the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2010) **62**, 39–51 doi: 10.1111/j.1365-313X.2010.04135.x

Genome-wide analysis of endogenous abscisic acid-mediated transcription in dry and imbibed seeds of Arabidopsis using tiling arrays

Masanori Okamoto[1], Kiyoshi Tatematsu[2], Akihiro Matsui[1], Taeko Morosawa[1], Junko Ishida[1], Maho Tanaka[1], Takaho A. Endo[3], Yoshiki Mochizuki[3], Tetsuro Toyoda[3], Yuji Kamiya[1], Kazuo Shinozaki[1], Eiji Nambara[4] and Motoaki Seki[1,5,*]

[1]*RIKEN Plant Science Center, Yokohama, Kanagawa 230-0045, Japan,*

[2]*National Institute for Basic Biology, Okazaki, Aichi 444-8585, Japan,*

[3]*RIKEN Bioinformatics and Systems Engineering Division, Yokohama, Kanagawa 230-0045, Japan,*

[4]*Department of Cell & Systems Biology, University of Toronto, Toronto, Ontario M5S 3B2, Canada, and*

[5]*Kihara Institute for Biological Research, Yokohama City University, Yokohama, Kanagawa 244-0813, Japan*

Received 27 August 2009; revised 14 December 2009; accepted 16 December 2009; published online 16 February 2010.
*For correspondence (fax +81 45 503 9586; e-mail mseki@psc.riken.jp).

SUMMARY

The phytohormone abscisic acid (ABA) plays important roles in the induction and maintenance of seed dormancy. Although application of exogenous ABA inhibits germination, the effects of exogenous ABA on ABA-mediated gene transcription differ from those of endogenous ABA. To understand how endogenous ABA regulates the transcriptomes in seeds, we performed comprehensive expression analyses using whole-genome Affymetrix tiling arrays in two ABA metabolism mutants – an ABA-deficient mutant (*aba2*) and an ABA over-accumulation mutant (*cyp707a1a2a3* triple mutant). Hierarchical clustering and principal components analyses showed that differences in endogenous ABA levels do not influence global expression of stored mRNA in dry seeds. However, the transcriptome after seed imbibition was related to endogenous ABA levels in both types of mutant. Endogenous ABA-regulated genes expressed in imbibed seeds included those encoding key ABA signaling factors and gibberellin-related components. In addition, cohorts of ABA-upregulated genes partially resembled those of dormant genes, whereas ABA-downregulated genes were partially overlapped with after-ripening-regulated genes. Bioinformatic analyses revealed that 6105 novel genes [non-Aradopsis Genome Initiative (AGI) transcriptional units (TUs)] were expressed from unannotated regions. Interestingly, approximately 97% of non-AGI TUs possibly encoded hypothetical non-protein-coding RNAs, including a large number of antisense RNAs. In dry and imbibed seeds, global expression profiles of non-AGI TUs were similar to those of AGI genes. For both non-AGI TUs and AGI code genes, we identified those that were regulated differently in embryo and endosperm tissues. Our results suggest that transcription in Arabidopsis seeds is more complex and dynamic than previously thought.

Keywords: abscisic acid, Arabidopsis, seed dormancy, germination, tiling array, non-protein-coding RNA.

INTRODUCTION

Germination is a critical event for survival and reproductive success in the plant life cycle. Therefore, seed dormancy is an adaptive trait in unfavorable conditions. Primary dormancy is acquired during seed maturation and is broken during storage of dry seeds, i.e. at after-ripening (Karssen *et al.*, 1983). After-ripened seeds enter a secondary dormant state in unfavorable germination conditions. In most plant species, seed dormancy and germination are controlled by several environmental factors such as light, temperature, nutrients and other seed storage conditions (Baskin and Baskin, 1998). Many environmental factors alter the metabolism and signaling of two plant hormones, abscisic acid (ABA) and gibberellins (GAs) (Finkelstein *et al.*, 2008; Holdsworth *et al.*, 2008). Abscisic acid promotes the induction and maintenance of seed dormancy, whereas GA is required for the initiation and completion of germination. It is thought that germination is regulated by the antagonistic effects of ABA and GA.

In typical angiosperm seeds, germination is controlled by both the growth potential of the embryo and the restrictive potential of tissues surrounding it. Embryo growth is triggered by increasing the pressure potential and extensibility of the cell wall. These changes allow protrusion of the radicle, which marks the completion of germination (Bewley, 1997). Abscisic acid arrests embryo growth by inhibiting the extensibility of the embryonic cell wall (Schopfer and Plachy, 1985; da Silva *et al.*, 2004). On the other hand, the endosperm and testa tissues surrounding the embryo act as a mechanical barrier to germination (Bewley, 1997; Muller *et al.*, 2006). Although the testa is a dead tissue in imbibed seeds, abnormal testa mutants in Arabidopsis (*Arabidopsis thaliana*) show a reduced dormancy phenotype, indicating that testa components contribute to seed dormancy and germination (Debeaujon and Koornneef, 2000; Léon-Kloosterziel *et al.*, 1994). In contrast, endosperm is a living tissue in Arabidopsis and produces enzymes related to cell wall modification. Activity of these enzymes is induced in response to GA prior to germination (Halmer *et al.*, 1976; Chen and Bradford, 2000; Nonogaki *et al.*, 2000; Wu *et al.*, 2001). Interestingly, endosperm weakening is also antagonistically regulated by ABA and GA in cress (*Lepidium sativum*), a close relative of Arabidopsis (Muller *et al.*, 2006).

In many cases, high levels of endogenous ABA are associated with physiologically dormant states. Levels of ABA change drastically during seed development and seed imbibition in response to developmental and environmental cues. The ABA accumulated during maturation is essential for inducing and maintaining seed dormancy (Karssen *et al.*, 1983; Koornneef *et al.*, 1989). Therefore, freshly harvested seeds of ABA-deficient mutants of Arabidopsis, tomato (*Lycopersicon esculentum*) and tobacco (*Nicotiana plumbaginifolia*) fail to induce lasting seed dormancy (Karssen *et al.*, 1983; Groot and Karssen, 1992; Grappin *et al.*, 2000). In several plant species, imbibed dormant seeds accumulate more ABA than imbibed non-dormant seeds (Grappin *et al.*, 2000; Jacobsen *et al.*, 2002; Ali-Rachedi *et al.*, 2004). Seed dormancy is maintained via activation of *de novo* ABA biosynthesis after seed imbibition. Thus, seeds of ABA biosynthesis mutants can germinate in unfavorable conditions, such as darkness after irradiance of far-red light, high temperature and high salinity (González-Guzmán *et al.*, 2004; Seo *et al.*, 2006; Tamura *et al.*, 2006). In contrast, activation of ABA catabolism is enhanced by after-ripening in imbibed seeds, and the levels of ABA decrease and remain at low levels until germination (Millar *et al.*, 2006). Therefore, ABA catabolism is one of the factors that regulate seed dormancy and germination.

Genetic screening by application of exogenous ABA has identified a large number of ABA signaling factors associated with seed dormancy and germination. Several protein phosphatase 2Cs (e.g. ABI1, ABI2, AHG1 and AHG3) and transcription factors (e.g. ABI3, ABI4, ABI5 and CHO1) are key components in the ABA signal transduction pathway (Finkelstein *et al.*, 2008; Holdsworth *et al.*, 2008; Yamagishi *et al.*, 2009). Application of exogenous ABA to imbibed seeds is useful for identification of ABA signaling components and experimental manipulation to mimic dormancy states. Indeed, exogenous ABA can reduce germination potential and delay endosperm rupture in non-dormant Arabidopsis seeds (Chibani *et al.*, 2006; Muller *et al.*, 2006). Moreover, exogenous ABA repressed the expression of a GA biosynthesis gene in the embryo of sorghum (*Sorghum bicolor*) (Pérez-Flores *et al.*, 2003). In contrast, it has been pointed out that application of exogenous ABA to non-dormant seeds does not reflect transcriptome and proteome levels in dormant seed states (Chibani *et al.*, 2006; Carrera *et al.*, 2008). For example, metabolic and physiological processes in seeds treated with exogenous ABA differ from those of dormant seeds (Pritchard *et al.*, 2002; Penfield *et al.*, 2004; Muller *et al.*, 2006).

To date, transcriptome data from Arabidopsis seeds in various different physiological states have been obtained using Affymetrix ATH1 arrays. These data are available from the Arabidopsis eFP Browser (Winter *et al.*, 2007). Although they are useful, recent transcriptome analyses have revealed that in Arabidopsis there are a large number of novel transcripts including non-protein-coding RNAs (Yamada *et al.*, 2003; Stolc *et al.*, 2005; Zhang *et al.*, 2006; Matsui *et al.*, 2008; Zeller *et al.*, 2009; Hazen *et al.*, 2009). To understand how endogenous ABA controls seed dormancy and germination, we performed comprehensive transcriptome analyses of ABA-deficient and ABA over-accumulation mutants using whole-genome Affymetrix tiling arrays. In addition, we also compared the transcriptome profiles between embryo and endosperm tissues to understand tissue-specific expression patterns. We characterized the expression patterns of both Arabidopsis Genome Initiative (AGI) code genes and novel genes (non-AGI transcriptional units) including a large number of non-protein-coding RNAs in Arabidopsis seeds.

