RESEARCH PAPER
https://doi.org/10.1071/FP21315

Functional Plant Biology

DELAY OF GERMINATION 1 (DOG1) regulates dormancy in dimorphic seeds of *Xanthium strumarium*

Iman Nemati[A], Mohammad Sedghi[A,*], Ghasem Hosseini Salekdeh[B,C], Reza Tavakkol Afshari[D], Mohammad Reza Naghavi[E] and Somayeh Gholizadeh[F]

For full list of author affiliations and declarations see end of paper

***Correspondence to**:
Mohammad Sedghi
Department of Plant Production and Genetics Engineering, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran
Email: m_sedghi@uma.ac.ir

Handling Editor:
Peter Bozhkov

Received: 12 June 2021
Accepted: 23 March 2022
Published: 16 May 2022

Cite this:
Nemati I *et al.* (2022)
Functional Plant Biology
doi:10.1071/FP21315

ABSTRACT

Seed dormancy ensures plant survival but many mechanisms remain unclear. A high-throughput RNA-seq analysis investigated the mechanisms involved in the establishment of dormancy in dimorphic seeds of *Xanthium strumarium* (L.) developing in one single burr. Results showed that *DOG1*, the main dormancy gene in *Arabidopsis thaliana* L., was over-represented in the dormant seed leading to the formation of two seeds with different cell wall properties. Less expression of *DME/EMB1649*, *UBP26*, *EMF2, MOM, SNL2, and AGO4* in the non-dormant seed was observed, which function in the chromatin remodelling of dormancy-associated genes through DNA methylation. However, higher levels of *ATXR7/SDG25, ELF6*, and *JMJ16/PKDM7D* in the non-dormant seed that act at the level of histone demethylation and activate germination were found. Dramatically lower expression in the splicing factors *SUA, PWI*, and *FY* in non-dormant seed may indicate that variation in RNA splicing for ABA sensitivity and transcriptional elongation control of *DOG1* is of importance for inducing seed dormancy. Seed size and germination may be influenced by respiratory factors, and alterations in ABA content and auxin distribution and responses. TOR (a serine/threonine-protein kinase) is likely at the centre of a regulatory hub controlling seed metabolism, maturation, and germination. Over-representation of the respiration-associated genes (*ACO3*, *PEPC3*, and *D2HGDH*) was detected in non-dormant seed, suggesting differential energy supplies in the two seeds. Degradation of ABA biosynthesis and/or proper auxin signalling in the large seed may control germinability, and suppression of endoreduplication in the small seed may be a mechanism for cell differentiation and cell size determination.

Keywords: ABA, *DOG1*, dormancy, germination, phytohormone, respiration, transcriptome, *Xanthium strumarium*.

Introduction

Controlling the timing of germination, seed dormancy is vital to enhance the chances of plant survival in a changing environment. Even though its biological significance is clear, many aspects of molecular mechanisms underlying seed dormancy induction, maintenance, and alleviation remain largely elusive. The reason for this arises from special complexities for studying seed dormancy and the profound impact of environmental conditions (Penfield and MacGregor 2017; Klupczyńska and Pawłowski 2021). Therefore, it seems that a complex genetic and environmental factors are involved in seed dormancy and germination, which require further investigation.

Hormonal regulation may be a highly conserved mechanism of seed dormancy among spermatophytes. It has been demonstrated that dormancy is induced by ABA during seed development on the parent plant (Graeber *et al.* 2012). After seed dispersal, germination is preceded by a decline in ABA in imbibed seeds, which results from ABA catabolism (Arc *et al.* 2013). It has long been known that a precise spatial and temporal crosstalk between ABA and other phytohormones such as GA (Liu *et al.* 2016), auxin (Liu and Schmidt 2012), brassinosteroids (Hu and Yu 2014), ethylene (Wang *et al.* 2013), jasmonic acid (JA) (Singh and Singh 2012) and cytokinin (Shu *et al.* 2016) is required

to regulate seed dormancy and germination. Other mechanisms, which might be independent of hormones, or specific to the seed dormancy pathway, are also emerging from genetic analysis of 'seed dormancy mutants'. For example, the Arabidopsis *thaliana* L. *DELAY OF GERMINATION1* (*DOG1*) gene, whose precise function is still unknown, has been identified as a master regulator of seed dormancy (Bentsink *et al.* 2006). Several studies have also uncovered the molecular mechanisms up or downstream of *DOG1* (Graeber *et al.* 2014; Huo *et al.* 2016; Née *et al.* 2017). Although several studies have suggested that the *DOG1*-mediated mechanism may be distinct from the ABA (reviewed by Shu *et al.* 2016), a recent report demonstrated at least one converge between the two mechanisms (Née *et al.* 2017). Furthermore, chromatin remodelling through histone ubiquitination, methylation, and acetylation, which could lead to gene silencing or transcription elongation, may play a significant role in seed dormancy regulation (Nonogaki 2014). Genetic analyses have also identified many transcription factors (TFs) involved in seed development and dormancy induction (Vaistij *et al.* 2013; Weiste and Dröge-Laser 2014; Née *et al.* 2017) that target the promoter of dormancy-associated genes. In addition, the regulation of mRNA processing and function, including splicing, modification, transport, translation, and RNA decay is crucial to the control of gene expression (Köster *et al.* 2017). For instance, Cyrek *et al.* (Cyrek *et al.* 2016) revealed that mutants in RNA 3′ processing factors like *fy-2* display weakened seed dormancy in parallel with defects in *DOG1* proximal polyadenylation site selection, suggesting that the short *DOG1* transcript is functional.

Despite these categorisations, the simultaneous presence of multiple dormancy classes in one seed makes the understanding of molecular mechanisms of seed dormancy very difficult. *Xanthium strumarium* (L.) is an annual herb belonging to the Asteraceae family, which grows almost all over the world. Because of its anti-cancer and anti-tumor bioactive properties, this plant is considered to develop new drugs against cancer (Vaishnav *et al.* 2015). *X. strumarium* also produces two seeds in one single burr, each differing in size and dormancy status. The relationship between seed size and dormancy level has already been studied (Rubio de Casas *et al.* 2017). Seed dormancy and size were hypothesised as alternative adaptive strategies, determining the environment in which a certain lineage will be predominant. Producing a larger non-dormant seed beside a small dormant seed in one burr, *X. strumarium* recruits a strong strategy to define a wider window in which the maximum seed survival, germination, and seedling establishment achieve. Therefore, *X. strumarium* can be an interesting plant to study seed dormancy. In this study, the transcriptome profiles of dormant and non-dormant seeds of *X. strumarium* grown on the same parent plant experiencing the same environmental conditions were characterised to answer this question what regulatory mechanisms underlying the differential induction of seed dormancy. This approach helps us to better understand more about the fundamental biological processes of dormancy in a wider range of species, and enables us to further develop new models for further research.

Materials and methods

Plant materials and growth condition

Seeds of *Xanthium strumarium* (L.) were collected from the research site of Tehran University, Karaj, Iran (35.8°92′21″N, 50.96°70′44″E) at 1300 m altitude. All the large seeds in each burr were precisely excised, washed and surface sterilised followed by sowing on two layers of water-saturated filter paper (Whatman, GE Healthcare, UK) in Petri dishes and kept in an incubator at 23°C in the darkness. After 14 days, seedlings were transferred to the growth chamber, planted in 4 L pots filled with field soil:peat (1:1 v/v) and grown under the controlled condition at 16–18°C (night/day) photoperiod with 8 h light and relative humidity of 40% for 4 months. Pots were then transferred to a long-day greenhouse (16-h photoperiod) at 20–25°C for the reproductive phase stimulation. We tagged burrs with coloured ribbon from the time of emergence, and samples were collected at 3, 10, 20, and 30 days after flowering as well as fully matured seeds (Fig. 1). Dormant and non-dormant seeds were accurately excised, immediately frozen using liquid nitrogen, and stored at −80°C until RNA extraction.

