematodes, brine shrimp cysts, bakers' yeast, resurrection plants and plant seeds (Tunnacliffe and Lapinski, 2003).

Embryo growth and seed maturation are controlled by master regulators such

as LEAFY COTYLEDON 1 (LEC1), LEC1-LIKE (L1L), transcription factors with B3 domain LEAFY COTYLEDON 2 (LEC2), FUSCA 3 (FUS3) and ABSCISIC ACID INSENSITIVE 3 (ABI3) (Gutierrez *et al.*, 2007; Santos-Mendoza *et al.*, 2008; Verdier *et al.*, 2013). These master regulators which share a common B3 domain are directly involved in the production of seed reserves during development, of which FUS3 is the transcription factor controlling lipid content (Roscoe *et al.*, 2015). These authors showed that lipid storage is high compromised during seed development in the triple mutant *fus3lec2abi3*. However, overexpression of one of these master regulators could complement the seed maturation trait, such as lipid storage. Moreover, a gene co-expression network of seed development in *Medicago truncatula* revealed that the master regulator *ABI3* was highly correlated with *ABI4* and *ABI5* in the DT gene cluster (Verdier *et al.*, 2013). These authors suggested that these *ABI* genes play an important role in ABA-signalling to ensure survival of the seed in a dry state. Germinated seeds can also survive the dry state after proper pre-conditioning, using a mild osmotic stress before desiccation. This has been shown for a number of species including Arabidopsis (Dekkers *et al.*, 2015). Although re-establishment of DT can be induced by an osmotic treatment with polyethylene glycol (PEG) (Maia *et al.*, 2011), there may be undesirable adverse effects during the treatment, including toxicity, blocking of water movement and a reduced transport of oxygen (Lagerwerff *et al.*, 1961; Lawlor, 1970). For these reasons I developed, in **Chapter 4,** the Mild Air Drying Treatment (MADT), an easier, faster, robust and less labour intensive protocol to re-induce and evaluate DT in Arabidopsis seeds. Germinated Arabidopsis seeds are sensitive to a fast drying treatment. It has been shown by Maia *et al.* (2011) that seeds transferred to a dry filter paper and dried for 72 hours under a relative humidity of 32%, were unable to continue with their developmental program after rehydration. It is likely that fast drying prevents the activation of protective mechanisms required to cope with desiccation in germinated seeds. However, in **Chapter 4** I showed that Arabidopsis seeds at the radicle protrusion stage are able to become DT if a slow drying method is applied. Interestingly, seeds submitted to MADT attained within 2 hours a water content comparable to the water content obtained in Arabidopsis seeds using a 3-day -2.5 MPa PEG treatment (Maia *et al.*, 2011). This demonstrates that to re-establish DT at the radicle protrusion stage, 2 hours in MADT is enough, whereas in PEG this takes 72 hours. By using the *aba2-1* and *aba3-1* mutants, I showed that the capacity of seeds of these mutants at the radicle protrusion stage to re-induce DT was also compromised in the MADT protocol, similarly to the osmotic treatment, as previously illustrated (Maia *et al.*, 2014). Therefore, in order to confirm whether ABA biosynthesis is restricted to the stages in which DT can be re-induced, I measured the ABA contents of germinated

seeds with the competence to induce DT (at radicle protrusion) and in seeds at the root hair stage, which are clearly incompetent. Although the root hair stage seeds were not able to re-establish DT, ABA content after 72 hours of dehydration was comparable with seeds at radicle protrusion. Thus, I also confirmed that an enhanced ABA accumulation is part of the DT response induced by MADT.

Taken into consideration that ABA is a master regulator of DT acquisition during seed development (Gutierrez *et al.*, 2007; Verdier *et al.*, 2013) and during its re-establishment (**Chapter 4**; Maia et al., 2014) the role of *ABI* genes was studied during the re-establishment of DT in germinated Arabidopsis seeds by -2.5 MPa PEG (Maia *et al.*, 2014). Interestingly, the ABA-insensitive *abi3-9*, *abi3-10*, *abi4-3* and *abi5-7* mutants produced desiccation tolerant seeds whereas the same mutants were compromised in their capacity to re-establish DT at radicle protrusion (Maia *et al.*, 2014). In spite of the importance of these genes in the re-establishment of DT, their genetic interactions have not been extensively studied yet. Using a yeast one-hybrid approach a molecular interaction between *ABI4* and *ABI5* was shown to be synergistic (Reeves *et al.*, 2011). In **Chapter 4** I showed that by comparing the single and double mutants, *ABI4* interacts also genetically in a synergistic manner with *ABI5* and *ABI3* during the re-establishment of DT. However, a similar synergistic effect could not be detected between *ABI3* and *ABI5*. With this data I hypothesized that either *ABI3* acts epistatically over *ABI5* or they interact genetically in an additive manner. A study of a genetic relationship between *ABI3* and *ABI5* showed that *ABI5* acts downstream of *ABI3* to arrest early seedling growth (Lopez-Molina *et al.*, 2002). Although the relationship between *ABI3* and *ABI5* is clear, it could be interesting to investigate whether overexpression lines of *ABI5* and *ABI3* would extend the developmental window in which DT can still be re-established. The observation that *ABI5* acts downstream from *ABI3* is confirmed in **Chapter 5**. In this study, employing a more in-depth analysis, I used the *abi3-9* mutant at the radicle protrusion stage (which has an 85% reduced ability to re-induce DT) and compared it with Col-0 at the same stage (with a 100% ability to re-induce DT). With this approach I found that *ABI5* is also one of the main regulators of DT re-establishment. These observations suggest that *ABI3* and *ABI5* act in a similar signalling pathway to regulate the ABA response during re-establishment of DT. The role of the *ABI* genes during the re-establishment of DT extended also to *ABI4* and *ABI5*. I showed that *ABI4* and *ABI5* interact synergistically during the re-establishment of DT (**Chapter 4**). However, it is interesting to speculate whether an overexpression line of *ABI5* in *abi4-3* would rescue the *abi4-3* phenotype during the re-establishment of DT. It is clear that ABA plays a role during re-establishment of DT (**Chapter 4**). Next, I checked whether there is cross-talk between ABA-dependent and ABA-independent genes (**Chapter 5**).