RESULT AND DISCUSSION

Control of seed dormancy and germination by endogenous ABA

In many plant species determination of seed dormancy and germination is related to the endogenous ABA levels, which are controlled by a combination of ABA biosynthesis and catabolism (Nambara and Marion-Poll, 2005). A key enzyme in ABA catabolism, CYP707A, was reported to be important for breaking seed dormancy (Kushiro *et al.*, 2004; Millar *et al.*, 2006; Okamoto *et al.*, 2006). However, seed dormancy of *cyp707a* single and double mutants is completely broken by stratification or after-ripening for a short period (Okamoto *et al.*, 2006). To find out if CYP707A contributes to breaking

seed dormancy, we generated a triple mutant defective in *CYP707A1*, *CYP707A2* and *CYP707A3*. The ABA levels in dry seeds of the *cyp707a1a2a3* triple mutant were 70-fold higher than those in the wild type and remained at high levels after seed imbibition (Figure 1a). This over-accumulation of ABA was more prominent in the *cyp707a1a2a3* triple mutant than in *cyp707a* double mutants (Okamoto *et al.*, 2006). Freshly harvested *cyp707a1a2a3* triple-mutant seeds did not germinate at 22°C after stratification for 3 days (Figure 1b,c). In addition, a long period of dry storage (6 months) could not rescue hyper-dormancy of the *cyp707a1a2a3* triple-mutant seeds. Seed dormancy of the *cyp707a1a2a3* triple mutant is stronger than that of Cape Verde Islands (Cvi), a dormancy accession of Arabidopsis, because Cvi seeds germinated after 6 months of dry storage in our conditions (Figure S1 in Supporting Information). A previous study showed that the *cyp707a1a2* double mutant was hyper-dormant, but this dormancy was easily released by after-ripening. The present results demonstrate that ABA 8′-hydroxylases are essential for regulating seed dormancy, and loss of function of the three CYP707A genes led to prominent dormancy, which was not easily released by after-ripening. In fact, germinating seeds showed high levels of *CYP707A* expression while non-germinating seeds did not (Millar *et al.*, 2006; Toh *et al.*, 2008). It is worth mentioning that seed dormancy of the *cyp707a1a2a3* triple mutant was broken by a combination of after-ripening and stratification or GA treatment (Figure 1c). It is possible that dormancy release in the *cyp707a1a2a3* triple mutant is accomplished by an ABA 8′-hydroxylase-independent mechanism, for example the ABA conjugate pathway, decreased ABA sensitivity or increased GA sensitivity and/or biosynthetic ability. On the other hand, *de novo* ABA biosynthesis is required to maintain seed dormancy after seed imbibition. Therefore, the ABA biosynthesis mutant *aba2* showed a lesser degree of seed dormancy than the wild type as reported previously (Figure 1a,b; Léon-Kloosterziel *et al.*, 1996).

Next, we examined whether endogenous ABA levels correlate with transcript levels of ABA-responsive and signaling-related genes. In Arabidopsis seeds, *ABI4*, *ABI5*, *AHG1* and *AHG3* are involved in ABA signal transduction, whereas *AtEM6* and *RD29B* are ABA-responsive genes. Interestingly, in dry seeds, transcript levels of these genes did not differ remarkably between the wild type and the mutants (Figure 2). However, after seed imbibition, transcript levels of these genes reflected endogenous ABA levels (Figure 2). In 12, 24 and 36-h imbibed seeds, transcript levels of *ABI5*, *AtEM6*, *RD29B*, *AHG1* and *AHG3* were higher in the *cyp707a1a2a3* triple mutant and lower in the *aba2* mutant, compared with those in the wild type (Figure 2). On the other hand, induction of *ABI4* in the *cyp707a1a2a3* triple mutant was repressed compared with that in the *aba2* mutant and wild-type after seed imbibition (Figure 2). Application of exogenous (+)-*S*-ABA (3 μM) inhibited germination of freshly harvested wild-type seeds (data not shown), but its effect on gene expression was relatively minor at early seed imbibition. Transcript levels of *AHG3* and *RD29B*, but not *ABI5*, *AHG1* and *AtEM6*, in the wild type treated with exogenous ABA were slightly higher than those in the control at 24 h after seed imbibition. Application of exogenous ABA to after-ripened seeds does not mimic dormant seed states with respect to transcriptomic and proteomic traits (Chibani *et al.*, 2006; Carrera *et al.*, 2007). In addition, exogenous ABA cannot completely prevent catabolism of fatty acids or lipids, which is necessary for embryo and seedling growth during germination (Pritchard *et al.*, 2002; Penfield *et al.*,

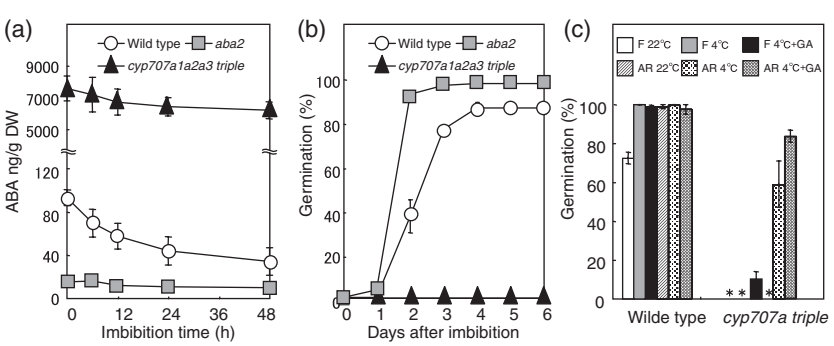

Figure 1. Phenotypic analysis of *aba2* and *cyp707a1a2a3* mutant seeds.
(a) Changes in ABA levels after seed imbibition. Experiments were performed four times using independent seed batches. Average values are shown with standard errors.
(b) Germination tests of freshly harvested seeds of the *aba2* mutant and *cyp707a1a2a3* triple mutant.
(c) Effects of stratification, gibberellin (GA) and after-ripening on *cyp707a1a2a3* triple-mutant seeds. F and AR indicate freshly harvested and after-ripened seeds, respectively. Seeds were stratified for 3 days at 4°C in the dark. GA_4 (10 μM) was applied during stratification. Seeds after-ripened for 6 months were used in these experiments. Wild type and mutants were both *Arabidopsis thaliana* accession Columbia. Asterisk indicates no germination.
In (b) and (c), seeds were sown on 0.5% agarose gel, and kept at 22°C under continuous light for 6 days. The germination rate was scored based on radicle emergence, and approximately 50 seeds were used in the each experiment. Experiments were performed five times using independent seed batches. Average values are shown with standard errors.

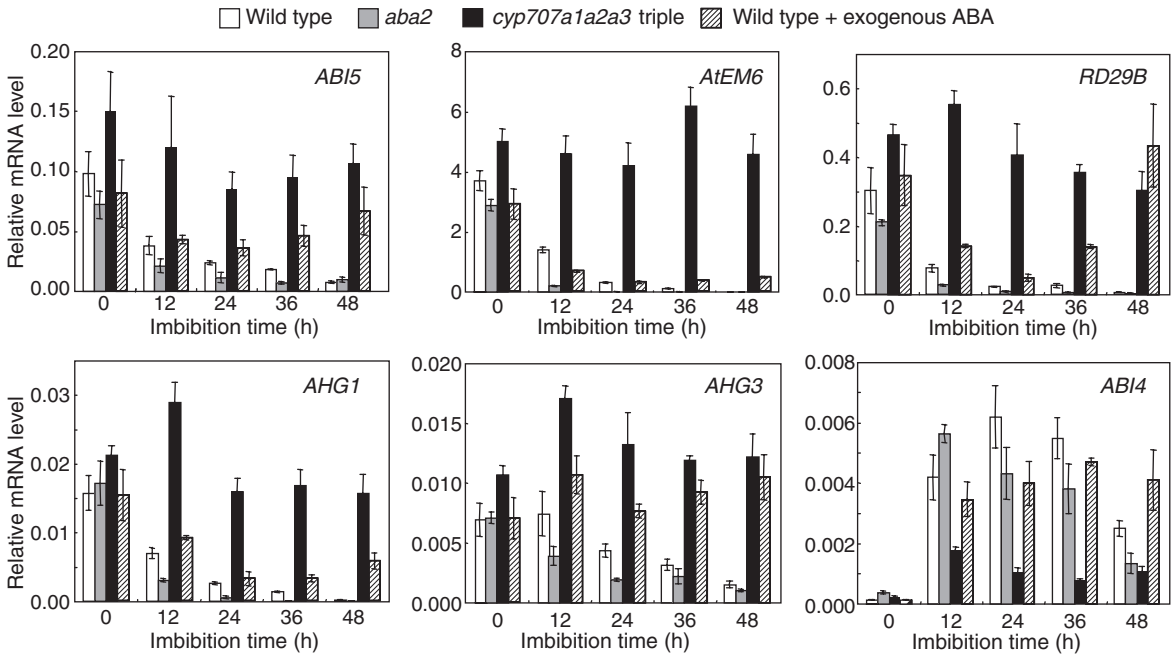

Figure 2. Changes in transcript levels of ABA-responsive and signaling-related genes in *aba2* and *cyp707a1a2a3* mutants and effects of exogenous ABA treatment. Freshly harvested seeds were imbibed on the filter paper containing water. Exogenous ABA was applied as an aqueous solution (3 μM (+)-*S*-ABA) to wild-type seeds. Transcript levels of the genes in dry seeds and imbibed seeds at indicated time points (h) after imbibition were determined by quantitative reverse transcription polymerase chain reaction (qRT-PCR). Experiments were performed four times using independent seed batches. Average values are shown with standard errors.

2004). Our expression analyses indicate that exogenous ABA markedly affected the expression of several genes at a later stage, in particular *RD29B* (Figure 2).

Global expression profiling in seeds by tiling array analysis

To reveal the global transcription profiles regulated by endogenous ABA in Arabidopsis seeds, we carried out comprehensive transcriptome analyses of *aba2* and *cyp707a1a2a3* triple mutants using whole-genome tiling arrays. We also examined tissue-specific transcript accumulation in embryo and endosperm tissues, because interactions between these tissues are a key process in the control of seed dormancy and germination. Total mRNA was extracted from whole dry seeds and 24-h imbibed seeds of the wild type and mutants (*aba2* and *cyp707a1a2a3* triple mutant). We also extracted total mRNA from dissected embryo and endosperm tissues of wild-type seeds after 24-h imbibition. Biotin-labeled cRNAs were hybridized to the whole-genome tiling array set (1.0 F array and 1.0 R array, Affymetrix, http://www.affymetrix.com/).