Germination essay

A standard germination test was performed on fully matured non-dormant and dormant seeds in four replications, with 100 seeds for each replication. Glass containers (210 mm × 27 mm) were kept at 23°C under an alternating cycle of 12–12 h (illumination–darkness). Seeds were placed onto filter papers fully moistened with sterile distilled water and kept in the incubator for 7 days (the time in which the number of germinated seeds is unchanged for 3 days). Germinated seeds were counted daily and the germination percentage was calculated according to the number of germinated seeds.

RNA isolation, library preparation, and transcriptome sequencing

All samples from non-dormant and dormant seeds were ground into a fine powder using a mixer mill (MM200, Retsch, Germany). Total RNA was extracted by using Invitrogen TRIzol Reagent (Invitrogen, California, USA) according to the manufacturer's instructions. The extracted RNAs were

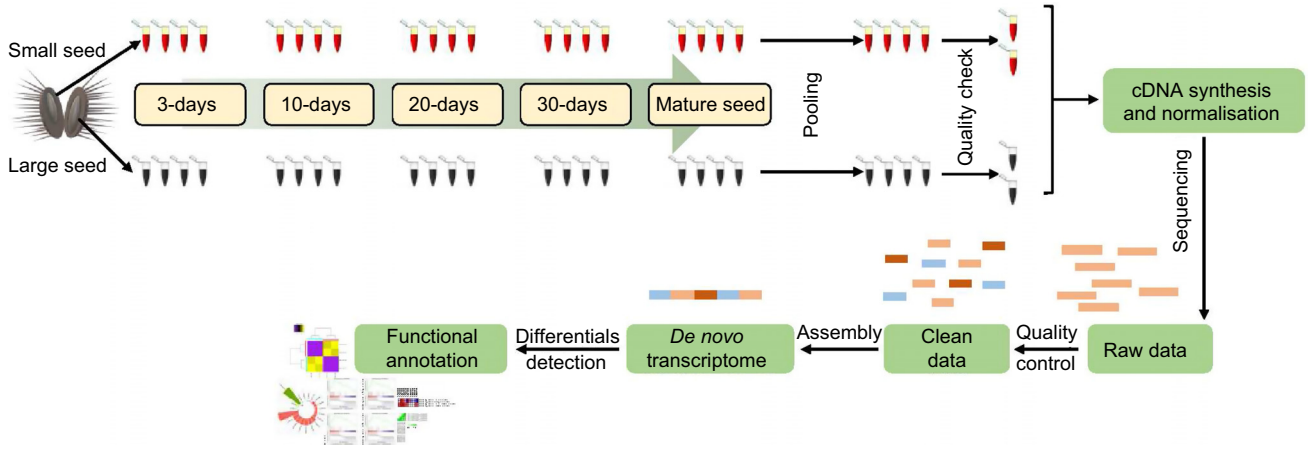


Fig. 1. Workflow of our experiment to identify the most important differentially expressed genes between developing non-dormant and dormant seeds of *X. strumarium*.

mixed based on their concentration to obtain pooled RNAs from the developmental stages of each seed, separately. Nanodrop Spectrophotometer (NP80 NanoPhotometer, IMPLEN, Munich, Germany), and Agilent 2100 (Agilent Technologies, Santa Clara, CA, USA) were used for the quality and quantity of RNA assessment. All the RNA samples had A260/A280 nm ratios between 2 and 2.1 and RNA integrity number (RIN) between 7.9 and 9.5. The qualified samples were sequenced using an Illumina Hiseq 2500 analyser (Beijing Genomics Institute, Shenzhen, China) with 150-bp paired-end reads on four samples (seed × replicate).

De novo transcriptome assembly

After the removal of low-quality reads using FastQC (ver. 0.11.5), high-quality reads were obtained for subsequent analysis. The clean reads were generated by removing adaptor sequences and low-quality reads deposited in NCBI Sequence Read Archive (SRA) Sequence Database with accession number PRJNA809034. The 'Trinity' programme (Grabherr *et al.* 2011) was used to assemble the clean reads and obtain non-redundant unigenes of the *X. strumarium*. In short, reads that overlapped were assembled to generate contigs, and then contigs were joined into scaffolds that were further assembled through gap-filling to generate unigenes. In this study, a default k-mer size of 25 bp was set for the *de novo* transcriptome assembly. All other parameters were set as default values and the length of the assembled unigenes used for further study was ≥200 bp. The CD-HIT-EST (ver. 4.6.1) (Li and Godzik 2006) was further used to cluster the assembled contigs based on sequence identity threshold 0.95 to full-length transcripts generation. The clean reads were subsequently mapped on the transcriptome using Bowtie2 (Liu and Schmidt 2012).

Identification of DEGs

The RPKM (Reads Per kb per Million reads) of each unigene in dormant and non-dormant samples was calculated by normalisation of the number of mapped clean reads per unigene. The edgeR programme was used to determine the differential expression genes (DEGs) with a log-fold expression change (log FC) > 2 or <−2 using a threshold of false discovery rates (FDR < 0.05).

Transcriptome annotation

Functional annotations were carried out by comparison of assembled transcripts corresponding to DEGs against the public databases. The BLASTX tool was employed to search the sequence similarity in the NCBI non-redundant protein sequences (nr) database using *Arabidopsis Q*as a reference organism. All the BLAST results against the Nr database were loaded onto the STRING website for Gene Ontology (GO) enrichment analysis and protein-protein interactions prediction. Results were prepared using Cytoscape 3.8.1. GSEAPreranked test was performed to identify the enriched pathways using GSEA software followed by loading results into Cytoscape 3.8.1 to provide the enrichment map (Reimand *et al.* 2019). The protein sequences corresponded to up-and downregulated genes were separately loaded onto to the transcription factors detection. Furthermore, loading protein sequences onto http://pfam.xfam.org/search#tabview=tab1 allowed us to predict protein domains. Figures were drawn using R (ver. 4.0.2) and Microsoft Excel (ver. 2019).

Quantitative real-time PCR (qRT-PCR) analysis

For qRT-PCR, cDNA was synthesised from 1 μL of each RNA sample using Omniscript reverse transcriptase (Qiagen, Germany). The sequence information from the RNA-seq

data was utilised for primer design using the Primer-BLAST online programme (https://www.ncbi.nlm.nih.gov/tools/primer-blast). qRT-PCR assays were performed on a LightCycler 96 Real-Time PCR System (Roche Life Science, Germany) with SYBR Premix EX TaqII (Takara Bio Inb, Japan), and for four biological replicates of each developmental stage (seed × developmental time × replicate) to assess the expression levels of five DEGs involved in seed dormancy- and development-associated processes along with *TUBULIN BETA CHAIN 4* (*TUB4*) as a reference gene. The *TUB4* gene was recognised as an optimal reference gene in different crops and wild plants during fruit and seed development (Wei *et al.* 2013; Ferraz dos Santos *et al.* 2016; Liu *et al.* 2016; Pereira *et al.* 2017; Jaiswal *et al.* 2019; Zhou *et al.* 2019). The equation $2^{-\Delta\Delta C_T}$ was used for the calculation of relative transcription levels (Livak and Schmittgen 2001).

Results

Dimorphic seeds of *X. strumarium* differed in germinability

The germination test highlighted a delayed radicle protrusion and less germination percentage in the small seed ($P < 0.05$; Fig. 2). There were no significant differences germination between non-dormant and dormant seeds after Day 4. The ~86% germination in large seed was ~58% more than the small one at the end of the test period (Fig. 2).

De novo assembly of transcriptome revealed the differences between two seeds

We detected 189.22 million clean reads and the Trinity assembler allowed us to produce 191 192 high-quality contigs with a length of 201 to 13 863 with an average of 989.69 bp. Fig. 3*a* pinpointed the distribution of identified contigs based on their size, and the similarity/dissimilarity of sample libraries from large and small seeds was shown in Fig. 3*b*. Differentially heatmap representing the DEGs with high fold changes (>2) between large and small seeds of *X. strumarium* in the four clusters with distinct expression patterns was depicted in Fig. 3*c*. Based on our volcano plot analysis and the comparison of DEGs affected by different seed types, we identified differentially expressed genes in the large compared to the small seeds (Fig. 3*d*). The comparison between the large and small seeds identified 324 sequences with different expression levels (Fig. 4*a*), including 188 under-represented and 136 over-represented transcripts in the large compared to the small seed (Fig. 4*b*). As detailed in Fig. 4*b*, the most frequently upregulated and downregulated sequences in large seed were those having a $\log_2$FC = ±(8–10).