Control of re-establishment of DT through ABA-dependent and ABA-independent genes

In order to acquire DT during seed development, plants employ a signal transduction network in which diverse transcription factors control the expression of many genes (Gutierrez *et al.*, 2007; Verdier *et al.*, 2013; Dekkers *et al.*, 2015). Moreover, transcripts of the re-establishment of DT in Arabidopsis have been associated with transcripts of the seed development program (Maia *et al.*, 2011). A recent study of the re-induction of DT in germinated Arabidopsis seeds, mediated by ABA, showed that both drought and desiccation stress responses are largely mediated through ABA signalling pathways (Costa *et al.*, 2015). However, the role of ABA-independent genes in DT remains elusive. The best understood regulatory network of DT, however, is in Medicago seed development, and it is known to be ABA-dependent rather than ABA-independent (Verdier *et al.*, 2013). However, cross-talk between ABA-dependent and ABA-independent pathways is likely because stress-responsive genes, including *DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN 2A* (*DREB2A*), are regulated by both dependent- and independent pathways (Yoshida *et al.*, 2014). Considering that in Arabidopsis Col-0, DT at the root hair stage is only re-established in about 20% of the seeds and at radicle protrusion in about 100% (**Chapter 4** and **Chapter 5**), I used these seeds across dehydration intervals in a further attempt to unravel the mechanism of re-establishment of DT. However, when employing only Col-0 seeds at radicle protrusion and root hair stages, only genes associated with the germination program rather than the program of re-establishment of DT are showing up. To exclude that, I used the *abi3-9* mutant at the radicle protrusion stage, in which DT is re-established only in about 15% of the seeds (Maia et. al. 2014; **Chapter 4** and **Chapter 5**). Thus, in **Chapter 5,** I found a putative transcriptional module of re-establishment of DT in germinated Arabidopsis seeds, clearly suggesting a cross-talk between ABA-dependent and ABA-independent signalling pathways. In this study, two putative ABA-independent transcription factors *INDUCER OF CBF EXPRESSION 1* (*ICE1*) and *CRY2-INTERACTING BHLH 4* (*CIB4*), and two ABA-dependent ones, *ABI5* and At1g949 were found. A recent study also showed that *ICE1* is important in embryo development (Denay *et al.*, 2014), possibly through regulation of one of these *ABI* genes. Moreover, *ICE1* has been suggested to act as a negative regulator of the ABA-dependent pathway (Liang and Yang, 2015). These authors demonstrated that expression of *ABI3* and *ABI4* is high in the *ice1-2* mutant upon ABA treatment. This treatment with ABA also arrested seedling growth of *ice1-2*, and the high expression of *ABI* genes may have contributed to this arrest. Arrest of seedling growth has also been associated with the interaction between *ABI3* and *ABI5* (Lopez-Molina *et al.* (2002). They showed that *ABI3* acts upstream from

ABI5 which corroborates my result of **Chapter 5**, in which I identified the *ABI5* DT sub-network by using the *abi3* mutant. Considering that *ABI5* is downstream of *ABI3* (Lopez-Molina *et al.*, 2002) and that *ABI5* plays a role during acquisition of DT in *Medicago truncatula* seed development (Verdier *et al.*, 2013), I expected to find *ABI5* in the DT sub-network in **Chapter 5**. However, the *abi5* mutant is able to re-establish DT, comparable to the wild type (**Chapter 5**), which suggests that genes found in the transcriptional module may act redundantly to *ABI5* in the re-establishment of DT.

Further considerations

In my thesis I made considerable progress made in understanding the seed-to-seedling transition and the re-establishment of DT during this phase transition. In **Chapters 2** and **3** the transcript and metabolic profiles of the seed-to-seedling transition highlight new perspectives to understand this complex phenomenon. In **Chapter 2** the putative novel regulators of the transition provide a new dimension to the transcriptional regulation of the seed-to-seedling transition. The results of **Chapter 2** can be extended to crop species in which controlled expression of these regulators of seedling development can help to improve this phase transition in crops. Crop transformations may be considered for species with a great economic importance such as soybean and maize. Besides the identification of an important gene expression pattern for seedling establishment, I also correlated transcripts with metabolites in **Chapter 3**. With this I narrowed down the list of putative genes that control energy metabolism during the phase transition. It will be interesting to see whether these metabolic profiles extend to the seed-to-seedling transition in crops, and, if this is the case, the genes that were tightly correlated with the metabolites in Arabidopsis could be used as starting point to study this phase transition in crops (particularly in crops for which extensive microarray or RNAseq data is not available). The transition from seed to plant is a hallmark of the plant's life cycle. In general, it is advantageous for a species to keep the period of seedling establishment as short as possible since young seedlings are highly sensitive to biotic and abiotic stresses (Pandey *et al.*, 2012; Zhang *et al.*, 2014). Within the context of the molecular basis of the seed-to-seedling transition, as described in **Chapters 2** and **3,** the stress-inducible genes involved in the transition where assessed in **Chapter 5** by using a novel approach to re-establish DT under mild air drying treatment (MADT) (**Chapter 4**). With this approach I have generated a reliable system to study DT. In a more detailed study (**Chapter 5**) I propose that re-establishment of DT is mediated through a cross-talk between ABA-dependent and ABA-independent genes. Further studies in Arabidopsis are

needed to confirm this hypothesis. Overexpression of the putative key regulators of the DT mechanism may be feasible to maintain the seed's stress tolerance well into the seedling stage.

References

Aoyama T, Dong CH, Wu Y, Carabelli M, Sessa G, Ruberti I, Morelli G, Chua NH.1995 Ectopic expression of the Arabidopsis transcriptional activator Athb-1 alters leaf cell fate in tobacco. *Plant Cell* **7:** 1773-1785

Arce AL, Raineri J, Capella M, Cabello JV, Chan RL.2011 Uncharacterized conserved motifs outside the HD-Zip domain in HD-Zip subfamily I transcription factors; a potential source of functional diversity. *BMC Plant Biol* **11:** 42

Ariel FD, Manavella PA, Dezar CA, Chan RL.2007 The true story of the HD-Zip family. *Trends Plant Sci* **12:** 419-426

Bassel GW, Lan H, Glaab E, Gibbs DJ, Gerjets T, Krasnogor N, Bonner AJ, Holdsworth MJ, Provart NJ.2011 Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. *Proceedings of the National Academy of Sciences* **108:** 9709-9714

Bewley JD, Bradford K, Hilhorst HM, Nonogaki H (2013) Germination. *In* Seeds. Springer New York, pp 133-181

Bouyer D, Roudier F, Heese M, Andersen ED, Gey D, Nowack MK, Goodrich J, Renou JP, Grini PE, Colot V, Schnittger A.2011 Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. *PLoS Genet* **7:** e1002014

Brunoud G, Wells DM, Oliva M, Larrieu A, Mirabet V, Burrow AH, Beeckman T, Kepinski S, Traas J, Bennett MJ, Vernoux T.2012 A novel sensor to map auxin response and distribution at high spatio-temporal resolution. *Nature* **482:** 103-106

Chen Y-z, Pang Q-Y, He Y, Zhu N, Branstrom I, Yan X-F, Chen S.2012 Proteomics and Metabolomics of Arabidopsis Responses to Perturbation of Glucosinolate Biosynthesis. *Molecular Plant* **5:** 1138-1150

Costa M, Righetti K, Nijveen H, Yazdanpanah F, Ligterink W, Buitink J, Hilhorst HM.2015 A gene co-expression network predicts functional genes controlling the re-establishment of desiccation tolerance in germinated Arabidopsis thaliana seeds. *Planta***:** 1-15

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

Denay G, Creff A, Moussu S, Wagnon P, Thévenin J, Gérentes M-F, Chambrier P, Dubreucq B, Ingram G.2014 Endosperm breakdown in Arabidopsis requires heterodimers of the basic helix-loop-helix proteins ZHOUPI and INDUCER OF CBP EXPRESSION 1. *Development* **141:** 1222-1227

Ding Z, Friml J.2010 Auxin regulates distal stem cell differentiation in Arabidopsis roots.