Tiling array data were analyzed based on the TAIR8 gene model, and the signal intensity of an AGI code gene was calculated by using the Arabidopsis tiling array-based detection of exons (ARTADE) method (Toyoda and Shinozaki, 2005; Matsui *et al.*, 2008). Of a total of 32 144 AGI code genes, 18 848 genes were expressed in at least one condition (P initial value $<10^{-8}$) (Table S1). The 18 848 expressed AGI code genes were analyzed using hierarchical clustering analysis (Figure 3a). There were large differences in endogenous ABA levels among the various seed types. However, in dry seeds, transcript levels of AGI code genes were similar among the wild -type, *aba2* and *cyp707a1a2a3* triple mutant (Figure 3a). After seed imbibition, the global expression patterns of these mutants drastically differed from that of the wild type (Figure 3a). These expression patterns indicate that endogenous ABA levels affect global transcription in association with seed dormancy and germination states after seed imbibition rather than dry seed states. Moreover, the global expression patterns in embryo and endosperm differed from that of imbibed whole wild-type seeds. This result indicates that the transcriptome of seeds is made up of integrated transcriptomes from the embryo and thin endosperm layers. In other words, although the endosperm is a thin single-celled layer in Arabidopsis seeds, expression of endosperm-specific genes contributes to global expression of imbibed whole seeds.

To further compare these transcriptomes, 15 331 differentially expressed AGI code genes (P initial value $<10^{-8}$; false discovery rate (FDR) $\alpha = 0.05$; Table S2) in at least one condition were analyzed using a principal components analysis (PCA). Dry seeds of the wild type, the *aba2* mutant and the *cyp707a1a2a3* triple mutant were located at similar positions, suggesting that these seeds are similar at the transcriptome level compared with imbibed seeds (Figure 3b). Interestingly, the position on the first PCA axis of 24-h imbibed seeds of the *cyp707a1a2a3* triple mutant was

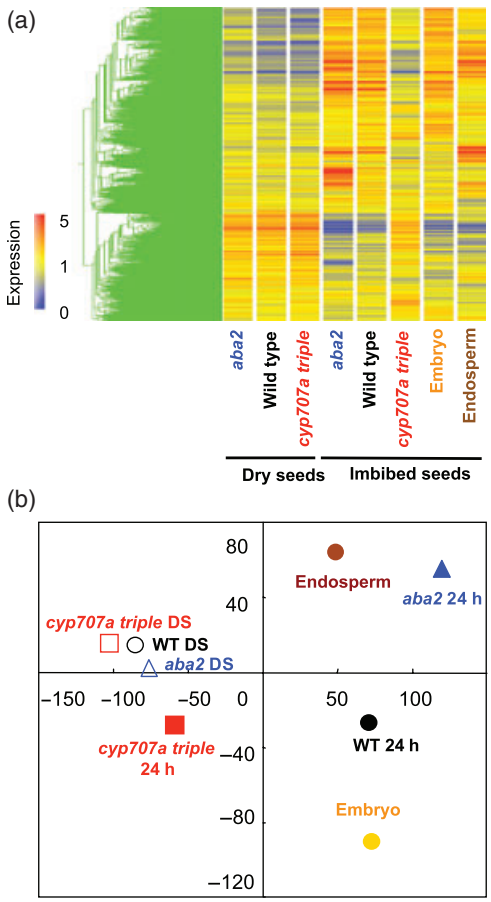

Figure 3. Global gene expression profiling of Arabidopsis Genome Initiative (AGI) genes in dry seeds and imbibed seeds.
(a) Hierarchical clustering analysis of 18 848 expressed AGI genes in dry and imbibed seeds with embryo and endosperm tissues. Of all the AGI genes (32 144), 18 849 were expressed in seeds. Signal intensity values were obtained using the ARTADE program (*P* initial value $<10^{-8}$) from three independent biological replicates. Colored bars indicate relative expression levels. Embryo and endosperm tissues were isolated from 24-h imbibed seeds.
(b) Principal components analysis (PCA) of genes expressed in dry seeds, imbibed seeds, embryo and endosperm tissues. DS and 24 h indicate dry seeds and 24-h imbibed seeds, respectively. Variance in PC1 and PC2 was 31 and 17%, respectively.

located close to that of dry seeds, whereas the positions of 24-h imbibed wild-type and *aba2* mutant seeds were located distantly from that of dry seeds. From these observations, the first dimension (PC1) is likely to reflect dormancy or germination states. Indeed, *LEA* and *RD29B* genes were observed in low scores on the first PCA axis (Table S2). These genes are expressed at high levels in dry or dormant seeds (Bassel *et al.*, 2008; Carrera *et al.*, 2008). Conversely, several 40S and 60S ribosomal protein genes, which are indicators of growth status, were observed in high scores on the first PCA axis (Tatematsu *et al.*, 2008; Table S2). The second dimension (PC2) is likely to reflect the differences between embryo and endosperm. As reported previously, genes related to protein metabolism were observed in high scores (Penfield *et al.*, 2006; Table S2). In contrast, a large number of pentatricopeptide repeat (PPR) proteins and RNA metabolism-related genes were observed in low scores in PC2 (Table S2). The PPR proteins have a range of essential functions in post-transcriptional processing of RNA in mitochondria and chloroplasts (Schmitz-Linneweber and Small, 2008). A subset of mutants of PPR genes show embryo-defective phenotypes or severely reduced germination, indicating that these proteins are required for normal embryo growth (Lurin *et al.*, 2004; de Longevialle *et al.*, 2007).

Gene expression analysis in *aba2* and *cyp707a1a2a3* triple mutants

As shown in Figure 2, expression of ABA-responsive and signaling-related genes is associated with endogenous ABA levels in the *aba2* mutant and the *cyp707a1a2a3* triple mutant. To identify ABA-responsive AGI code genes, significant differences were judged by the Mann–Whitney *U*-test (FDR $\alpha = 0.05$; Storey, 2002; Storey and Tibshirani, 2003). In dry seeds, there were few genes showing an opposite expression pattern between the *aba2* mutant and *cyp707a1a2a3* triple mutant (i.e. an expression ratio for *aba2*/wild type >2.0 and *cyp707a1a2a3*/wild type <2.0; Figure S2 and Table S3). This result suggests that a large number of stored mRNAs in dry seeds might be regulated by independent mechanisms in these mutants. It was also reported that the global expression patterns were similar among dry seeds of the wild type and two ABA-insensitive mutants, *abi4* and *abi5* (Nakabayashi *et al.*, 2005). However, severely ABA-insensitive mutants, such as *abi3-6* or *snrk2d/e/i* triple mutants, cannot enter seed maturation processes, and numerous genes in these dry seeds showed different transcript levels from those in the wild-type seeds (Nakashima *et al.*, 2009). From these results, a large number of stored mRNAs might be predominantly regulated by the threshold of ABA sensitivity concurrent with the developmental signal rather than endogenous ABA levels, because the seeds of the ABA metabolism-related mutants, *aba2* and *cyp707a1a2a3*, can progress towards seed maturation in a similar manner as wild-type seeds. However, it is worth mentioning that many photosynthesis- and cell wall-related genes were observed among the upregulated genes of the *aba2* dry seeds (Table S3). These genes are known as feature genes that are expressed during germination. This observation indicates that endogenous ABA levels are required for repressing the expression of these genes during the seed maturation process. In contrast to dry seeds, at 24 h after seed imbibition we identified 336 ABA-upregulated genes, including ABA signaling factors, *ABI5*, *AHG1*, *ATAF1* and *EEL*, and a regulator of ABA and GA metabolism, *SOMNUS* (Figure 4a and Table S3). As expected, CACGTG-related sequences, one of the most typical ABA-responsive

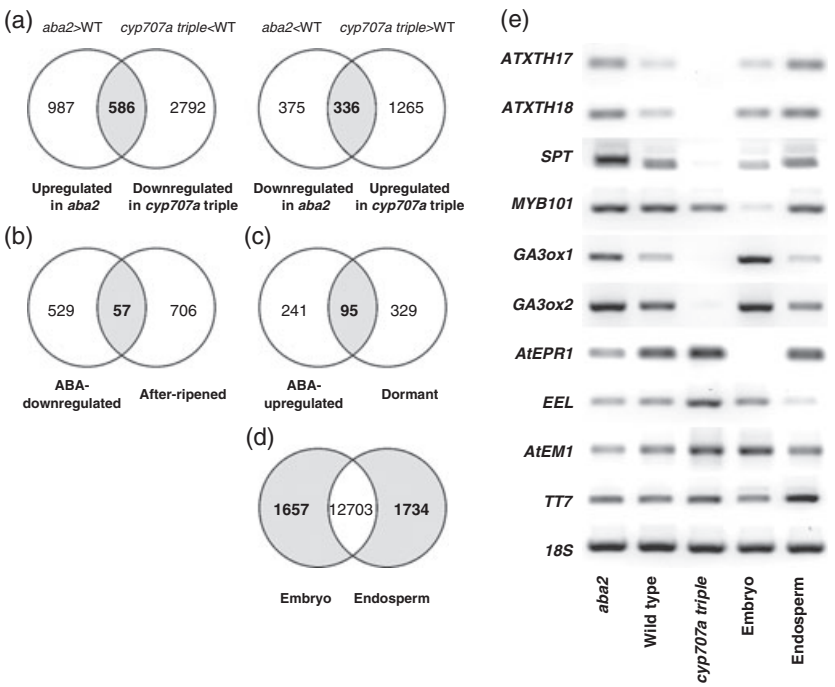

Figure 4. ABA-responsive genes and genes differentially expressed in embryo and endosperm.
(a) ABA-responsive genes in *aba2* mutant and *cyp707a1a2a3* triple mutant at 24 h after seed imbibition.
(b) Comparison of ABA-downregulated genes and after-ripened genes.
(c) Comparison of ABA-upregulated genes and dormant genes.
(d) Comparison of differential gene expression in embryo and endosperm at 24 h after seed imbibition. Genes differentially expressed between (a) and (d) judged by Mann–Whitney *U*-test (false discovery rate (FDR) $\alpha = 0.05$) were further selected using an expression ratio cut-off of twofold higher or lower.
(e) Semi-quantitative RT-PCR for ABA-responsive genes and differentially regulated genes in embryo and endosperm.

elements, were frequently observed in the 0.5-kb upstream regions of 335 ABA-upregulated AGI genes (Table S4; Shen and Ho, 1995). We also identified 586 ABA-downregulated AGI code genes at 24 h after seed imbibition, including the GA metabolism-related genes *GA3ox1* and *GA3ox2* and their regulator *SPATULA* (Figure 4a and Table S3). Indeed, expression of *GA3ox1* and *GA3ox2* is repressed by elevated endogenous ABA in unfavorable conditions (Seo *et al.*, 2006; Toh *et al.*, 2008).