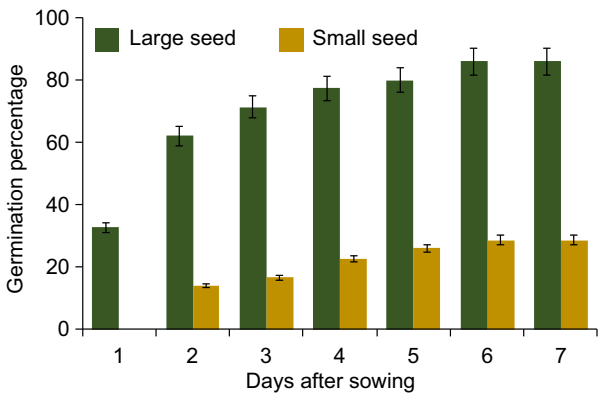


Fig. 2. Germination percentage of fully matured large (non-dormant) and small (dormant) seeds of *X. strumarium* (means of four independent replication ±5%).

Most of the enriched gene ontologies were seed-specific

BLASTX searches followed by introducing over- and under-represented sequences to STRING database were adopted for identification of proteins corresponding to the detected sequences and subsequent classification into three major categories: (1) biological process; (2) molecular function; and (3) cellular component (Fig. 5*a*, *b*). We observed 46, 1, and 13 GO terms that positively enriched in the 83, 75, and 83 sequences of biological process, molecular function, and cellular component, respectively (Fig. 5*a*). While there were 97, 12, and 91 under-represented sequences that belonged to 34, 15, and 9 GO terms of biological process, molecular function, and cellular process, respectively (Fig. 5*b*). The analysis revealed that 300 of 324 detected sequences were assigned on at least one protein. In the biological process category, upregulated and downregulated transcripts in non-dormant seed belonged to 46 and 34 GOs, respectively (Fig. 5*c*). As regards the molecular function, the only GO was 1,3-beta-D-glucan synthase under-represented in non-dormant seed (Fig. 5*c*), while the upregulated transcripts were mainly associated with binding processes (Fig. 5*c*). Moreover, *CALLOSE SYNTHASEs* and *GLUCAN SYNTHASE-LIKE* involved in molecular functions were negatively enriched in the non-dormant seed. Gene ontology showed that most of the enriched GOs (~85%) were seed-specific, indicating their special roles in the physiological processes of the dimorphic seeds of *X. strumarium*. For example, the non-dormant seed was specifically enriched for transcripts involved in response to the hormone, cellular response to stress, regulation of cell cycle, cellular respiration, and regulation of lipid and ABA biosynthesis processes, while dormant seed highly upregulated genes for response to post-embryonic development, cell wall organisation, seed development, cellular carbohydrate metabolic process, and regulation of cell shape (Fig. 6).

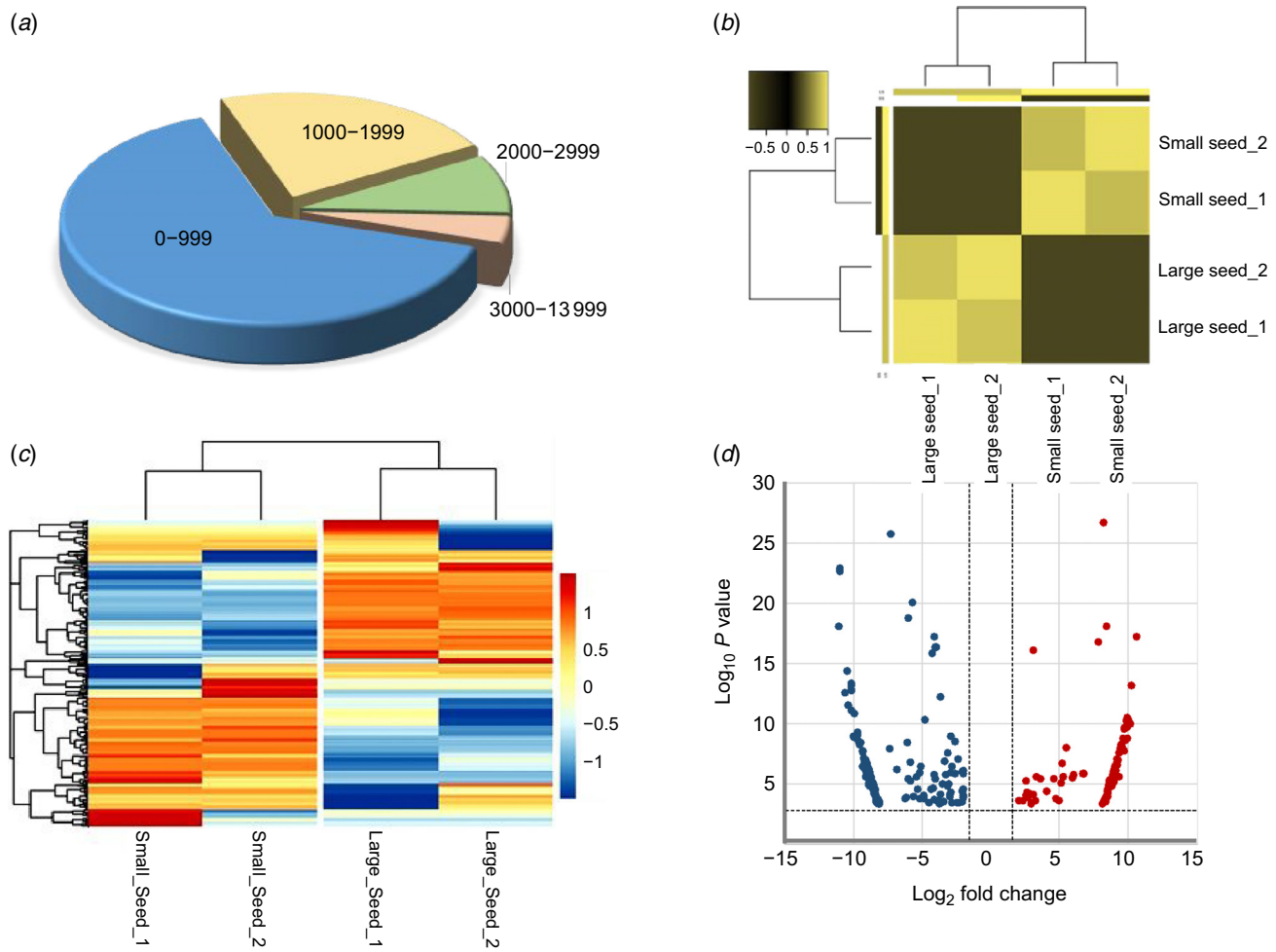


Fig. 3. Results of detected sequences analysis; (*a*) distribution of identified transcripts according to sequence length; (*b*) heatmap of sample-to-sample distance clustering, illustrating similarities and dissimilarities between individual samples; (*c*) heatmap of differentially expressed genes between dormant and nondormant developing seeds of *X. strumarium*; (*d*) volcano plot showing the total distribution of differentially expressed genes according to $\log_2$ fold-change versus $\log_{10}$ *P*-value of identified differentially expressed genes in developing non-dormant compared with the dormant seed of *X. strumarium*.

DOG1 was over-represented in the dormant seed of *X. strumarium*

Less abundance of the *DOG1* mRNAs (log2FC = −3.7), the central regulator of seed dormancy in *A. thaliana*, were detected in the non-dormant seed (Fig. 6*b*). Gene interaction network (Fig. 7) highlighted an interaction between *DOG1* and a 12*S* globulin storage protein involved in seed maturation, CRUCIFERIN 2 (CRU2), which showed less expression in the non-dormant seed (log2FC = −5.33).