Proc Natl Acad Sci U S A **107:** 12046-12051

Fait A, Angelovici R, Less H, Ohad I, Urbanczyk-Wochniak E, Fernie AR, Galili G.2006 Arabidopsis Seed Development and Germination Is Associated with Temporally Distinct Metabolic Switches. *Plant Physiology* **142:** 839-854

Gallardo K, Job C, Groot SP, Puype M, Demol H, Vandekerckhove J, Job D.2002 Importance of methionine biosynthesis for Arabidopsis seed germination and seedling growth. *Physiol Plant* **116:** 238-247

Gibon Y, Usadel B, Blaesing OE, Kamlage B, Hoehne M, Trethewey R, Stitt M.2006 Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in Arabidopsis rosettes. *Genome Biology* **7:** R76-R76

Grubb CD, Abel S.2006 Glucosinolate metabolism and its control. *Trends Plant Sci* **11:** 89-100

Gutierrez L, Van Wuytswinkel O, Castelain M, Bellini C.2007 Combined networks regulating seed maturation. *Trends in Plant Science* **12:** 294-300

Hanson J, Johannesson H, Engstrom P.2001 Sugar-dependent alterations in cotyledon and leaf development in transgenic plants expressing the HDZhdip gene ATHB13. *Plant Mol Biol* **45:** 247-262

Jia H, Suzuki M, McCarty DR.2014 Regulation of the seed to seedling developmental phase transition by the LAFL and VAL transcription factor networks. *Wiley Interdisciplinary Reviews: Developmental Biology* **3:** 135-145

Johannesson H, Wang Y, Hanson J, Engström P.2003 The Arabidopsis thaliana homeobox gene ATHB5 is a potential regulator of abscisic acid responsiveness in developing seedlings. **51:** 719-729

Junker A, Monke G, Rutten T, Keilwagen J, Seifert M, Thi TM, Renou JP, Balzergue S, Viehover P, Hahnel U, Ludwig-Muller J, Altschmied L, Conrad U, Weisshaar B, Baumlein H.2012 Elongation-related functions of LEAFY COTYLEDON1 during the development of Arabidopsis thaliana. *Plant J* **71:** 427-442

Kim H, Park PJ, Hwang HJ, Lee SY, Oh MH, Kim SG.2006 Brassinosteroid signals control expression of the AXR3/IAA17 gene in the cross-talk point with auxin in root development. *Bioscience, Biotechnology and Biochemistry* **70:** 768-773

Lagerwerff JV, Ogata G, Eagle HE.1961 Control of osmotic pressure of culture solutions with polyethylene glycol. *Science* **133:** 1486-1487

Lawlor DW.1970 ABSORPTION OF POLYETHYLENE GLYCOLS BY PLANTS AND THEIR EFFECTS ON PLANT GROWTH. *New Phytologist* **69:** 501-513

Liang CH, Yang CC.2015 Identification of ICE1 as a negative regulator of ABA-dependent pathways in seeds and seedlings of Arabidopsis. *Plant Mol Biol* **88:** 459-470

Ljung K, Hull AK, Celenza J, Yamada M, Estelle M, Normanly J, Sandberg G.2005 Sites and regulation of auxin biosynthesis in arabidopsis roots. *Plant Cell* **17:** 1090-1104

Lopez-Molina L, Mongrand S, McLachlin DT, Chait BT, Chua NH.2002 ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination. *Plant J* **32:** 317-328

Lourens Poorter.**2007** Are Species Adapted to Their Regeneration Niche, Adult Niche, or Both? *The American Naturalist* **169:** 433-442

Maia J, Dekkers BJW, Dolle MJ, Ligterink W, Hilhorst HWM.**2014** Abscisic acid (ABA) sensitivity regulates desiccation tolerance in germinated Arabidopsis seeds. *New Phytologist***:** n/a-n/a

Maia J, Dekkers BJW, Provart NJ, Ligterink W, Hilhorst HWM.**2011** The Re-Establishment of Desiccation Tolerance in Germinated *Arabidopsis thaliana* Seeds and Its Associated Transcriptome. *PLos ONE* **6**

McAllister CH, Good AG.**2015** Alanine aminotransferase variants conferring diverse NUE phenotypes in Arabidopsis thaliana. *PLos ONE* **10**

Meurs C, Basra AS, Karssen CM, Van Loon LC.**1992** Role of abscisic acid in the induction of desiccation tolerance in developing seeds of Arabidopsis thaliana. *Plant Physiology* **98:** 1484-1493

Orzechowski S, Socha-Hanc J, Paszkowski A.**1999** Alanine aminotransferase and glycine aminotransferase from maize (Zea mays L.) leaves. *Acta Biochim Pol* **46:** 447-457

Pandey PK, Yadav SK, Singh A, Sarma BK, Mishra A, Singh HB.**2012** Cross-Species Alleviation of Biotic and Abiotic Stresses by the Endophyte Pseudomonas aeruginosa PW09. *Journal of Phytopathology* **160:** 532-539

Pinto JT, Krasnikov BF, Alcutt S, Jones ME, Dorai T, Villar MT, Artigues A, Li J, Cooper AJL.**2014** Kynurenine aminotransferase III and glutamine transaminase L are identical enzymes that have cysteine S-conjugate β-lyase activity and can transaminate L-selenomethionine. *Journal of Biological Chemistry* **289:** 30950-30961

Ravanel S, Gakiere B, Job D, Douce R.**1998** The specific features of methionine biosynthesis and metabolism in plants. *Proc Natl Acad Sci U S A* **95:** 7805-7812

Reed LK, Lee K, Zhang Z, Rashid L, Poe A, Hsieh B, Deighton N, Glassbrook N, Bodmer R, Gibson G.**2014** Systems Genomics of Metabolic Phenotypes in Wild-Type Drosophila melanogaster. *Genetics*

Reeves W, Lynch T, Mobin R, Finkelstein R.**2011** Direct targets of the transcription factors ABA-Insensitive(ABI)4 and ABI5 reveal synergistic action by ABI4 and several bZIP ABA response factors. *Plant Molecular Biology* **75:** 347-363

Reintanz B, Lehnen M, Reichelt M, Gershenzon J, Kowalczyk M, Sandberg G, Godde M, Uhl R, Palme K.**2001** Bus, a bushy Arabidopsis CYP79F1 knockout mutant with abolished synthesis of short-chain aliphatic glucosinolates. *Plant Cell* **13:** 351-367

Ribeiro PR, Fernandez LG, de Castro RD, Ligterink W, Hilhorst HWM.**2014** Physiological and biochemical responses of Ricinus communis seedlings to different temperatures: A metabolomics approach. **14**

Ribone PA, Capella M, Chan RL.**2015** Functional characterization of the homeodomain leucine zipper I transcription factor AtHB13 reveals a crucial role in Arabidopsis development. *Journal of Experimental Botany*

Roscoe TT, Guilleminot J, Bessoule JJ, Berger F, Devic M.**2015** Complementation of Seed Maturation Phenotypes by Ectopic Expression of ABSCISIC ACID INSENSITIVE3,