To understand how ABA-responsive genes affect biological processes in seeds, we used TAGGIT ontology, which is a seed biology-related gene ontology (GO) annotation (Carrera *et al.*, 2007). TAGGIT ontology analysis revealed that seed storage proteins and dormancy-related genes were only observed among the ABA-upregulated genes (Table S5). Genes associated with photosynthesis, glycolysis and gluconeogenesis, cell wall modification, the cell cycle and the Krebs cycle were more commonly found among the ABA-downregulated genes than among ABA-upregulated genes (Table S5). From these observations, cohorts of ABA-upregulated and -downregulated genes possibly reflected those of dormant and after-ripened states, respectively (Carrera *et al.*, 2007). Therefore, our ABA-responsive gene list was compared with dormancy or after-ripened genes as reported previously (Cadman *et al.*, 2006). Among 586 ABA-downregulated genes, 57 genes, including *GA3ox1* and *SPATULA*, were overlapped with after-ripened genes, whereas *GA3ox2*, an ABA-downregulated gene, was not among the after-ripened genes (Figure 4b and Table S6). On the other hand, of 336 ABA-upregulated genes, 95 genes including *EEL*, *ATAF1*, *RD29B*, *LEA*s, *AtEM1* and *AtEM6*, overlapped with dormancy genes but not after-ripened genes (Figure 4c and Table S6). Interestingly, two major ABA signaling factors in Arabidopsis seeds, *ABI5* and *AHG1*, were not included among the dormancy genes, suggesting that seed dormancy might be regulated by a different pathway of ABA signal transduction. It is reported that the *abi5* mutant exhibits a similar degree of seed dormancy as the wild type (Finkelstein, 1994). In contrast, the *abi5* mutant can germinate in the presence of a GA biosynthesis inhibitor (Piskurewicz *et al.*, 2008). In the case of secondary seed dormancy at high temperatures, only a subset of ABA signaling factors is involved in thermoinhibition of Arabidopsis seed germination (Tamura *et al.*, 2006). Consistent with our speculation, these observations imply that seed dormancy is controlled by various ABA signaling factors that respond to multiple external cues.

We have identified many ABA-responsive genes using ABA-deficient and ABA-overaccumulation mutants. However, we also cannot exclude the possibility that part of these genes is regulated by developmental transition, because endogenous ABA levels are generally associated with seed dormancy and germination states. Therefore, it appears that ABA-responsive genes are composed of ones directly regulated by ABA and ones regulated by both ABA and developmental transition.

Embryo- and endosperm-specific expression of regulators for seed dormancy and germination

Tiling array analyses using dissected embryo and endosperm tissues revealed that 16 094 AGI code genes were expressed in these tissues at 24 h after seed imbibition

(*P* initial value $<10^{-8}$; Table S1). *ABI4* was predominantly expressed in the embryo (for embryo, 10 463; for endosperm, 1526; *P*-value <0.0001), whereas *AtEPR1* was almost exclusively expressed in the endosperm (for embryo, 918; for endosperm, 86 944; *P*-value <0.0001). This result indicates that there was no significant cross-contamination between the two fractions. The Mann–Whitney *U*-test (FDR $\alpha = 0.05$; signal value cut-off twofold difference) revealed that 1657 genes, including embryo-specific markers, *PDF1*, *ABI4*, *GA3ox1* and *GA3ox2*, were predominantly expressed in embryo tissues (Figure 4d and Table S7). It is worth mentioning that the ABA-upregulated genes, *EEL* and *AtEM1*, and the ABA-downregulated genes, *GA3ox1* and *GA3ox2*, were predominantly expressed in the embryo, suggesting that endogenous ABA regulates their transcription in specific tissues (Figure 4e). TAGGIT ontology analysis revealed that more genes related to DNA repair/RNA metabolism, respiration and photosynthesis were expressed in the embryo than in the endosperm (Table S5). Notably, a subset of mutants in mRNA metabolism-related genes show embryo-defective or seedling-lethal phenotypes (SeedGenes Project, http://www.seedgenes.org/). Interestingly, seeds of a mutant lacking mRNA processing-related components exhibit an ABA-hypersensitive phenotype (Finkelstein *et al.*, 2008). A large number of mRNA species change drastically during germination in association with ABA levels (Ogawa *et al.*, 2003; Nakabayashi *et al.*, 2005). Therefore, ABA is likely to be involved in tissue-specific expression patterns of these genes.

On the other hand, we identified 1734 genes including endosperm-specific markers, *AtEPR1* and *MYB10*, that were expressed in the endosperm rather than the embryo (Figure 4d,e and Table S7). *AtEPR1* is expressed in the endosperm during germination (Dubreucq *et al.*, 2000). Our study revealed that this gene was highly expressed even in *cyp707a1a2a3* triple mutants after seed imbibition. It is possible that AtEPR1 affects the composition of the cell wall in the endosperm and contributes to maintenance of dormant seeds. Indeed, *AtEPR1* expression is observed from late-maturation to mature dry seeds, and its expression levels are maintained in dormant seeds rather than non-dormant seeds (Schmid *et al.*, 2005; Cadman *et al.*, 2006; Bassel *et al.*, 2008). Among ABA-downregulated genes, expression of *SPATULA* and the cell wall-related genes *ATXTH17* and *ATXTH18* was higher in the endosperm than in the embryo (Figure 4e). These genes are probably involved in endosperm weakening during germination. TAGGIT ontology analysis revealed that the Krebs cycle, β-oxidation and stress-related genes were more abundant in the endosperm than in the embryo (Table S5). Mobilization of endosperm lipid reserve by the Krebs cycle and β-oxidation is important for embryo growth during germination (Penfield *et al.*, 2004, 2006). Stress-related genes included the flavonoid biosynthetic enzymes *TT6* and *TT7* (Table S5). Most *tt* mutants show reduced dormancy, and *tt*7 mutants were also isolated as resistant to high temperature and germinating at cold temperatures, indicating that flavonoid biosynthesis contributes to the regulation of germination (Debeaujon and Koornneef, 2000; Salaita *et al.*, 2005; Tamura *et al.*, 2006).

Penfield *et al.* (2006) has reported transcriptome analysis of embryo and endosperm in Arabidopsis. Therefore, a list of the embryo- and endosperm-specific genes identified in this study was compared with that of Penfield *et al.* (2006). About 40% of our embryo- and endosperm-specific genes overlapped with the previous data set (Figure S3 and Table S8). This low percentage of overlap might be due to differences in the preparation of the seed materials, i.e. in this study, the data set was obtained from 24-h imbibed seeds without stratification, whereas the Penfield data set was obtained from 24-h imbibed seeds after 3 days of stratification.

Identification and characterization of novel transcriptional units in Arabidopsis seeds

Next, we focused on the novel transcriptional units (TUs; i.e. non-AGI TUs) in the TAIR8 gene model, because the tiling array can detect all transcripts from the entire genome. We identified 6105 non-redundant groups of non-AGI TUs (Table S9) as the expressed ones using the ARTADE method (*P* initial value $<10^{-8}$) (Toyoda and Shinozaki, 2005). These non-AGI TUs were located across the whole chromosome similarly to AGI code genes (Figure S4). Among non-AGI TUs, 706 (11.6%) and 3333 (55.6%) were overlapped with Arabidopsis community full-length cDNAs, including RIKEN full-length cDNAs and reliable signatures of massively parallel signature sequencing (MPSS) tag sequences, respectively (Figure S4 and Table S9; Seki *et al.*, 2002; Meyers *et al.*, 2004). Furthermore, 2517 (41.2%) non-AGI TUs were novel, and 3588 (58.8%) non-AGI TUs were previously identified as the non-AGI TUs expressed in Arabidopsis seedlings under stress or exogenous ABA treatment (Figure 5a; Matsui *et al.*, 2008). Interestingly, 5900 (97%) non-AGI TUs did not have sequence similarity with any proteins in the National Institutes of Health data set (Table S9). This result suggests that these non-AGI TUs probably encode hypothetical non-protein-coding RNA. In addition, 5256 (86%) and 849 (14%) of the non-AGI TUs were defined as antisense and intergenic types, respectively (Figure 5b,d,e and Table S9). In the case of protein-coding non-AGI TUs, approximately 50% of TUs were expressed in intergenic regions, and their sequence similarities were moderately conserved in the expressed sequence tag (EST) data sets from other plant species in the NCBI Unigene database (Figure 5b,c, Tables S9 and S10). In contrast, non-protein-coding non-AGI TUs in intergenic regions were less