Different chromatin remodelling was observed in two seeds of *X. strumarium*

Less expression of *DEMETER* (*DME*), also known as *EMB1649*, ($\log_2$FC = −8.8; Fig. 6*b*) and *UBIQUITIN-SPECIFIC PROTEASE 26* (*UBP26*) (log2FC = −9.92) in the non-dormant seed (Fig. 6*b*) was observed. In the case of *UBP26*, qRT-PCR showed that there was a distinct difference for UBP26 between samples from non-dormant and dormant seeds in different developmental times. Although relative expression levels of non-dormant seed revealed a more abundance at early developmental stages, a significant increase was observed in dormant seed at the late stage of seed development (Fig. 8). Histone deubiquitination by UBP26 action maintains the H3K9 methylation induced by the histone methyltransferase KYP (KRYPTONITE), a negative regulator of dormancy, leading to DNA methylation resulting in gene silencing in heterochromatin (Zheng *et al.* 2012). *DME* also is a transcription activator, acting to remove methylated bases, by a glycosylase/lyase mechanism, from maternal genes in the central cell of the embryo sac. *dme-2* mutants display delayed seed germination, aberrant cellular proliferation,

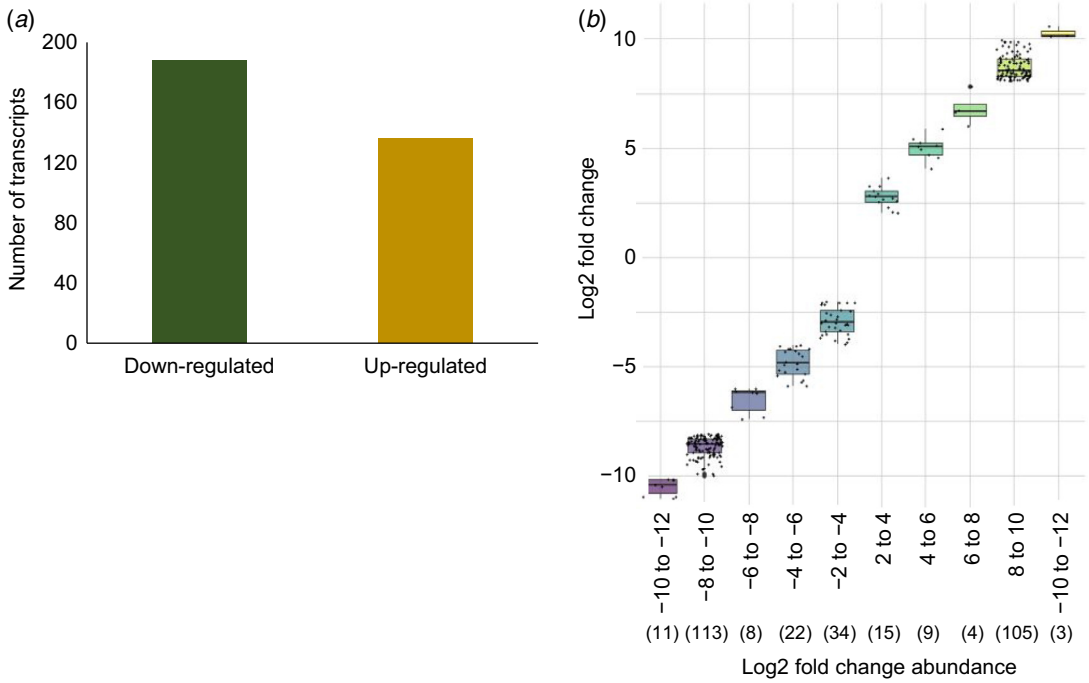


Fig. 4. Total number (*a*) and boxplot of distribution (*b*) of differentially expressed genes of developing non-dormant in comparison with dormant seeds of *X. strumarium*.

and differentiation (Kim *et al.* 2021). Furthermore, *EMBRYONIC FLOWER 2* (*EMF2*) ($log_2FC = -8.17$; Fig. 6*b*), *MAINTENANCE OF METHYLATION* (*MOM*) ($log_2FC = -8.48$), a methylation activator, and *ARGONAUTE4* (*AGO4*) ($log_2FC = -8.94$) showed less expression levels in the non-dormant seeds. *EMF2* encodes a Polycomb group that maintains gene silencing via histone modification (Kim *et al.* 2010). *AGO4* also is a translation regulator, which acts for gene silencing by RNA-directed DNA methylation through H3K9 methylation and affects the expression of dormancy genes (Qi *et al.* 2006; Nonogaki 2014). However, we found higher levels of SET DOMAIN PROTEIN 25 (SDG25) transcripts ($log_2FC = 6$), encoding ARABIDOPSIS TRITHORAX-RELATED7 (ATXR7), and *EARLY FLOWERING 6* (*ELF6*), a transcription factor that functions in chromatin remodelling through H3K9 demethylase activity, in the non-dormant seed. Another over-represented chromatin remodelling gene in the non-dormant seeds was *JUMONJI DOMAIN-CONTAINING PROTEIN 16* (*JMJ16*), also known as *PKDM7D*, ($log_2FC = 8.1$), interacts with *EMF2*, *ELF6*, and *ATXR7/SDG25* (Fig. 7). In addition to activation of dormancy-related genes, repression of seed germination genes, a further regulation for dormancy maintenance, was found in the non-dormant seed (Shu *et al.* 2016). *SIN3-LIKE 2* (*SNL2*) that positively regulates seed dormancy, interacted with *DOG1* (Fig. 7), and showed under-representation ($log_2FC = -8.26$) in the non-dormant seed. *ETHYLENE RESPONSE FACTOR 110* (*ERF110*) transcripts, which is one of the members of the ERFs subfamily and appears to be controlled by ethylene at both the transcriptional and posttranscriptional level (Heyman *et al.* 2018), represented a similar expression pattern ($log_2FC = -2.07$). All together, these confirmed the important role of epigenetic regulations in dormancy induction during seed development of *X. strumarium*.

Genes involved in cellular respiration showed higher transcription in non-dormant seed

We found a positive enrichment of energy production and cellular respiration in the non-dormant seed (Fig. 6*a*). Over-representation of three respiration-associated genes, *ACOTINATE HYDRATASE 3* (*ACO3*) (log FC = 8.12), *PHOSPHOENOLPYRUVATE CARBOXYLASE 3* (*PEPC3*) (log FC = 8.42), and *D-2-HYDROXYGLUTARATE DEHYDROGENASE* (*D2HGDH*) (log FC = 8.48), was detected in the non-dormant seed, suggesting differences in energy supplies between two seeds. In the case of ACO3, the qRT-PCR showed that the relative expression levels of non-dormant seed accounted for a greater abundance at early stages of seed development and maturation than dormant seed (Fig. 8). We also showed a higher expression level of *KIN 10* in the non-dormant seed (log2FC = 2.28; Fig. 6*a*). Sensing and signalling deprivation of sugar and energy, KIN10 targets a remarkably broad array of genes that orchestrate transcription networks,