FUSCA3 and LEAFY COTYLEDON2 in Arabidopsis. *Plant Cell Physiol* **56:** 1215-1228

Saddic LA, Huvermann B, Bezhani S, Su Y, Winter CM, Kwon CS, Collum RP, Wagner D.2006 The LEAFY target LMI1 is a meristem identity regulator and acts together with LEAFY to regulate expression of CAULIFLOWER. *Development* **133:** 1673-1682

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

Schena M, Davis RW.1992 HD-Zip proteins: members of an Arabidopsis homeodomain protein superfamily. *Proc Natl Acad Sci U S A* **89:** 3894-3898

Schwartz YB, Pirrotta V.2008 Polycomb complexes and epigenetic states. *Curr Opin Cell Biol* **20:** 266-273

Siloto RMP, Findlay K, Lopez-Villalobos A, Yeung EC, Nykiforuk CL, Moloney MM.2006 The accumulation of oleosins determines the size of seed oilbodies in Arabidopsis. *Plant Cell* **18:** 1961-1974

Son O, Hur YS, Kim YK, Lee HJ, Kim S, Kim MR, Nam KH, Lee MS, Kim BY, Park J, Park J, Lee SC, Hanada A, Yamaguchi S, Lee IJ, Kim SK, Yun DJ, Soderman E, Cheon CI.2010 ATHB12, an ABA-inducible homeodomain-leucine zipper (HD-Zip) protein of Arabidopsis, negatively regulates the growth of the inflorescence stem by decreasing the expression of a gibberellin 20-oxidase gene. *Plant Cell Physiol* **51:** 1537-1547

Tian H, Wabnik K, Niu T, Li H, Yu Q, Pollmann S, Vanneste S, Govaerts W, Rolčík J, Geisler M, Friml J, Ding Z.2014 WOX5-IAA17 feedback circuit-mediated cellular auxin response is crucial for the patterning of root stem cell niches in arabidopsis. *Molecular Plant* **7:** 277-289

Tunnacliffe A, Lapinski J.2003 Resurrecting Van Leeuwenhoek's rotifers: A reappraisal of the role of disaccharides in anhydrobiosis. *Philosophical Transactions of the Royal Society B: Biological Sciences* **358:** 1755-1771

Van Zanten M, Zöll C, Wang Z, Philipp C, Carles A, Li Y, Kornet NG, Liu Y, Soppe WJJ.2014 Histone deacetylase 9 represses seedling traits in Arabidopsis thaliana dry seeds. *Plant Journal* **80:** 475-488

Verdier J, Lalanne D, Pelletier S, Torres-Jerez I, Righetti K, Bandyopadhyay K, Leprince O, Chatelain E, Vu BL, Gouzy J, Gamas P, Udvardi MK, 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

Yadav UP, Ayre BG, Bush DR.2015 Transgenic approaches to altering carbon and nitrogen partitioning in whole plants: assessing the potential to improve crop yields and nutritional quality. *Front Plant Sci* **6:** 275

Yoo H, Widhalm JR, Qian Y, Maeda H, Cooper BR, Jannasch AS, Gonda I, Lewinsohn E, Rhodes D, Dudareva N.2013 An alternative pathway contributes to phenylalanine

biosynthesis in plants via a cytosolic tyrosine:phenylpyruvate aminotransferase. *Nature communications* **4**

Yoshida T, Mogami J, Yamaguchi-Shinozaki K.**2014** ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. *Curr Opin Plant Biol* **21:** 133-139

Zhang Z, Chen Y, Zhao D, Li R, Wang H, Zhang J, Wei J.**2014** X1-homologous genes family as central components in biotic and abiotic stresses response in maize (Zea mays L.). *Functional and Integrative Genomics* **14:** 101-110

Zheng ZL.**2009** Carbon and nitrogen nutrient balance signaling in plants. *Plant Signaling and Behavior* **4:** 584-591

Summary

One of the most important developmental processes in the life-cycle of higher plants is the transition from seed to plant, from a generative to a vegetative developmental program. The major hallmark or end-point of the transition from seed to plant is the onset of photosynthesis and the concomitant shift from heterotrophy to autotrophy. It is advantageous for a species to keep the period of seedling establishment as short as possible since young seedlings are highly sensitive to biotic and abiotic stresses, since they have lost the extreme stress tolerance that seeds have, to for example desiccation, upon germination. If the regulatory principles of the seed-to-seedling transition are better understood it may become feasible to maintain the seed's stress tolerance well into the seedling stage.

Despite the profound impact of seedling performance on crop establishment and yield, the seed-to-seedling transition has hardly been studied at the molecular level. This thesis therefore aimed at deciphering and understanding the molecular processes that govern this transition in *Arabidopsis thaliana*. A high-resolution study of the molecular events that occur during these successive transitional stages may provide clues as to the regulatory principles that drive this transition. It may also yield information about the factors that are critical for the maintenance and loss of desiccation tolerance and other stress responses.

In **Chapter 1** I describe important processes involved in germination and seedling development as well as in dehydration tolerance such as abscisic acid signalling and their regulation and I discuss in what way the seed-to-seedling transition may have links to a trait such as desiccation tolerance. I present an overview of the current knowledge of the seed-to-seedling phase transition and the existence of a point-of-no-return that can be manipulated by osmotic treatment and by abscisic acid which results in the re-establishment of desiccation tolerance, suggesting the re-activation of the seed maturation program.

In **Chapter 2** the regulation of gene expression in the seed-to-seedling transition is studied by a detailed transcriptional analysis across seven developmental stages. It describes the inference of a gene co-expression network and several transcriptional modules. We show that such an approach highlights important molecular processes during seedling development, which would not likely be derived from comparative transcript profiling. Moreover, we show that a putative key regulator in one of the transcriptional modules affects late seedling establishment.

In **Chapter 3** we show how the seed-to-seedling phase transition is reflected in the primary metabolite profiles and how the latter correlate with gene expression. Metabolite-metabolite correlation analysis suggests two different profiles occur, which point at the metabolic preparation of seed germination and of vigorous seedling establishment, respectively. Furthermore, a linear correlation between metabolite

contents and transcript abundance provides a global view of the transcriptional and metabolic changes during the seed-to-seedling transition. It creates new perspectives for the identification of the regulatory complexes underlying the seed-to-seedling transition.

Chapter 4 describes a novel method to study the re-establishment of desiccation tolerance during the seed-to-seedling transition without adverse effects such as those caused by an osmotic treatment with polyethylene glycol. By using this method, named 'Mild Air Drying Treatment' (MADT), we show that desiccation tolerant seeds (germinated seeds at radicle protrusion) and desiccation sensitive seeds (germinated seeds at root hair stage) accumulated the same amount of ABA. We conclude that the re-establishment of desiccation tolerance is not linked to a reduced ability to accumulate ABA in the desiccation sensitive seeds. I also present a genetic interaction study of *ABSCISIC ACID INSENSITIVE* (*ABI*) genes in their response to the re-establishment of desiccation tolerance using the MADT. The interaction between *ABI3* and *ABI4,* and between *ABI4* and *ABI5* act synergistically in the re-establishment of DT.