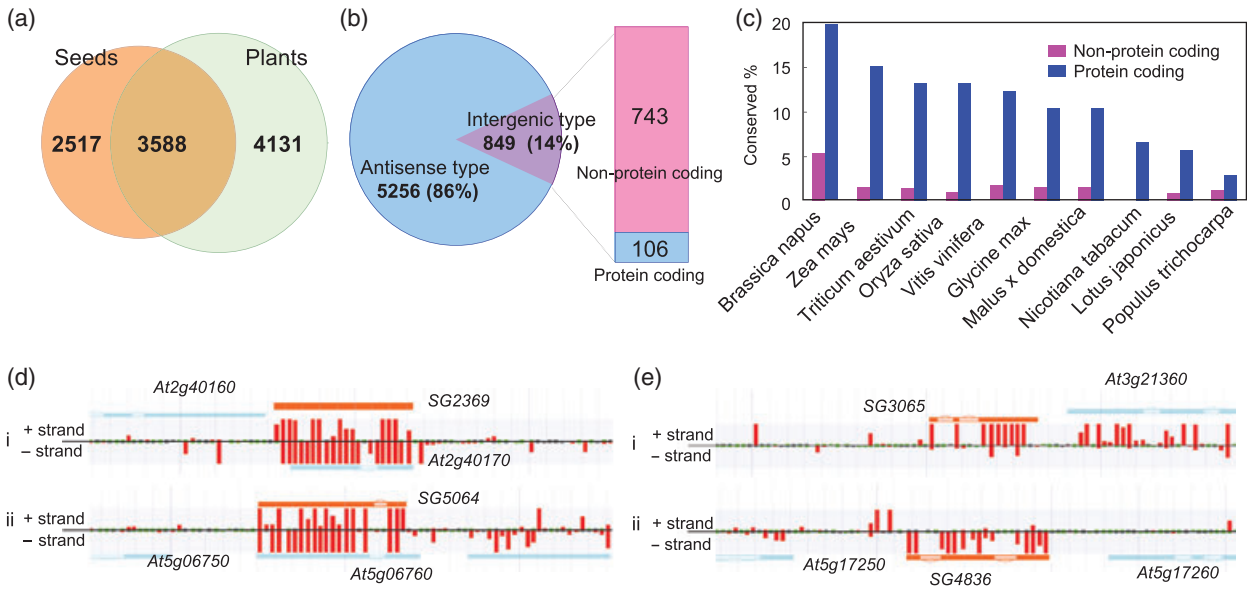

Figure 5. Overview of novel [non-Arabidopsis Genome Initiative (AGI)] transcriptional units (TUs) in Arabidopsis seeds.
(a) Comparison of non-AGI TUs between plants and seeds. Of 6105 non-AGI TUs, 2517 were novel, and 3588 have been detected in plants previously (Matsui *et al.*, 2008).
(b) Types of non-AGI TUs. Among all non-AGI TUs, 5256 (86%) TUs were antisense and 849 (14%) were intergenic. In intergenic regions, 743 non-AGI TUs were classified as non-protein coding and 106 as protein coding. Typical antisense and intergenic types are shown in (d) and (e), respectively.
(c) Evolutionary conservation of non-AGI TUs in intergenic regions. Conserved non-AGI TUs in intergenic regions were searched using BlastN (*E* value $<10^{-10}$) against expressed sequence tag (EST) data sets from 10 plant species.
(d) Example of antisense non-AGI TUs. Expression profiles of *SG2369* in dry seeds and *SG5064* in imbibed seeds are shown in (i) and (ii), respectively.
(e) Example of intergenic non-AGI TUs. Expression profiles of *SG3065* in dry seeds and *SG4836* in imbibed seeds are shown in (i) and (ii), respectively.
In (d) and (e) orange and blue horizontal bars indicate non-AGI TU and AGI code genes, respectively. Red and green bars indicate the signal intensity of probes (red > 400, green < 400). Tiling array data are available at OmicBrowse (http://omicspace.riken.jp/gps/group/psca5).

conserved than protein-coding TUs (Figure 5c and Table S10), suggesting that non-protein-coding TUs have evolved more rapidly than protein-coding TUs.

Expression of these non-AGI TUs was also analyzed by using strand-specific RNA probes. The expression of non-AGI TUs was observed in dry seeds and/or imbibed seeds (Figure 6a,b). These results indicate that non-AGI TUs are not artifacts in tiling array experiments and are expressed in Arabidopsis seeds. In northern analyses, the sense and antisense transcripts showed similar expression patterns, except for *At1g05190*/*SG0123* (Figure 6a). There was a linear correlation between sense and antisense transcripts of 5705 AGI/non-AGI TU pairs ($r = 0.64$) but not between those of 1435 AGI/AGI pairs ($r = 6.3 \times 10^{-4}$) (Figure S5 and Table S11). We also detected antisense transcripts for many major ABA-related genes including *AHG1*, *ABI5*, *RD29B* and *AtEM6*, which were downregulated after seed imbibition similarly to sense transcripts (Table S11). A linear correlation between expression of AGI/non-AGI TUs was also observed in stress responses (Matsui *et al.*, 2008). Recently, Hazen *et al.* (2009) reported that the expression patterns of some antisense transcripts differed from those of sense transcripts in circadian clock regulation. These observations suggest that antisense transcripts might be regulated dynamically in response to periodic changes in the environment. Interestingly, there was also a linear correlation ($r = 0.61$) between the 123 non-AGI/non-AGI TU pairs, as reported previously (Figure S5; Matsui *et al.*, 2008). Indeed, expression patterns of the pairs *SG1645*/*SG1646* and *SG3258*/*SG3259* were similar during seed imbibition (Figure 6b).

Expression of non-AGI TUs in seeds

We performed a hierarchal clustering analysis on 6105 non-AGI TUs (Figure 7a). Expression patterns of non-AGI TUs in dry seeds were similar among the wild type, the *aba2* mutant and the *cyp707a1a2a3* triple mutant (Figure 7a). However, after seed imbibition there were significant differences in expression patterns of many non-AGI TUs among the three types. Moreover, expression patterns of non-AGI TUs in embryos differed from those in endosperm tissues. In fact, ABA-responsive non-AGI TUs in dry seeds were barely identified by the Mann–Whitney *U*-test (Figure S6 and Table S12). On the other hand, we identified 42 ABA-upregulated and three ABA-downregulated non-AGI TUs at 24 h after seed imbibition (Figure 7b and Table S12), whereas expression of 61 and 775 non-AGI TUs was identified as embryo- and endosperm- specific in 24-h imbibed

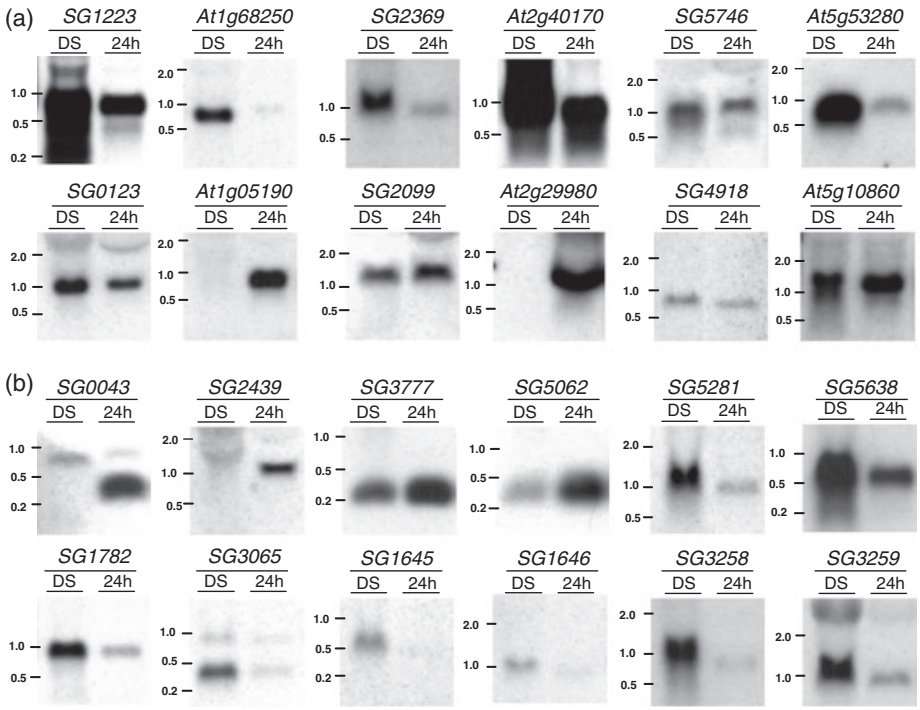

Figure 6. Northern blot analysis of the non-Arabidopsis Genome Initiative (AGI) transcriptional units (TUs) in dry seeds and imbibed seeds.
(a) Northern blot analysis of antisense non-AGI TUs. Pairs of sense/antisense transcripts are as follows: *At1g68250*/*SG1223*, *At2g40170*/*SG2369*, *At5g53280*/*SG5746*, *At1g05190*/*SG0123*, *At2g29980*/*SG2099* and *At5g10860*/*SG4918*.
(b) Northern blot analysis of intergenic-type non-AGI TUs. For northern blot analysis, 40 μg total RNA from dry seeds and imbibed seeds at 24 h was loaded into each lane and hybridized with strand-specific RNA probes.

seeds, respectively (Figure 7c and Table S13). A RT-PCR expression analysis was performed to confirm ABA-responsive and tissue-specific expression of several non-AGI TUs in intergenic regions. In agreement with tiling array data, after seed imbibition, expression of ABA-responsive non-AGI TUs was associated with endogenous ABA levels, and expression of tissue-specific non-AGI TUs depended on the tissues used (Figure 7d,e and Tables S12 and S13). Notably, ABA-upregulated *SG1158* and *SG3065* and ABA-downregulated *SG3872* were predominantly expressed in the endosperm (Figure 7d,e). These results indicate that expression of stored non-AGI TUs in dry seeds is hardly influenced by endogenous ABA levels, and that the expression patterns of ABA-responsive and tissue-specific non-AGI TUs after seed imbibition are similar to those of AGI genes. However, significantly fewer ABA-responsive non-AGI TUs were expressed compared with ABA-responsive AGI code genes in 24-h imbibed seeds. These results imply that the changes in expression patterns of non-AGI TUs are less dynamic than those of AGI code genes after seed imbibition (Figure S5).