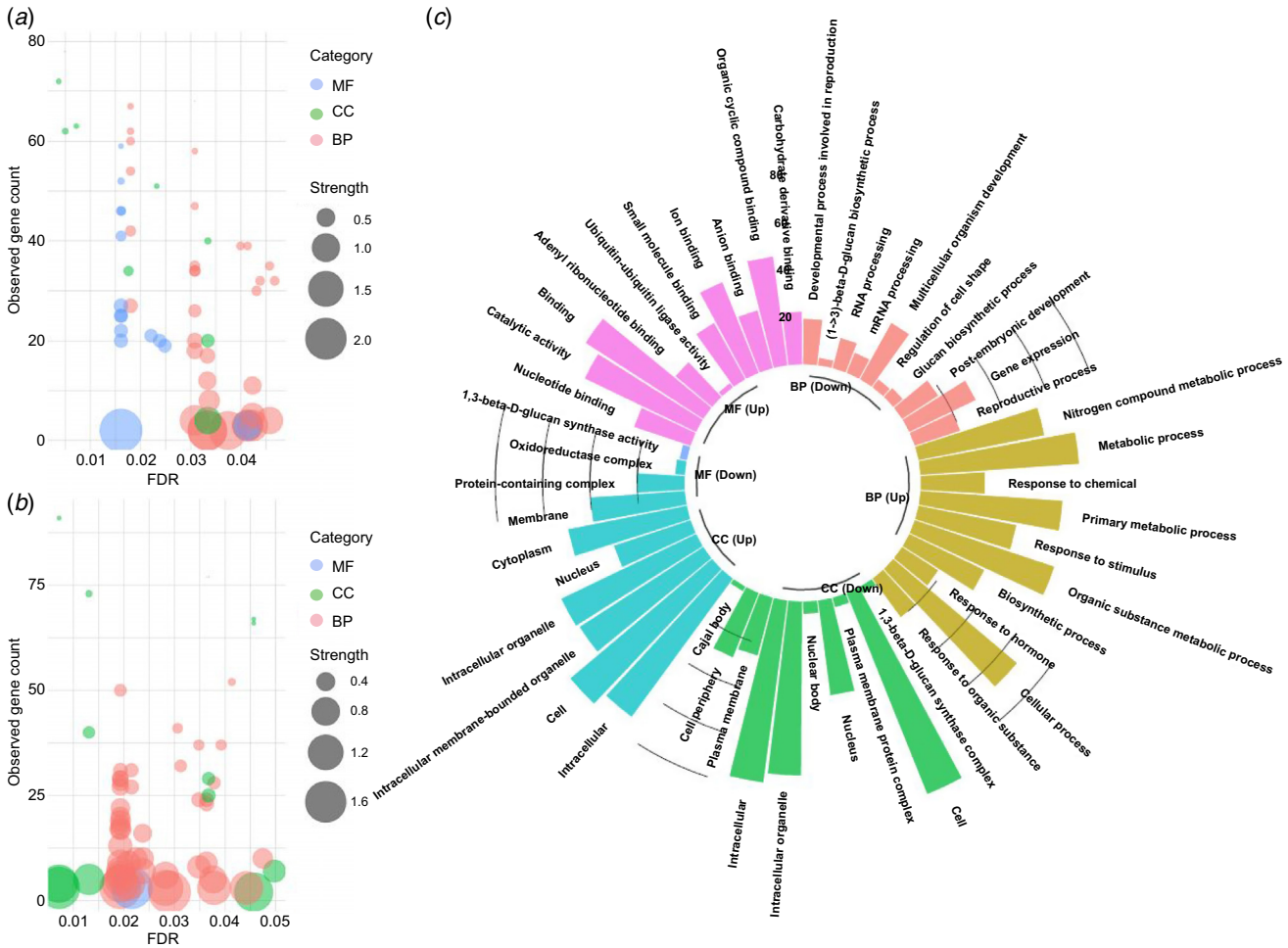


Fig. 5. Distribution of upregulated (*a*) and downregulated (*b*) genes in developing non-dormant versus dormant seeds of *X. strumarium*. Three categories identified: (1) biological process (BP); (2) molecular function (MF); and (3) cellular process (CC). Top 10 significantly (FDR < 0.05) enriched GOs of upregulated and down regulated genes in developing non-dormant and dormant seeds of *X. strumarium*.

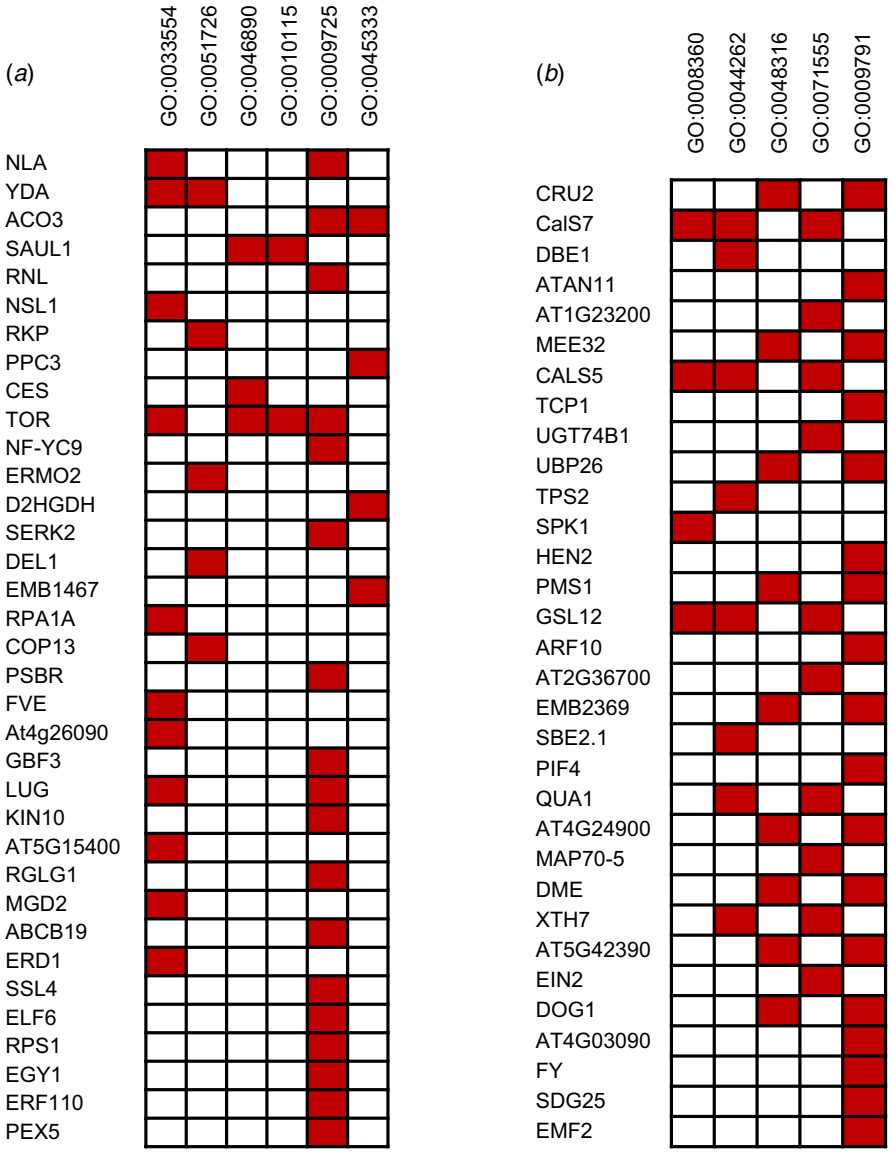
promote catabolism and suppress anabolism (Baena-González *et al.* 2007).

ABA and auxin biosynthesis and signalling were differently regulated between the two seeds

Results showed enrichment of the ABA regulatory pathway for over-represented transcripts in non-dormant seed (Fig. 6*a*). Our data revealed a higher accumulation of SENESCENCE-ASSOCIATED E3 UBIQUITIN LIGASE 1 (SAUL1) (log FC = 8.31; Fig. 6*a*), also known as ARABIDOPSIS THALIANA PLANT U-BOX 44 (PUB44), and TARGET OF RAPAMYCIN (TOR) (log FC = 9.87; Fig. 6*a*) transcripts in the non-dormant seed. Moreover, data analysis confirmed the interaction between *TOR* and *SNF1-RELATED PROTEIN KINASE 1.1* (*SnRK1.1*), also known as *KIN10*, which showed a higher expression level in the non-dormant seed (Fig. 7). Two auxin response regulators *ARABIDOPSIS THALIANA RNA LIGASE* (*RNL*) (log2FC = 3.26; Fig. 6*a*) and *ARABIDOPSIS THALIANA ATP-BINDING CASSETTE B19* (*ABCB19*) (log2FC = 8.12; Fig. 6*a*) were also enriched in the non-dormant seeds. We found interactions between *RNL* with *AGO4* and *AUXIN RESPONSE FACTOR 10* (*ARF10*), which showed less ($log_2FC = -8.26$) transcript level in the non-dormant seed (Fig. 7). In addition, under-representation ($log_2FC = -4.32$) of a modulator of auxin levels, *UDP-GLUCOSYL TRANSFERASE7B1* (*UGT7B1*), was found in the non-dormant. All together, these indicated that biosynthesis and signalling pathways of two hormones ABA and auxin may differently regulate between the two seeds of *X. strumarium*.