In a more in depth study in **Chapter 5** I carried out transcriptome analysis to infer possible mechanisms of the re-establishment of desiccation tolerance. Using the MADT protocol, possible mechanisms underlying the re-establishment of desiccation tolerance were inferred by employing a time-series comparison of germinated desiccation tolerant and -sensitive seeds. Early-response genes of the re-establishment of desiccation tolerance may play a role in events that promote the initial protection to dehydration stress, whereas the late-response genes may play a role in events that help seeds to respond to the changes in water dynamics. Moreover, using a gene co-expression network and transcriptional module I concluded that crosstalk between ABA-dependent and ABA-independent transcription factors regulates the re-establishment of desiccation tolerance.

In **Chapter 6** I discuss how the results presented in this thesis contribute to our knowledge of the molecular basis of the seed-to-seedling transition and the re-establishment of desiccation tolerance during its phase changes. Finally, new possibilities for further research are discussed, as well as the further use of the data sets to delineate the mechanisms underlying the seed-to-seedling transition and desiccation tolerance. Possible applications of the results for crop improvement are addressed

Samenvatting

Eén van de belangrijkste ontwikkelingsprocessen in de levenscyclus van de hogere plant is de transitie van zaad naar plant, ofwel van een generatief naar een vegetatief ontwikkelingsprogramma. Het belangrijkste kenmerk, of eindpunt, van de transitie van zaad naar plant is het op gang komen van de fotosynthese en de gelijktijdige overgang van heterotrofie naar autotrofie. Het is voor een soort voordelig om de periode van vestiging van de zaailing zo kort mogelijk te houden omdat jonge zaailingen erg gevoelig zijn voor biotische en abiotische stress. Dit omdat zij de extreme stresstolerantie voor o.a. uitdroging, die zaden bezitten, verloren hebben. Als de regulerende principes van de transitie van zaad naar zaailing beter kunnen worden begrepen, ontstaat wellicht de mogelijkheid om de stresstolerantie van het zaad te behouden tot in het zaailingstadium.

Ondanks de enorme impact van zaailingeigenschappen op vestiging en opbrengst van het gewas, is de zaad naar zaailing transitie nog nauwelijks bestudeerd op moleculair niveau. Daarom richt dit proefschrift op het ontrafelen en begrijpen van de moleculaire processen die deze transitie reguleren in *Arabidopsis thaliana*. Een hoge-resolutie studie van de moleculaire gebeurtenissen die zich voordoen tijdens de opeenvolgende overgangsstadia kan aanwijzingen opleveren ten aanzien van de regulerende principes die aanzetten tot de transitie. De studie kan ook informatie opleveren over de factoren die kritisch zijn voor het handhaven en verliezen van uitdroogtolerantie en andere stressresponsen.

In **Hoofdstuk 1** beschrijf ik belangrijke processen die betrokken zijn bij kieming en zaailingontwikkeling en bij dehydratietolerantie, waaronder abscicinezuur (ABA) signalering, en regulatie van deze processen. Ik behandel de wijze waarop de zaad-zaailing transitie gerelateerd kan zijn aan een eigenschap zoals uitdroogtolerantie. Ik presenteer een overzicht van de huidige kennis omtrent de zaad-zaailing transitie en het bestaan van een 'point-of-no-return' die gemanipuleerd kan worden door een osmotische behandeling en door abscisinezuur resulterend in de herinductie van uitdroogtolerantie en daarmee mogelijk ook een reactivatie van het zaadafrijpingsprogramma.

In **Hoofdstuk 2** wordt de regulatie van genexpressie in de zaad-zaailing transitie bestudeerd door middel van een gedetailleerde transcriptionele analyse van de zeven ontwikkelingsstadia. Het hoofdstuk beschrijft tevens de afleiding van een gen coexpressie netwerk en verscheidene transcriptionele modules. We laten zien dat een dergelijke benadering belangrijke moleculaire processen tijdens de zaadontwikkeling aan het licht kan brengen die waarschijnlijk niet afgeleid zou kunnen worden uit een vergelijkende transcriptionele profilering. Bovendien laten wij zien dat een vermeende sleutelregulator in één van de transcriptionele modules latere zaailingvestiging kan beïnvloeden.

In **Hoofdstuk 3** laten wij zien hoe de transitie van zaad tot zaailing wordt

weergegeven in de primaire metabolietprofielen en hoe deze correleren met genexpressie. Metaboliet-metaboliet correlatieanalyse suggereren het bestaan van twee verschillende profielen die, respectievelijk, verwijzen naar de metabolische voorbereiding tot kieming en tot een krachtige vestiging van de zaailing. Bovendien verschaft lineaire correlatie tussen metaboliet- en transcripthoeveelheden en een globaal inzicht in de transcriptionele en metabolische veranderingen tijdens de transitie.

Hoofdstuk 4 beschrijft een nieuwe methode om de herinductie van uitdroogtolerantie te bestuderen tijdens de transitie van zaad naar zaailing, zonder nadelige effecten zoals die kunnen worden veroorzaakt door een osmotische behandeling met polyethyleen glycol. Door gebruik van deze methode, 'Mild Air Drying Treatment' (MADT) genaamd, laten wij zien dat uitdroogtolerante zaden (gekiemde zaden in het stadium van worteldoorbraak) en uitdrooggevoelige zaden (in het wortelhaarstadium) vergelijkbare hoeveelheden ABA accumuleren. Wij concluderen dat de herinductie van uitdroogtolerantie niet gerelateerd is aan het afgenomen vermogen van uitdrooggevoelige zaden om ABA te accumuleren. Gebruik makend van het MADT systeem presenteer ik ook een genetische studie van de interactie van *ABSCISIC ACID INSENSITIVE* (*ABI*) genen in hun respons op de herinductie van uitdroogtolerantie. De interactie tussen *ABI3* en *ABI4* en tussen *ABI4* en *ABI5* is synergistisch in de herinductie van uitdroogtolerantie.

In een meer diepgaande studie in **Hoofdstuk 5** heb ik een transcriptoomanalyse uitgevoerd om daaruit mogelijke mechanismen van de herinductie van uitdroogtolerantie te kunnen afleiden. Met gebruik van het MADT protocol werden mogelijke onderliggende mechanismen van uitdroogtolerantie afgeleid door het toepassen van een tijdafhankelijke vergelijking van gekiemde uitdroogtolerante en –gevoelige zaden. Vroege-respons genen in de herinductie van uitdroogtolerantie spelen mogelijk een rol in gebeurtenissen die de initiële bescherming tegen dehydratatiestress bevorderen terwijl de late-respons genen een rol kunnen spelen in processen die het zaad ondersteunen in de reactie op veranderingen in waterdynamica. Door het gebruik van een gen-coexpressie netwerk en een transcriptionele module concludeer ik bovendien dat de 'crosstalk' tussen ABA-afhankelijke en ABA-onafhankelijke transcriptiefactoren de herinductie van uitdroogtolerantie reguleert.