Recent transcriptome studies have shown that many mRNAs resembling non-protein-coding RNA are present in addition to microRNAs and small interfering RNAs in Arabidopsis and rice (*Oryza sativa*) (Li *et al.*, 2006; Zhang *et al.*, 2006; Matsui *et al.*, 2008). We also have found more than 5900 non-protein-coding RNAs in Arabidopsis seeds in this study. Although the functions of non-protein-coding RNA, including natural antisense RNA or pseudo genes, are largely unknown, it is thought that non-protein-coding RNA has a broad range of functions including roles in gene silencing, epigenetic control and ribozymes in eukaryotes (Amaral *et al.*, 2008). Recent studies showed that several non-protein-coding RNAs participate in stress responses and organ development in plants (Dai *et al.*, 2007; Franco-Zorrilla *et al.*, 2007; Ben Amor *et al.*, 2009). It appears that identification of conserved or unique non-protein-coding RNA is one way to find functional non-protein-coding RNA in Arabidopsis seeds (Rymarquis *et al.*, 2008).

Conclusions and future perspectives

In this study, we have demonstrated that the effects of exogenous ABA on ABA-mediated transcription at early stages of seed imbibition differ from those of endogenous ABA. The effects of exogenous ABA were prominent in the expression of several ABA-related genes at a later stage of imbibition. We performed large-scale expression studies on Arabidopsis seeds using tiling arrays. A large number of stored mRNA species in dry seeds were not regulated by endogenous ABA levels, suggesting that

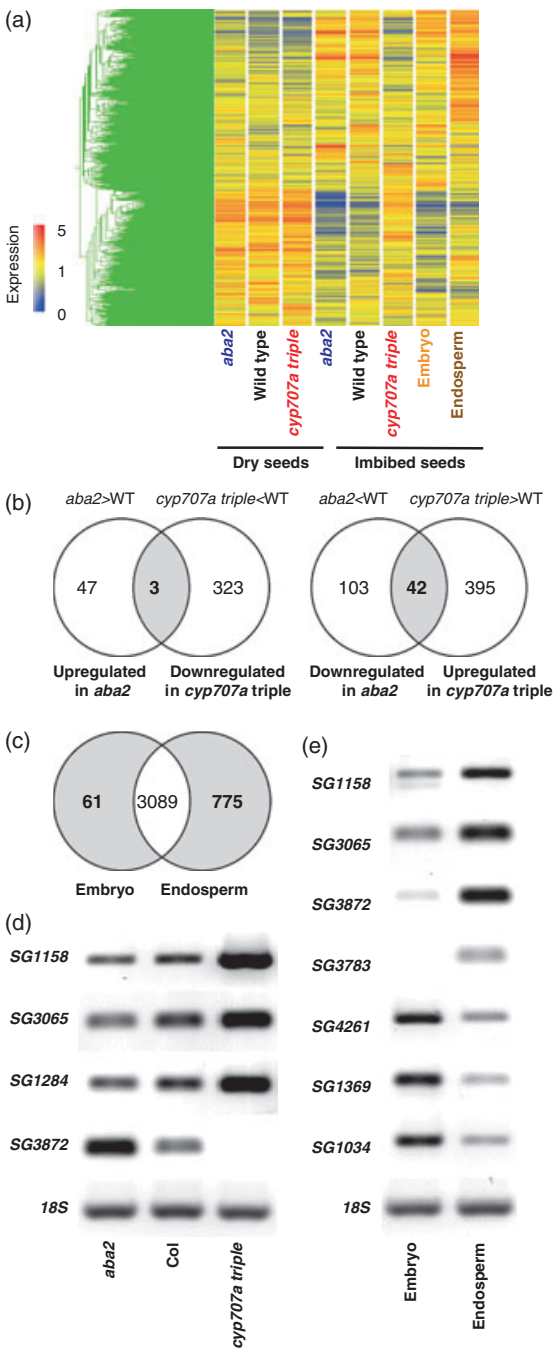

Figure 7. Global expression profiling of non-Arabidopsis Genome Initiative (AGI) transcriptional units (TUs) and differentially regulated non-AGI TUs in Arabidopsis seeds.
(a) Hierarchical clustering analysis of 6105 non-AGI TUs in Arabidopsis seeds. Colored bars indicate relative expression levels.
(b) ABA-responsive non-AGI TUs at 24 h after seed imbibition.
(c) Differential expression of non-AGI TUs in embryo and endosperm at 24 h after seed imbibition. Genes differentially expressed between (b) and (c) judged by Mann–Whitney *U*-test (false discovery rate (FDR) $\alpha = 0.05$) were further selected using an expression ratio cut-off of twofold higher or lower.
(d) Semi-quantitative RT-PCR for ABA-responsive genes at 24 h after seed imbibition.
(e) Semi-quantitative RT-PCR for differentially regulated genes in embryo and endosperm at 24 h after seed imbibition.

these stored mRNAs might not influence seed dormancy and germination. After seed imbibition, endogenous ABA affected the expression of critical components, e.g. ABA signaling, photosynthesis, physiological and metabolic genes including a GA biosynthesis enzyme. Some of these genes were differentially regulated in the different seed tissues, suggesting that they determine whether seeds remain dormant or germinate after imbibition. In addition, our tiling array analysis identified a large number of non-AGI TUs including non-protein-coding RNA. Numerous non-protein-coding RNAs were defined as antisense transcripts, and sense/antisense pairs showed similar expression patterns in dry and imbibed seeds. Like AGI code genes, ABA-responsive and tissue-specific non-AGI TUs were regulated in imbibed seeds. Here, we show that dynamic and complex transcriptional regulation occurs in seeds. Our large-scale data set will contribute to future TAIR gene models and will increase our understanding of the molecular basis of how endogenous ABA controls seed dormancy and germination.

EXPERIMENTAL PROCEDURES

Plant materials and growth conditions

In this study, both the wild type and mutants were *A. thaliana* accession Columbia. The *aba2-1* mutant was isolated previously (LeonKloosterziel *et al.*, 1996). The *cyp707a1a2a3* triple mutant was isolated by crossing the *cyp707a1-1 cyp707a2-1* double mutant and the *cyp707a2-1 cyp707a3-1* double mutant (Okamoto *et al.*, 2006). Plants were grown in a growth chamber at 22°C and 50–60% relative humidity under a 16-h light/8-h dark cycle. Seeds were harvested from yellow-brown siliques and were immediately used in this study. For germination tests, freshly harvested seeds were sown on a 0.5% agarose gel (LO3; TAKARA, http://www.takara-bio.com/) and the plates were kept at 22°C under continuous light conditions. To obtain seed samples for ABA measurement and RNA extraction, 30–60 mg seeds were imbibed in 8.5 cm Petri dishes containing two layers of filter paper (approximately 7 cm diameter) and 2 ml water and harvested. Dissection of embryo and testa/endosperm from 24-h imbibed seeds was carried out under a stereoscopic microscope.

ABA determinations

Extraction, purification and quantification of ABA were carried out as described in Saika *et al.* (2007).

RNA isolation

For RT-PCR and tiling array analyses, total RNA was isolated as described in Kushiro *et al.* (2004). For northern blot analysis and RNA isolation from the dissected embryo and endosperm tissues, extraction was carried out as described previously (Martin *et al.*, 2005).

RT-PCR analysis

First-strand cDNA synthesis and quantitative RT-PCR (qRT-PCR) using SYBR Green I were performed as described previously (Okamoto *et al.*, 2006). Sequences of primers used for RT-PCR are shown in Table S14.

Whole-genome tiling array analysis

The GeneChip Arabidopsis tiling array set (1.0 F array and 1.0 R array, Affymetrix) was used in this study (Zhang *et al.*, 2006; Matsui *et al.*, 2008). The tiling array analysis was carried out as described previously (Matsui *et al.*, 2008).

Whole-genome tiling array data analysis

Arabidopsis genome sequence and annotation information from TAIR 8 was mapped to the probes of the Affymetrix Arabidopsis whole-genome tiling array. The tiling array data analysis was carried out essentially as described previously (Matsui *et al.*, 2008). The expressed AGI code genes and non-AGI TUs were detected using the ARTADE program (*P* initial value $<10^{-8}$) (Toyoda and Shinozaki, 2005). To detect non-protein-coding RNAs, we carried out homology searches of non-AGI TUs against registered protein sequence data sets (NIH NR database) using the BLASTX program (Matsui *et al.*, 2008). Comparisons of non-AGI TUs with cDNAs and MPSS tags and identification of sense/antisense transcript pairs were also carried out as described previously (Matsui *et al.*, 2008). In analysis of evolutionary conservation of non-AGI TUs in intergenic regions, homology searches of the non-AGI TU against the EST sequences of 10 plant species (NCBI Unigene database, http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene) was carried out using the BLASTN program (*E* value $<10^{-10}$). To conduct hierarchical clustering analysis for the AGI code genes and non-AGI TUs, tiling array data (signal intensity, physical position and *P* initial value) were entered into GENESPRING 7.3. In the PCA, 15 331 genes that were significantly differentially expressed in at least one condition (FDR <0.05, *P* initial value $<10^{-8}$) were subjected to *Z*-scaling using genefilter, prcomp, princomp and the *R* statistical analysis packages (Ihaka and Gentleman, 1996). Genes were ranked in order of their PC1 and PC2 scores (Table S2). To identify ABA-responsive and differentially regulated AGI code genes and non-AGI TUs among samples, significant differences were judged by the Mann–Whitney *U*-test (FDR $\alpha = 0.05$) using the all probes (5.8 million perfect match and 5.8 million mismatch probes) as described previously (Storey, 2002; Storey and Tibshirani, 2003; Matsui *et al.*, 2008). Regulatory *cis* elements for differentially regulated AGI code genes and non-AGI code TUs were identified using GENESPRING 7.3 software. Developmental signatures associated with differentially regulated genes were defined using the TAGGIT workflow list (Carrera *et al.*, 2007). Arabidopsis tiling array data used in this study is available at GEO (http://www.ncbi.nlm.nih.gov/geo/info/linking.html) under the accession number GSE15700.