mRNA processing was a mechanism associated with seed dormancy

In our study, mRNA processing in non-dormant seed was negatively enriched compared to the dormant seed. We found dramatically fewer expression levels in three splicing factors *SUPPRESSOR OF ABI3–5* (*SUA*) ($log_2FC = -10.18$) as well as PWI ($log_2FC = -10.94$), and AT3G12640



Positively enriched	GO:0009725	Response to hormone
	GO:0033554	Cellular response to stress
	GO:0051726	Regulation of cell cycle
	GO:0045333	Cellular respiration
	GO:0046890	Regulation of lipid biosynthetic process
	GO:0010115	Regulation of abscisic acid biosynthetic process
Negatively enriched	GO:0009791	Post-embryonic development
	GO:0071555	Cell wall organisation
	GO:0048316	Seed development
	GO:0044262	Cellular carbohydrate metabolic process
	GO:0008360	Regulation of cell shape

Fig. 6. Detected transcripts involved in significantly (FDR < 0.05) enriched development- and dormancy-associated GOs for (*a*) over-represented and (*b*) under-represented transcripts in non-dormant seed in comparison with the dormant seed of *X. strumarium*. The table below describes the GOs mentioned in (*a*) and (*b*) in detail.

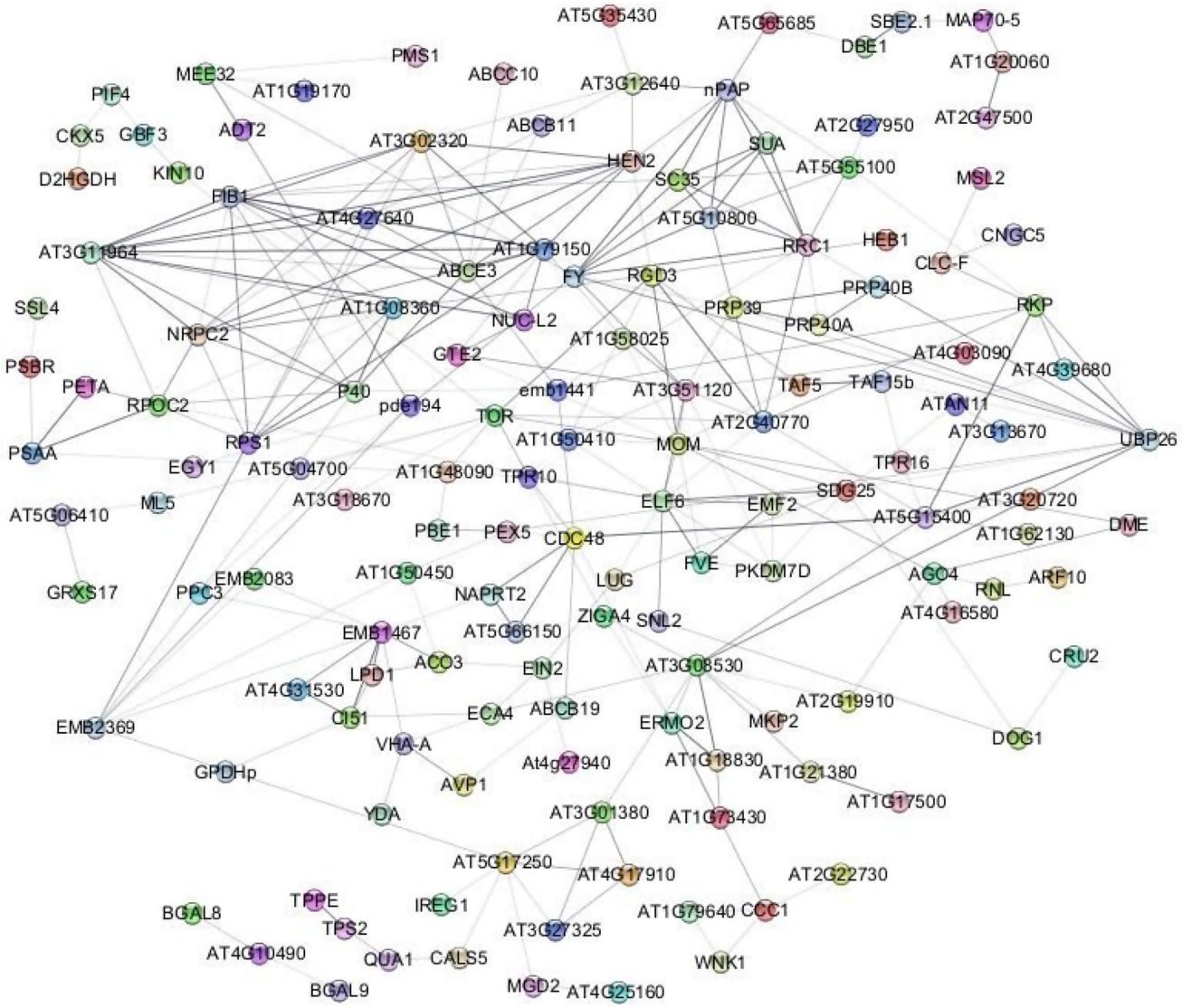


Fig. 7. Gene network showing the interaction between significant (FDR < 0.05) upregulated and downregulated genes in developing non-dormant compared to the dormant seed of *X. strumarium*. Disconnected nodes were removed from the figure.

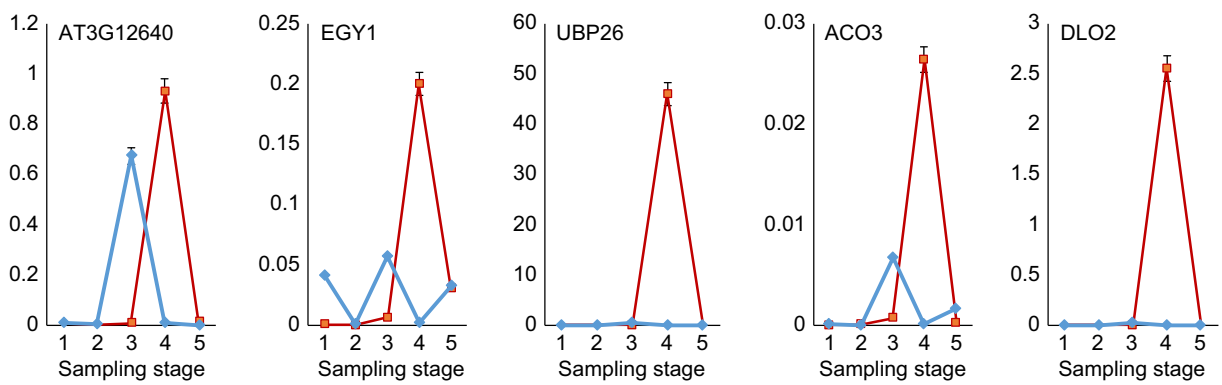


Fig. 8. Relative expression levels of development involved genes for qRT-PCR experiment at five-time intervals from dormant and non-dormant developing seeds of *X. strumarium*. Data are relative expression in non-dormant versus dormant seeds. Sampling stages are shown by 1 to 5, which represent 3, 10, 20, and 30 days after burr emergence, and fully matures seeds, respectively.

($log_2FC = -10.9$), both containing RNA binding (RRM/RBD/RNP motifs) family protein in the non-dormant seed. qRT-PCR also showed that there was a significant difference for AT3G12640 between samples from non-dormant and dormant seeds in different developmental times; however, relative expression levels of non-dormant seed revealed a less abundance at late developmental stages compared to the dormant seed. The expression of *FY*, encoding flowering time control protein FY with a direct effect on DOG1 transcripts polyadenylation, also revealed a less expression in non-dormant seed (log FC = −8.4; Fig. 6*b*). These results suggest that RNA splicing factors may have a key role in dormancy induction to *X. strumarium* seeds.