In **Hoofdstuk 6** bespreek ik hoe de in dit proefschrift gepresenteerde resultaten bijdragen aan onze kennis over de moleculaire basis van de transitie van zaad naar zaailing en de herinductie van uitdroogtolerantie tijdens de faseveranderingen. Tenslotte worden nieuwe mogelijkheden voor de toekomst besproken alsmede het verdere gebruik van de datasets om de mechanismen te ontrafelen die ten grondslag liggen aan de transitie van zaad naar zaailing en uitdroogtolerantie. Ook worden mogelijke toepassingen voor gewasverbetering besproken.

Acknowledgements

Conducting a PhD project is an arduous and enjoyable experience. It is just like walking up a high mountain; step by step, accompanied with hardships, frustration, bitterness, encouragement and confidence and, of course, with the kind support of many people. Now that I find myself at the enjoyable scenery, I realize that it was, in fact, teamwork that got me here. I am very thankful to all who helped me.

I thank the "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior" (CAPES, Brazil) for financial support of my PhD study at Wageningen University.

My sincere thanks go to my honorific supervisor, Dr. Henk Hilhorst, who accepted me as his PhD student without any hesitation. Thereafter, during my PhD, you gave me so much advice and patiently supervised me, which helped me to get the right direction. I have learned so much about seed science from you; without your help I could not have finished my thesis successfully. Besides, words are not enough to expresses what your friendship means to me. I am waiting for you in Brazil, so we can have a typically Brazilian BBQ.

A special thanks is also given to Dr. Wilco Ligterink, my daily supervisor. You are one of the most kind and generous persons I have met in science. Your encouragement and help, made me feel confident to achieve my desire and to overcome every difficulty I encountered. I also appreciate the advice of professor Dr. Harro Bouwmeester, our group leader. At the last stage of my dissertation, he suggested possible improvements. It is not enough to express my gratitude with only a few words. My appreciation goes also to Dr. Leonie Bentsink with whom, in the early stages of my PhD, I had the privilege to be your neighbor in the open office. Many thanks Leonie for all your advise and talks about science. It was a real pleasure to have somebody like you next to my desk. The quality of this thesis is also due to you.

I also appreciate the advice of other PPH members from the PhD meetings; Dr. Sander van der Krol, Dr. Dick Vreugdenhil and Dr. Richard Immink, for your critical comments during my PhD, which enabled me to notice the weaknesses of my research and to make the necessary improvements according to your comments.

Concerning the Seed Lab, I only can say that it is a very pleasant group to work in. Starting with the great technician Leo Willems; Mr. Leo thanks for the day-to-day support in the lab. I think you also should, at least once, record yourself singing. You will hear how bad it is. Besides, every party would have been boring without you (*let's finish this whole crate of beer?)*. I will never forget that day. Concerning the other technician Juriaan Rienstra (*the cloning guy*); We have talked about many *"crazy"* science. Juriaan, you are good in molecular cloning, but forget about asking him something easy as taking photos with the germinator. *Did you flip the photos?* uɐɹnɟ ; this was the reason why I only had half of my photos. Anyway, thanks Juriaan. I also would like to express my appreciation to each one of you and your

uniqueness. Bing Bai (*are you ok? Keep track in the gym*), Gonda Buijs, Alexandre (*thanks for helping me with designing the cover! You have to accept science "it is what it is"*), Deborah (*thanks for all the moments with the little kids*), Nafiseh, Hanzi (*your wedding was lovely! It was a great pleasure to be there*), Melissa (*you are a really kind person*), Natalia Moreno (*chica*), Phuong (*you are not anymore in the seed lab, but how can I forget you. You are a great person! Keep running!*), Harm Nijveen (*thanks for our discussions about data analysis*), Renake (*did you get used to Leo's desk chaos?*), Dr. Paulo Ribeiro (*"I agree with you to not disagree". It was always good to discuss the scientific results with you. I wish you all the best in your further career*), Paulo Teixeira (*"the truco man". Thanks to bring the cards!*) Elise, Mariana and Farzaneh. You all contributed to the nice work atmosphere. Yes! I did not forget you: my especial thanks goes to Dr. Bas Dekker (*"the maniatis man"*). Thanks for all our talks about good science and experiments, which mostly occurred in the weekends. I have learned a lot from you. Your talent for science is impressive. I am really glad that I had the opportunity to work with you. I really respect your knowledge. You contributed a lot to my work and also as being my friend; I am glad to have you as my paranymph. Thanks! I also would like to thank Rina. Many thanks for helping with all those forms and in the organization of all good PPH events.

Outside of the seed lab, I would like to express my appreciation to the PPH friends. Beatriz (*Bea, you are a great runner*), Arman (*we should play football*), Jun, Mark (*have you taken something to run that much in Veluweloop?*), Esmer, Karen (*are you being mean? You are great, I like your good mood*), Giovanni (*Bravo ragazzo! Keep the good things*), Nasr, Umidjon, Mahdere and Abi (*great that the kids could play together! I enjoyed all the moments that we had together),* Manus (*the great musician*) and Emilie (*thanks for helping with the formatting of this thesis and all the talks in the lab*). I also would like to thanks all PPH technicians; Kristýna, Diaan, Francel and Kerstin, to share their knowledge with me during my PhD.

I would like to thank Edvaldo Aparecido Amaral da Silva (Amaral), who introduced me to the seed science community and to Henk. Thanks!

I would like to convey my thanks to people from my home university (Universidade Federal de Lavras); Luciano, Douglas, Alan and Chalfun. You helped me in one way or another, offering me an ideal environment during my master. I reached this point today, due to all of you. With the knowledge acquired at LCBM, I felt free and could concentrate better on my research at WUR. I give my sincere thanks to all of you.

There are two Brazilians I need to mention especially, Julio Maia and Charles, for your friendship and unselfish help, particularly in my first months in the Netherlands, and thanks to share your house. I owe both of you my sincere gratitude for your generous and timely help.

I also had a pleasant time with the Brazilian people living in Wageningen. Starting with Fernanda and Mauricio (*guys thanks for the AC/DC concert, it was a pleasure to be in this concert with you*). I really enjoy your friendship. Anabele and Wiebe (*thanks to have the door of your house always open for me and my family*). All the parties and BBQs that we had together were just amazing. Cinara and Jânio (*thanks for all your help in Netherlands*). No! I did not forget the greatest couple, Carol and Dennis. Guys I have no words to express my gratitude to you. Dennis, your help in getting me and my family settled in Wageningen is unmeasurable. Besides, I also enjoyed the beers that we brewed together at home and, of course, all the others that we drank. I am really glad for your new family member (Bento) and to have you as my paranymph. I wish you and your family all the best.

I would also like to give my appreciation to the three musketeers (Rodrigo, Emerson and Vinicius) in Brazil; you are good friends guys. Although, you are completely outside of science, you gave all the support to do this PhD abroad. Despite the fact that Vinicius did not visit me, I still will be your friend. Thanks for everything guys.