Northern analysis

Northern analysis was performed using a DIG Northern Starter Kit (Roche Applied Science, http://www.roche-applied-science.com/). Complementary DNA fragments for non-AGI TUs and AGI code genes were subcloned into pSTBlue-1 vector (Novagen, http://www.merck.de/en/index.html) and were amplified by PCR using forward and reverse M13 primers. Sequences of cDNA fragments for non-AGI TUs and AGI code genes are shown in Table S14. The PCR products were used as templates for synthesis of digoxigenin (DIG)-labeled antisense RNA probes with T7 or SP6 RNA polymerase. Total RNA (40 μg) was separated by electrophoresis on a 1.2% agarose gel containing 2.2 M formaldehyde. After gel blotting, the membranes (Hybond N; Amersham Biosciences, http://www.gelifesciences.co.jp) were hybridized with the RNA probe using DIG Easy Hyb Granules (Roche Applied Science). The membranes were washed using the DIG Wash and Block Buffer Set (Roche Applied Science). The hybridized probes were immunologically detected with anti-DIG-AP and visualized with the chemiluminescence substrate CDP-*Star* (Roche Applied Science). Chemiluminescence was detected using Hyperfilm ECL (Amersham Biosciences).

ACKNOWLEDGEMENTS

We thank Dr Ian Graham for providing microarray data, Ms Sachiyo Harada for technical assistance and the Arabidopsis Biological Resource Center (ABRC) for providing T-DNA-tagged lines. This work was supported by the Special Postdoctoral Researcher's Program from RIKEN (to MO) and a grant from RIKEN Plant Science Center (to MS).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article or our website (http://pfgweb.psc.riken.jp/download.html):

Figure S1. Phenotypic analysis of *cyp707a1a2a3* triple mutant and Cape Verde Islands (Cvi) accession seeds.

Figure S2. ABA-responsive genes in dry seeds.

Figure S3. Comparison of embryo- and endosperm-specific genes between this study and the Penfield *et al.* (2006) data set.

Figure S4. Overview of non-Arabidopsis Genome Initiative (AGI) transcriptional units (TUs) in Arabidopsis seeds.

Figure S5. Correlations between expression ratios in Arabidopsis Genome Initiative (AGI)/AGI, AGI/non-AGI transcriptional unit (TU) and non-AGI TU/non-AGI TU gene pairs.

Figure S6. ABA-responsive non-Arabidopsis Genome Initiative (AGI) transcriptional units (TUs) in dry seeds.

Table S1. Arabidopsis Genome Initiative (AGI) code gene expression in this study.

Table S2. Principal components analysis (PCA) of 18 884 Arabidopsis Genome Initiative (AGI) code genes.

Table S3. A list of ABA-responsive genes at 24 h after seed imbibition.

Table S4. Regulatory *cis* elements of ABA-upregulated Arabidopsis Genome Initiative (AGI) code genes in 24-h imbibed seeds.

Table S5. TAGGIT ontology analysis of ABA-responsive and differentially regulated Arabidopsis Genome Initiative (AGI) code genes.

Table S6. Comparison between ABA-responsive genes at 24 h after seed imbibition and dormant or after-ripened genes.

Table S7. Differential gene expression in embryo and endosperm.

Table S8. A list of embryo- and endosperm-specific genes that were identified in both our and Penfield *et al.*'s analyses.

Table S9. Non-Arabidopsis Genome Initiative (AGI) transcriptional unit (TU) expression data in this study.

Table S10. Comparison between non-Arabidopsis Genome Initiative (AGI) transcriptional unit (TU) sequences and expressed sequence tag (EST) sequences from several plants species.

Table S11. Arabidopsis Genome Initiative (AGI)/AGI, AGI/non-AGI transcriptional unit (TU) and non-AGI TU/non-AGI TU pairs.

Table S12. ABA-responsive non-Arabidopsis Genome Initiative (AGI) transcriptional units (TUs) in dry seeds and imbibed seeds.

Table S13. Non-Arabidopsis Genome Initiative (AGI) transcriptional units (TUs) differentially expressed in embryo and endosperm tissues.

Table S14. Sequences used for quantitative RT-PCR (qRT-PCR) and semiquantitative RT-PCR (SQRT-PCR), and cDNA template sequences used for northern analyses.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online

delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

REFERENCES

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

Amaral, P.P., Dinger, M.E., Mercer, T.R. and Mattick, J.S. (2008) The eukaryotic genome as an RNA machine. *Science*, **319**, 1787–1789.

Baskin, C.C. and Baskin, J.M. (1998) *Seeds: Ecology, Biogeography, and Evolution of Dormancy and Germination*. San Diego: Academic Press.

Bassel, G.W., Fung, P., Chow, T.F.F., Foong, J.A., Provart, N.J. and Cutler, S.R. (2008) Elucidating the germination transcriptional program using small molecules. *Plant Physiol.* **147**, 143–155.

Ben Amor, B., Wirth, S., Merchan, F. *et al.* (2009) Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. *Genome Res.* **19**, 57–69.

Bewley, J.D. (1997) Seed germination and dormancy. *Plant Cell*, **9**, 1055–1066.

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

Carrera, E., Holman, T., Medhurst, A., Peer, W., Schmuths, H., Footitt, S., Theodoulou, F.L. and Holdsworth, M.J. (2007) Gene expression profiling reveals defined functions of the ATP-binding cassette transporter COMATOSE late in phase II of germination. *Plant Physiol.* **143**, 1669–1679.

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

Chen, F. and Bradford, K.J. (2000) Expression of an expansin is associated with endosperm weakening during tomato seed germination. *Plant Physiol.* **124**, 1265–1274.

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

Dai, X.Y., Yu, J.J., Ma, J.X., Ao, G.M. and Zhao, Q. (2007) Overexpression of Zm401, an mRNA-like RNA, has distinct effects on pollen development in maize. *Plant Growth Regul.* **52**, 229–239.

Debeaujon, I. and Koornneef, M. (2000) Gibberellin requirement for Arabidopsis seed germination is determined both by testa characteristics and embryonic abscisic acid. *Plant Physiol.* **122**, 415–424.

Dubreucq, B., Berger, N., Vincent, E., Boisson, M., Pelletier, G., Caboche, M. and Lepiniec, L. (2000) The *Arabidopsis AtEPR1* extensin-like gene is specifically expressed in endosperm during seed germination. *Plant J.* **23**, 643–652.

Finkelstein, R.R. (1994) Mutations at 2 new Arabidopsis Aba response loci are similar to the Abi3 mutations. *Plant J.* **5**, 765–771.

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

Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J.A. and Paz-Ares, J. (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet.* **39**, 1033–1037.

González-Guzmán, M., Abia, D., Salinas, J., Serrano, R. and Rodríguez, P.L. (2004) Two new alleles of the abscisic aldehyde oxidase 3 gene reveal its role in abscisic acid biosynthesis in seeds. *Plant Physiol.* **135**, 325–333.

Grappin, P., Bouinot, D., Sotta, B., Miginiac, E. and Jullien, M. (2000) Control of seed dormancy in *Nicotiana plumbaginifolia*: post-imbibition abscisic acid synthesis imposes dormancy maintenance. *Planta*, **210**, 279–285.

Groot, S.P.C. and Karssen, C.M. (1992) Dormancy and germination of abscisic acid-deficient tomato seeds - studies with the *sitiens* mutant. *Plant Physiol.* **99**, 952–958.

Halmer, P., Bewley, J.D. and Thorpe, T.A. (1976) An enzyme to degrade lettuce endosperm cell walls. Appearance of a mannanase following phytochrome- and gibberellin-induced germination. *Planta*, **130**, 189–196.

Hazen, S.P., Naef, F., Quisel, T., Gendron, J.M., Chen, H.M., Ecker, J.R., Borevitz, J.O. and Kay, S.A. (2009) Exploring the transcriptional landscape of plant circadian rhythms using genome tiling arrays. *Genome Biol.* **10**, R17.

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

Ihaka, R. and Gentleman, R. (1996) R: a language for data analysis and graphics. *J. Comput. Graph. Stat.* **5**, 299–314.

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

Karssen, C.M., Brinkhorstvanderswan, D.L.C., Breekland, a.E. and Koornneef, M. (1983) Induction of dormancy during seed development by endogenous abscisic acid : studies on abscisic acid deficient genotypes of *Arabidopsis thaliana* (L) Heynh. *Planta*, **157**, 158–165.

Koornneef, M., Hanhart, C.J., Hillhorst, H.W.A. and Karssen, C.M. (1989) *In vivo* inhibition of seed development and reserve protein accumulation in recombinants of abscisic acid biosynthesis and responsiveness mutants in *Arabidopsis thaliana*. *Plant Physiol.* **90**, 463–469.

Kushiro, T., Okamoto, M., Nakabayashi, K., Yamagishi, K., Kitamura, S., Asami, T., Hirai, N., Koshiba, T., Kamiya, Y. and Nambara, E. (2004) The *Arabidopsis* cytochrome P450 CYP707A encodes ABA 8′-hydroxylases: key enzymes in ABA catabolism. *EMBO J.* **23**, 1647–1656.

Léon-Kloosterziel, K.M., Keijzer, C.J. and Koornneef, M. (1994) A seed shape mutant of Arabidopsis that is affected in integument development. *Plant Cell*, **6**, 385–392.

Léon-Kloosterziel, K.M., Gil, M.A., Ruijs, G.J., Jacobsen, S.E., Olszewski, N.E., Schwartz, S.H., Zeevaart, J.A.D. and Koornneef, M. (1996) Isolation and characterization of abscisic acid-deficient Arabidopsis mutants at two new loci. *Plant J.* **10**, 655–661.