Mechanisms involved in the cell cycle may differ in dimorphic seeds of *X. strumarium*

The expression of *YODA/YDA* ($log_2FC = 8.6$), Somatic embryogenesis receptor kinase 2 (SERK2) ($log_2FC = 7.8$), and *ENDOPLASMIC RETICULUM MORPHOLOGY 2* (*ERMO2*)/*SEC24a* ($log_2FC = 9.08$; Fig. 6*a*) increased differentially in non-dormant seed of *X. strumarium* (Fig. 6*a*). *ERMO2* functions in setting up mechanisms involved in cell proliferation and mediating the membrane fusion through SNARE binding (Ebine *et al.* 2008), and SERK-mediated signals regulate division patterns of vascular precursors and ground tissue stem cells, likely via the YDA-MKK4/5 cascade, during embryo development. Furthermore, the abundance of *COP9 SIGNALOSOME COMPLEX SUBUNIT 3* (*CSN3*) ($log_2FC = 9.56$), also known as *FUS11*, as well as DP-E2F-like1 (*DEL1/E2Fe*) ($log_2FC = 8.5$) were significantly higher in the non-dormant seed. Similarly, we observed a higher expression of *ENHANCER OF VARIEGATION3* (*EGY1*) ($log_2FC = 10.14$), which is a probable zinc metalloprotease and interacts with ABA, in the non-dormant seed. qRT-PCR also showed that relative expression levels of *EGY1* in the non-dormant seed were more abundant at early developmental stages compared to the dormant seed (Fig. 8). However, the *DMR6-LIKE OXYGENASE 2* (*DLO2*), belonging to 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (2-ODD), transcript level was less in non-dormant seed ($log_2FC = -7.31$) and confirmed by qRT-PCR results (Fig. 8). All together, this indicates that cell cycle and expansion control may differ in dimorphic seeds of *X. strumarium*.

Validation by qRT-PCR

Technical and biological variations in the data were checked by performing qRT-PCR of four independent biological replicates for five developmental times of both dormant and non-dormant seeds of *X. strumarium*. We selected five DEGs including *UBP26*, *ACO3*, *DLO2*, *EGY1*, and *AT36122640*. qRT-PCR results were highly consistent with those of RNA sequencing and there was a simple linear regression ($R^2 = 0.8$) based on log_2FC obtained from RNAseq data (X), and the sum of values of different developmental stages from qRT-PCR (Y). Furthermore, qRT-PCR showed that although there were differences in relative expression levels between different seed types, and distinct developmental times, most of the selected genes (*AT3G12640*, *EGY1*, *UBP26*, and *ACO3*) were expressed at late stages (stages 4 and/or 5; Fig. 8) of seed development in dormant seed, suggesting a delay in developmental processes of the dormant seed of *X. strumarium* compared to the non-dormant seed.

Discussion

The differences between germinability of two fresh seeds of *X. strumarium* confirms different dormancy status, which is in line with previous results (Katoh and Esashi 1975). An 85% seed-specific GOs also highlighted a great difference in the triggered developmental processes between the two seeds. A compilation of different mechanisms ranging from dormancy-specific genes activation, epigenetic and hormonal regulations, cell wall modifications, and mRNA processing had differentially been activated between the two seeds. These mechanisms have previously been well described in *A. thaliana* (Sugliani *et al.* 2010; Graeber *et al.* 2012, 2014; Nonogaki 2014).

DOG1 functions in the regulation of dormancy in seeds of *X. strumarium*

We noticed less expression of *DOG1* in the non-dormant seed. Although several studies have been carried out on *DOG1* (Bentsink *et al.* 2006; Huo *et al.* 2016; Née *et al.* 2017), the molecular and biochemical functions of its protein remain largely unknown. Since the level of DOG1 protein showed no difference between the two seeds of *X. strumarium* in our study (unpubl. data), it seems that the DOG1 transcript level plays a fundamental role in dormancy regulation. A previous report also indicates for lack of correlation between the amount of DOG1 protein and seed dormancy in *A. thaliana* (Nakabayashi *et al.* 2012), and probably support the assumption that FY-mediated *DOG1* processing, which leads to producing two functional and non-functional isoforms of DOG1, has differentially taken place between the two seeds (Cyrek *et al.* 2016).

DOG1 acts as both ABA-independent (Graeber *et al.* 2014) and ABA-dependent (Née *et al.* 2017; Nishimura *et al.* 2018) manner to establish the seed dormancy. We observed a reduction in CRU2 transcripts abundance in the non-dormant seed, which interacted with *DOG1*. CRU2 has been known to be the most abundant stored protein in seeds of *A. thaliana* that functions in structural components of oil bodies. The higher CRU2 level in ABA-treated seeds of *A. thaliana* suggests that ABA contributes to the inhibition of CRU2 degradation (Ghelis *et al.* 2008). Although no

change has been observed in the accumulation of this protein in the seeds of *dog1–1 A. thaliana* mutants, DOG1 probably acts for the accumulation of N-rich compounds and amplification of ABA signalling (Dekkers *et al.* 2016). Therefore, our results confirm the likely effect of ABA signalling on high *CRU2* expression in the dormant seeds and also, at least in part, reflect a link between the *DOG1* function and ABA (Figs 9 and 10).

Dormancy control by *DOG1* through regulation of genes required for the biomechanical weakening of the coat encasing the embryo has previously been reported in *Lepidium sativum* (L.) (Graeber *et al.* 2014). Considering the higher abundance of callus biosynthesis-related transcripts and positive enrichment of the pathways related to cell wall organisation in dormant seed including *CALLOSE SYNTHASE*s and *GLUCAN SYNTHASE-LIKE*, we reasoned there are different cell wall properties between the dimorphic seeds of *X. strumarium*. There are two reasons why the callus deposition and following restriction of cell wall widening are important in our study: (1) determination of final seed size; and (2) suppression of radicle protrusion after imbibition. The importance of cell wall-associated processes of two dimorphic seeds became clear when the enrichment of the amylopectin biosynthetic process was detected for under-represented transcripts of the non-dormant seed, suggesting the formation of two seeds with different cell wall properties probably through *DOG1*-mediated processes may result in different dormancy status of *X. strumarium*.

Coordinated epigenetic modifications and their relevance to the dormancy

To date, many epigenetic modifications have been reported to regulate seed dormancy (reviewed in Klupczyńska and Pawłowski 2021). In this study, the importance of DNA methylation in the regulation of seed dormancy has been corroborated by the interaction of *DME* with *MAINTENANCE OF METHYLATION* (*MOM*) and *ARGONAUTE 4* (*AGO4*). *DME* plays an important role in gene imprinting in the endosperm, male fertility (Schoft *et al.* 2011), and seed viability through likely demethylation of the maternal allele of target genes (Choi *et al.* 2002; Morales-Ruiz *et al.* 2006). For example, it demethylates a group of Polycomb group regulators of seed development (Köhler *et al.* 2003). *AGO4* is a translation regulator which acts for gene silencing by RNA-directed DNA methylation through H3K9 methylation and affects the expression of dormancy genes (Qi *et al.* 2006; Nonogaki 2014). Although little information is available for silencing of seed dormancy genes through RNA-directed DNA methylation (RdDM), possible involvement of AGO4 in seed dormancy regulation has been suggested from studies of cereal seed dormancy. AGO1003, an *ARGONAUTE* gene in barley (*Hordeum vulgare* L.), is expressed differentially in the embryos of dormant and non-dormant seeds and is thought to function as a negative regulator of seed dormancy (Singh and Singh 2012). Moreover, mutations in *UBP26/SUP32* encoding a