I am very grateful to all my relatives in Brazil for their absolute love and support. Too numerous people that have contributed. To my wife's family: Sr. Alvaro (Vico) e Sra. Nené, muito obrigado pelo apoio de vocês, foram fundamentais para essa conquista. Neys e Danda, obrigado pelo carinho de vocês. Mara, Erbet, Ana Clara e Iasmin vocês são exceptionais, obrigado por tudo. Vocês todos são um belo exemplo de vida, das quais quero continar tirando as melhores lições. To my relatives: Vó Maria, que exemplo de pessoa, obrigado por todo apoio. Amo você vó. Tia Dedé, Tio Zé, Tia Marlene, Tio Toninho e Tia Nené, Obrigado! Júnior e Beth , foi ótimo nossas viagens. Paris ficou pequena. Obrigado por todo apoio de vocês. Edna e Du, sempre com palavras calmas, mesmo longe torceram por mim. Obrigado! To my parents: Pai e Mãe, não tenho palavras para agradecer à vocês. Obrigado pela educação que vocês me deram e por mostrar que a simplicidade é a essência da vida. A compreensão e o amor de vocês me encorajou a trabalhar duro durante meu doutorado. A personalidade firme e bondosa de vocês afetou-me para ser firme e nunca desistir diante das dificuldades. À minha irmã Andressa, obrigado por tudo. Obrigado por fazer parte de minha vida. Amo vocês.

At last, but not at least, I am greatly indebt to my dear wife Roseane and my daughter Rafaela. You form a backbone for my happiness. Your love and support without complain has enabled me to succeed in this PhD. I know that it was not an easy thing for you, Roseane, to take care of Rafaela, being both a mother and father, while I was away. All the responsibilities were on your shoulders. I owe every single achievement to both of you. I love you.

Muito Obrigado!

Many Thanks!

Bedankt!

About the author

Curriculum vitae

Anderson Tadeu Silva was born on 01st September, 1985 in São Paulo, Brazil. He studied Biology between 2005 and 2008 at Centro Universitário de Lavras (UNILAVRAS) in Brazil. During his studies he was actively involved in molecular biology in the Laboratório Central de Biologia Molecular (LCBM) at Universidade Federal de Lavras (UFLA). At LCBM, he worked in the big project of Genome sequencing of the bacterium *Corynebacterium pseudotuberculosis*. Later, he carried out project on gene molecular expression and cloning. For his bachelor thesis he worked on a molecular identification and characterization of endo-beta-mannanase in *Genipa americana* L. seeds. In his master thesis, he studied genetic expression of Baby Boom (*BBM*) and Somatic Embryogenesis Receptor Kinase (*SERK*) during acquisition of somatic embryos in *Coffea arabica* L. After his master thesis, he starting writing a PhD-proposal to work with seeds at Wageningen University, which was, after 10 months of patiently of waiting, rejected. Later, he starting writing another PhD-proposal together with Dr. Henk Hilhorst. Together, they decided to give it a second try. In the meantime, he had started PhD in plant biotechnology at UFLA. Finally, his PhD-proposal was accepted and he was able to start his PhD on 03rd January, 2012 at Wageningen University, Laboratory of Plant Physiology under supervision of Prof. Harro Bouwmeester, Dr. Henk Hilhorst and Dr. Wilco Ligterink. In the research he integrated approach of transcriptomics and metabolomics and focused on the seed-to-seedling transition and implications to desiccation tolerance, of which this thesis is the final result.

Publications

Queiroz SE, da Silva EA, Davide AC, Jose AC, **Silva AT**, Fraiz AC, Faria JM, Hilhorst HW. **2012**. Mechanism and control of *Genipa americana* seed germination. *Physiologia Plantarum* 144: 263-276.

Perreira LM, Resende MLV, Mathioni SM, Campos MA, Ribeiro JPM, Camilo FR, **Silva AT**, Paiva LV. **2013**. Protective effect and expression of defense-related ESTs induced by acibenzolar-S-methyl and a phosphorylated mannan oligosaccharide-based product against Moniliophthora perniciosa in *Theobroma cacao*. *African Journal of Biotechnology* 12: 1311-1317.

Silva AT, Paiva LV, Andrade AC, Barduche D. **2013**. Identification of expressed sequences in the coffee genome potentially associated with somatic embryogenesis. *Genetics and Molecular Research* 12: 1698-1709.

Silva AT, Barduche D, do Livramento K, Ligterink W, Paiva L. **2014**. Characterization of a Putative *Serk*-Like Ortholog in Embryogenic Cell Suspension Cultures of *Coffea arabica* L. *Plant Molecular Biology Reporter* 32: 176-184.

Silva AT, Barduche D, do Livramento K, Paiva L. 2014. A putative *BABY BOOM*-like gene (*CaBBM*) is expressed in embryogenic calli and embryogenic cell suspension culture of *Coffea arabica* L. **2014**. *In Vitro Cellular & Developmental Biology - Plant* 50: 1-9.

Abreu JRd, Paiva LV, Rodríguez MAD, **Silva AT**, Henriques AR, Chalfun-Junior A. **2015**. Identification and quantification of differentially expressed genes associated with citrus blight (*Citrus* spp.). *Ciência e Agrotecnologia* 39: 32-38.

Silva AT, Ligterink W, Hilhorst HWM. Combining co-expression networks and transcriptional modules identifies novel genes controlling the seed-to-seedling phase transition in *Arabidopsis thaliana*. (*Submitted*).

Drost HG, Bellstädt J, **Silva AT**, Ó›Maoiléidigh DS, Gabel A, Weinholdt C, Ryan PT, Dekkers BJW, Bentsink L, Hilhorst H, Ligterink W, Wellmer F, Grosse I, Quint M. Post-embryonic hourglass patterns mark ontogenetic transitions in plant development. (*Submitted*).

Silva AT, Ligterink W, Hilhorst HWM. Metabolie profilling reveals two metabolic shifts during the seed-to-seedling transition in *Arabidopsis thaliana*. (*in preparation for submission*).

Silva AT, Dekkers BJW, Ribeiro PR, Rienstra JA, Willems LAJ, Floková K, Ligterink W, Hilhorst HWM. ABA-dependent and independent transcription factors control the re-establishment of desiccation tolerance in germinated seeds of Arabidopsis. (*in preparation for submission*).