Li, L., Wang, X.F., Stolc, V. *et al.* (2006) Genome-wide transcription analyses in rice using tiling microarrays. *Nat. Genet.* **38**, 124–129.

de Longevialle, A.F., Meyer, E.H., Andres, C., Taylor, N.L., Lurin, C., Millar, A.H. and Small, I.D. (2007) The pentatricopeptide repeat gene OTP43 is required for trans-splicing of the mitochondrial nad1 intron 1 in *Arabidopsis thaliana*. *Plant Cell*, **19**, 3256–3265.

Lurin, C., Andres, C., Aubourg, S. *et al.* (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, **16**, 2089–2103.

Martin, R.C., Liu, P.P. and Nonogaki, H. (2005) Simple purification of small RNAs from seeds and efficient detection of multiple microRNAs expressed in *Arabidopsis thaliana* and tomato (*Lycopersicon esculentum*) seeds. *Seed Sci. Res.* **15**, 319–328.

Matsui, A., Ishida, J., Morosawa, T. *et al.* (2008) Arabidopsis transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. *Plant Cell Physiol.* **49**, 1135–1149.

Meyers, B.C., Tej, S.S., Vu, T.H., Haudenschild, C.D., Agrawal, V., Edberg, S.B., Ghazal, H. and Decola, S. (2004) The use of MPSS for whole-genome transcriptional analysis in Arabidopsis. *Genome Res.* **14**, 1641–1653.

Millar, A.A., Jacobsen, J.V., Ross, J.J., Helliwell, C.A., Poole, A.T., Scofield, G., Reid, J.B. and Gubler, F. (2006) Seed dormancy and ABA metabolism in Arabidopsis and barley: the role of ABA 8′-hydroxylase. *Plant J.* **45**, 942–954.

Muller, K., Tintelnot, S. and Leubner-Metzger, G. (2006) Endosperm-limited Brassicaceae seed germination: abscisic acid inhibits embryo-induced endosperm weakening of *Lepidium sativum* (cress) and endosperm rupture of cress and *Arabidopsis thaliana*. *Plant Cell Physiol.* **47**, 864–877.

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

Nakashima, K., Fujita, Y., Kanamori, N. *et al.* (2009) Three Arabidopsis SnRK2 protein kinases, SRK2D/SnRK2.2, SRK2E/SnRK2.6/OST1 and SRK2I/SnRK2.3, involved in ABA signaling are essential for the control of seed development and dormancy. *Plant Cell Physiol.* **50**, 1345–1363.

Nambara, E. and Marion-Poll, A. (2005) Abscisic acid biosynthesis and catabolism. *Annu. Rev. Plant. Biol.* **56**, 165–185.

Nonogaki, H., Gee, O.H. and Bradford, K.J. (2000) A germination-specific endo-beta-mannanase gene is expressed in the micropylar endosperm cap of tomato seeds. *Plant Physiol.* **123**, 1235–1245.

Ogawa, M., Hanada, A., Yamauchi, Y., Kuwahara, A., Kamiya, Y. and Yamaguchi, S. (2003) Gibberellin biosynthesis and response during Arabidopsis seed germination. *Plant Cell*, **15**, 1591–1604.

Okamoto, M., Kuwahara, A., Seo, M., Kushiro, T., Asami, T., Hirai, N., Kamiya, Y., Koshiba, T. and 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.

Penfield, S., Rylott, E.L., Gilday, A.D., Graham, S., Larson, T.R. and Graham, I.A. (2004) Reserve mobilization in the Arabidopsis endosperm fuels hypocotyl elongation in the dark, is independent of abscisic acid, and requires *PHOSPHOENOLPYRUVATE CARBOXYKINASE1*. *Plant Cell*, **16**, 2705–2718.

Penfield, S., Li, Y., Gilday, A.D., Graham, S. and Graham, I.A. (2006) Arabidopsis ABA INSENSITIVE4 regulates lipid mobilization in the embryo and reveals repression of seed germination by the endosperm. *Plant Cell*, **18**, 1887–1899.

Pérez-Flores, L., Carrari, F., Osuna-Fernández, R., Rodríguez, M.V., Enciso, S., Stanelloni, R., Sanchez, R.A., Bottini, R., Iusem, N.D. and Benech-Arnold, R.L. (2003) Expression analysis of a GA 20-oxidase in embryos from two sorghum lines with contrasting dormancy: possible participation of this gene in the hormonal control of germination. *J. Exp. Bot.* **54**, 2071–2079.

Piskurewicz, U., Jikumaru, Y., Kinoshita, N., Nambara, E., Kamiya, Y. and Lopez-Molina, L. (2008) The gibberellic acid signaling repressor RGL2 inhibits arabidopsis seed germination by stimulating abscisic acid synthesis and ABI5 activity. *Plant Cell*, **20**, 2729–2745.

Pritchard, S.L., Charlton, W.L., Baker, A. and Graham, I.A. (2002) Germination and storage reserve mobilization are regulated independently in Arabidopsis. *Plant J.* **31**, 639–647.

Rymarquis, L.A., Kastenmayer, J.P., Huttenhofer, A.G. and Green, P.J. (2008) Diamonds in the rough: mRNA-like non-coding RNAs. *Trends Plant Sci.* **13**, 329–334.

Saika, H., Okamoto, M., Miyoshi, K. *et al.* (2007) Ethylene promotes submergence-induced expression of OsABA8ox1, a gene that encodes ABA 8 '-hydroxylase in rice. *Plant Cell Physiol.* **48**, 287–298.

Salaita, L., Kar, R.K., Majee, M. and Downie, A.B. (2005) Identification and characterization of mutants capable of rapid seed germination at 10 degrees C from activation-tagged lines of *Arabidopsis thaliana*. *J. Exp. Bot.* **56**, 2059–2069.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D. and Lohmann, J.U. (2005) A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* **37**, 501–506.

Schmitz-Linneweber, C. and Small, I. (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci.* **13**, 663–670.

Schopfer, P. and Plachy, C. (1985) Control of Seed Germination by Abscisic-Acid. III. Effect on Embryo Growth Potential (Minimum Turgor Pressure) and Growth Coefficient (Cell Wall Extensibility) in *Brassica napus* L. *Plant Physiol.* **77**, 676–686.

Seki, M., Narusaka, M., Kamiya, A. *et al.* (2002) Functional annotation of a full-length Arabidopsis cDNA collection. *Science*, **296**, 141–145.

Seo, M., Hanada, A., Kuwahara, A. *et al.* (2006) Regulation of hormone metabolism in Arabidopsis seeds: phytochrome regulation of abscisic acid metabolism and abscisic acid regulation of gibberellin metabolism. *Plant J.* **48**, 354–366.

Shen, Q.X. and Ho, T.H.D. (1995) Functional dissection of an abscisic-acid (Aba)-inducible gene reveals 2 independent Aba-responsive complexes each containing a G-Box and a novel cis-acting element. *Plant Cell*, **7**, 295–307.

da Silva, E.A.A., Toorop, P.E., van Aelst, A.C. and Hillhorst, H.W.A. (2004) Abscisic acid controls embryonic growth potential and endosperm cap weakening during coffee (*Coffea Arabica* cv. Rubi) seed germination. *Planta*, **220**, 251–261.

Stolc, V., Samanta, M.P., Tongprasit, W. *et al.* (2005) Identification of transcribed sequences in *Arabidopsis thaliana* by using high-resolution genome tiling arrays. *Proc. Natl Acad. Sci. USA*, **102**, 4453–4458.

Storey, J.D. (2002) A direct approach to false discovery rates. *J. R. Stat. Soc. Series B Stat. Methodol.* **64**, 479–498.

Storey, J.D. and Tibshirani, R. (2003) Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA*, **100**, 9440–9445.

Tamura, N., Yoshida, T., Tanaka, A., Sasaki, R., Bando, A., Toh, S., Lepiniec, L. and Kawakami, N. (2006) Isolation and characterization of high temperature-resistant germination mutants of *Arabidopsis thaliana*. *Plant Cell Physiol.* **47**, 1081–1094.

Tatematsu, K., Kamiya, Y. and Nambara, E. (2008) Co-regulation of ribosomal protein genes as an indicator of growth status: comparative transcriptome analysis on axillary shoots and seeds in Arabidopsis. *Plant Signal. Behav.* **3**, 450–452.

Toh, S., Imamura, A., Watanabe, A. *et al.* (2008) High temperature-induced abscisic acid biosynthesis and its role in the inhibition of gibberellin action in Arabidopsis seeds. *Plant Physiol.* **146**, 1368–1385.

Toyoda, T. and Shinozaki, K. (2005) Tiling array-driven elucidation of transcriptional structures based on maximum-likelihood and Markov models. *Plant J.* **43**, 611–621.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V. and Provart, N.J. (2007) An ''electronic fluorescent pictograph'' browser for exploring and analyzing large-scale biological data sets. *PLoS ONE*, **2**, e718.

Wu, C.T., Leubner-Metzger, G., Meins, F. and Bradford, K.J. (2001) Class I beta-1,3-glucanase and chitinase are expressed in the micropylar endosperm of tomato seeds prior to radicle emergence. *Plant Physiol.* **126**, 1299–1313.

Yamada, K., Lim, J., Dale, J.M. *et al.* (2003) Empirical analysis of transcriptional activity in the Arabidopsis genome. *Science*, **302**, 842–846.

Yamagishi, K., Tatematsu, K., Yano, R., Preston, J., Kitamura, S., Takahashi, H., McCourt, P., Kamiya, Y. and Nambara, E. (2009) CHOTTO1, a double AP2 domain protein of *Arabidopsis thaliana*, regulates germination and seedling growth under excess supply of glucose and nitrate. *Plant Cell Physiol.* **50**, 330–340.

Zeller, G., Henz, S.R., Widmer, C.K., Sachsenberg, T., Ratsch, G., Weigel, D. and Laubinger, S. (2009) Stress-induced changes in the *Arabidopsis thaliana* transcriptome analyzed using whole-genome tiling arrays. *Plant J.* **58**, 1068–1082.

Zhang, X.Y., Yazaki, J., Sundaresan, A. *et al.* (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell*, **126**, 1189–1201.