Education Statement of the Graduate School

Experimental Plant Sciences

Issued to: **Anderson Tadeu Silva**
Date: **9 December 2015**
Group: **Laboratory of Plant Physiology**
University: **Wageningen University & Research Centre**

1) Start-up phase	*date*
► **First presentation of your project**	
The transition from seed to seedling and implications for stress resistance	Sep 17, 2012
► **Writing or rewriting a project proposal**	
The transition from seed to seedling and implications for stress resistance	Oct 08, 2012
► **Writing a review or book chapter**	
► **MSc courses**	
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*7.5 credits**

2) Scientific Exposure	*date*
► **EPS PhD student days**	
EPS PhD student day, University of Amsterdam	Nov 30, 2012
EPS PhD student day, Leiden University	Nov 29, 2013
► **EPS theme symposia**	
EPS theme 3 'Metabolism and Adaptation', Utrecht University	Apr 26, 2012
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', Wageningen University	Feb 10, 2015
► **NWO Lunteren days and other National Platforms**	
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 22-23, 2013
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 14-15, 2014
ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 13-14, 2015
► **Seminars (series), workshops and symposia**	
Lessons from photosynthetic analysis in three widely used Arabidopsis ecotypes - Cornelia Spetea Wiklund	Feb 21,2012
Can ecologists shape cities?' - Remco Daalder (City Ecologist of Amsterdam, Amsterdam, The Netherlands)	May 24, 2012
Parenting in plants: maternal control of seed dormancy - Steven Penfield (University of Exter, Exter, UK)	Jun 12, 2012
Genomic Selection - Osé Crossa	Jun 14, 2012
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 - Jill Farrant	Jun 26, 2012
Genetic modification for iron biofortification and drought tolerance in rice - Inez Hortenze Slamet-Loedin (International Rice Research Institute (IRRI), The Philippines)	Jun 29, 2012
Genotype–phenotype mapping in a post-GWAS world - Lauren McIntyre (University of Florida, USA)	Sep 17, 2012
Strigolactones, new plant hormones. Importance of their stereochemistry for bioactivity as germination stimulant - Prof. Yukihiro Sugimoto	Oct 16, 2012
ABA signaling networks in Arabidopsis - Ruth Finkelstein (University of California, Santa Barbara, USA)	Nov 14, 2012
Tackling natural variance in seed metabolism integrating metabolite profiles via network analysis - Aaron Fait (Ben-Gurin University of Negev, Ben-Gurion, Israel)	Dec 04, 2012
Arabidopsis thaliana as a model system for the study of evolutionary questions - Dr. Detlef Weigel (Max Planck Institute for Developmental Biology, Germany)	Feb 27, 2013
'Plant drug smugglers' about transport of secondary metabolites in plants - Prof. Marc Boutry (University de Louvain, Belgium)	Mar 13, 2013
Iridoid glycosides: biochemistry and role in biotic interactions in ribwort plantain - Dr. Arjen Biere (NIOO-KNAW, Wageningen)	Apr 10, 2013
Anticipating critical transitions - Marten Scheffer (Wageningen University, The Netherlands)	May 23, 2013
Metabolomics in the lab: a myriad of applications - Robert Hall, Plant Research International and Laboratory of Plant Physiology, Wageningen University	Oct 08, 2013
Elucidating the Biosynthetic Pathway for Vibralactone in the Basidiomycete Fungus Boreostereum vibrans - Ying Zeng (Kunming Institute of Botany, Chinese Academy of Sciences)	Oct 24, 2013
Towards understanding rice brown spot, a disease induced by physiological stress - Professor Monica Höfte (Ghent University, Belgium)	Feb 06, 2015
The evolutionary significance of gene and genome duplications - Prof.dr. Yves van de Peer (University of Ghent, Belgium)	Feb 03, 2015
Tomato metabolomics in 2015, the difference a genome makes - Dr. Alisdair Fernie (Golm, Germany)	Mar 11, 2015
Evolution of floral signals in plants: mechanisms and consequences - Prof.dr. Florian Schiestl (University of Zürich, Switzerland)	Mar 12, 2015

CONTINUED ON NEXT PAGE

Inferring species trees given coalescence and reticulation - Michael D. Pirie (Department of Biochemistry, University of Stellenbosch South Africa)	Mar 18, 2015
Epigenetic regulation of seed dormancy and germination in Arabidopsis - Dr. Yongxiu Liu (Institute of Botany, Chinese Academy of Sciences, China)	Mar 17, 2015
Structure and evolution of centromeres: lessons learned from plants - Dr. Jiming Jiang (University of Wisconsin, Madison, USA)	Apr 01, 2015
The uncovering of new transcription factors involved in ripening and post-harvest life of tomato and strawberry fruits - Dr. Sonia Osorio (University of Malaga, Spain)	Apr 15, 2015
Hormone-related Functions of LEC2 in Somatic Embryogenesis Induction in Arabidopsis - Dr. Malgorzata Gaj (Malgorzata Gaj, University of Silesia, Poland)	Apr 16, 2015
Polyploidy in wild relatives of soybean and other legumes: systematics, comparative and functional genomics, and nodulation - Prof. Dr. Jeff Doyle (Cornell University, USA)	May 12, 2015
► **Seminar plus**	
► **International symposia and congresses**	
4th Workshop on the Molecular Aspects of Seed Dormancy and Germination, Paris, France	Jul 09-12, 2013
11th Conference of the International Society for Seed Science, Chansha, China	Sep 15-19, 2014
► **Presentations**	
Oral Presentation	
Metabolic and transcriptional dynamics changes during seed-to-seedling transition in Arabidopsis - The 11th conference of the international society for seed science (ISSS), Ghangsha, China	Sep 15-19, 2014
Metabolic and transcriptional dynamics changes during seed-to-seedling transition in Arabidopsis - 3 Ducth seed symposium, Wageningen, NL	Oct 07, 214
Poster Presentation	
Metabolite profiling of the seed-to-seedling transition in Arabidopsis thaliana - 4th Workshop on the Molecular Aspects of Seed Dormancy and Germination, Paris, France	Jul 09-12, 2013
Metabolite profiling of the seed-to-seedling transition in Arabidopsis thaliana - NWO-ALW meeting ' Experimental Plant Science', Lunteren, NL	Apr 22-23, 2013
Analysis of regulatory networks during seed-to-seedling transition in Arabidopsis thaliana - NWO-ALW meeting 'Experimental Plant Science', Lunteren, NL	Apr 14-15, 2014
► **IAB interview**	
► **Excursions**	
Subtotal Scientific Exposure	*13.9 credits**

3) In-Depth Studies	*date*
► **EPS courses or other PhD courses**	
Systems Biology: statistical analysis of ~omics data	Dec 10-14, 2012
Microscopy and Spectroscopy in Food and Plant Sciences	May 07-11, 2012
Introduction to R for statistical analysis	Jun 10-11, 2013
► **Journal club**	
Literature discussion in Plant Physiology	2012-2015
► **Individual research training**	
Subtotal In-Depth Studies	*6.6 credits**

4) Personal development	*date*
► **Skill training courses**	
EndNote Introduction	Oct 30-31, 2012
Techniques for writing and Presenting a Scientifc Paper	Apr 21-24, 2015
Career Orientation	Mar 06,13, 20, 27, 2015
► **Organisation of PhD students day, course or conference**	
► **Membership of Board, Committee or PhD council**	
Subtotal Personal Development	*3.3 credits**

TOTAL NUMBER OF CREDIT POINTS*	**31.3**

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.*

This work was performed at the Laboratory of Plant Physiology, Wageningen University, and financially supported by the 'Cordenação de Aperfeiçoamento de Pessoal de Nivel Superior' (CAPES, Brazil).

Cover design by: Anderson Tadeu Silva and Alexandre Marques

Cover photos: Seven developmental stages of the seed-to-seedling transitionr

Thesis Layout by: Anderson Tadeu Silva

Drawings of the seed-to-seedling transition stages by: Julia Bellstädt

Printed by: Proefschriftmaken.nl || Uitgeverij BOXPress