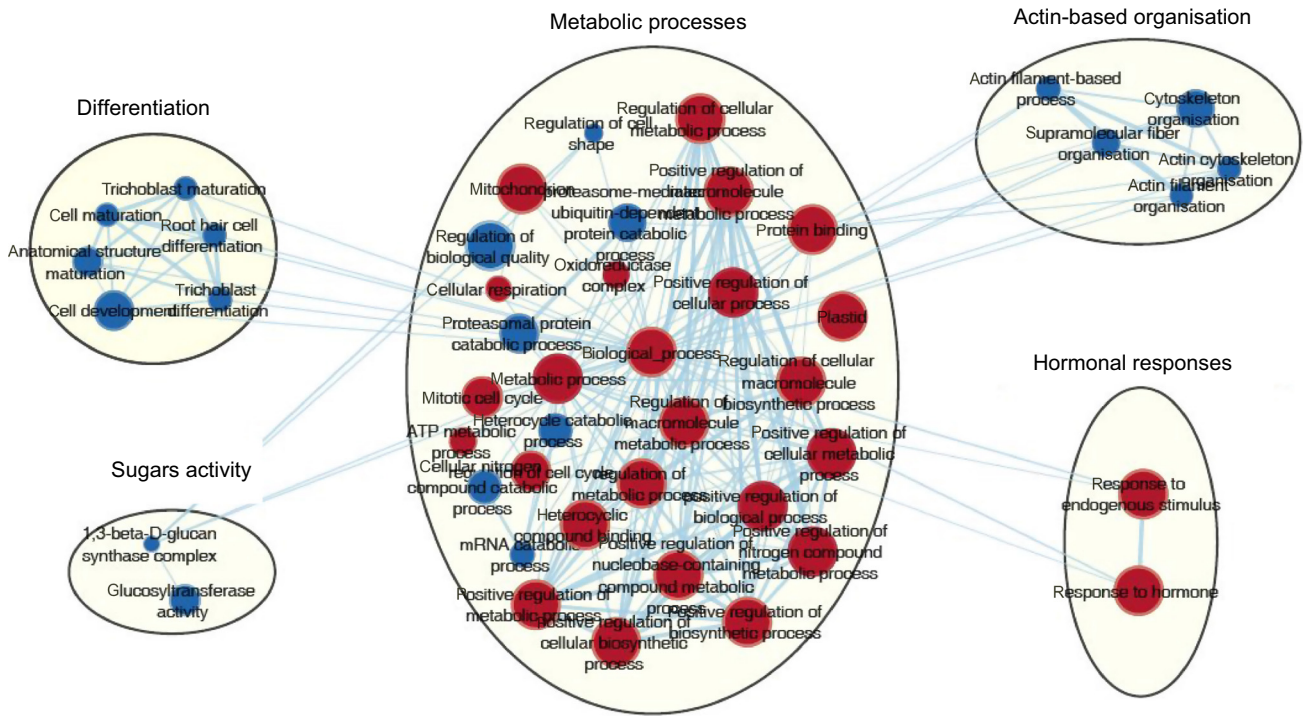


Fig. 9. Enrichment map of significantly ($P < 0.05$) enriched pathways for upregulated and downregulated genes of developing non-dormant and dormant seeds of *X. strumarium*. Blue and red circles show positively and negatively enriched pathways, respectively.

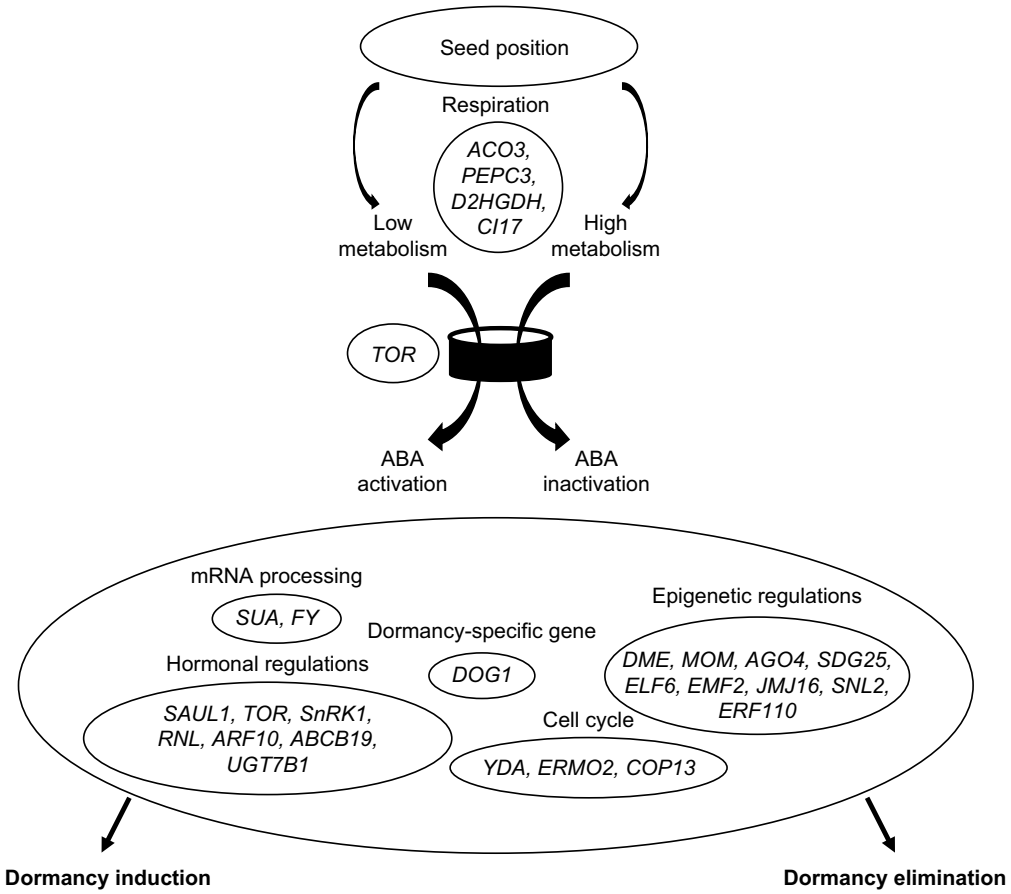


Fig. 10. Schematic representation of processes involved in dormancy induction to seeds of *X. strumarium*.

ubiquitin-specific protease executing H2B deubiquitination, cause a dramatic increase in monoubiquitinated H2B (Sridhar *et al.* 2007). Since the *ubp26* mutant shows enhanced seed dormancy, indicating H2B monoubiquitination may influence seed dormancy via histone H3K9 methylation and DNA methylation (Zheng *et al.* 2012). In plants, Polycomb group proteins (PcG), like EMF2, are known to repress seed master regulator and maturation genes during germination via chromatin modifications (Xu *et al.* 2018). Kim *et al.* (2010) reported the role of *EMFs* in dormancy genes silencing through H3K27me3. The interaction between *EMF2* and long noncoding RNAs not only results in H3K27me3 but also inactivates seed dormancy and maturation genes such as *DOG1* and *ABI3* (Chen *et al.* 2020).

Trithorax group (TrxG) proteins like *ATXR7/SDG25* activate gene expression through depositing active histone marks, H3K4me3 and H3K36me3 at target chromatin, which often overlap with PcG target chromatin, thus antagonising PcG function (Kim and Sung 2014). For example, loss of *ATXR7* function in H3K4 methylation, previously reported by Tamada *et al.* (2009) results in the elimination of seed dormancy in *Arabidopsis* (Liu *et al.* 2011). PcG silencing can be additionally counteracted by histone demethylases such as *ELF6* and *PKDM7D* that actively remove H3K27me3 and H3K4me2/3 marks, respectively (Qian *et al.* 2015), suggesting that over-expression of histone demethylation activity can also be correlated with the activation of germination genes during seed development. It has been confirmed that *SNL2* acts redundantly with *SNL1* for seed dormancy regulation. Increased acetylation of H3K9/18 and H3K14 was observed in the *snl1 snl2* double mutant (Wang *et al.* 2013). *SNL* suppresses transcription of ethylene biosynthesis (*ACOs*) and signalling (*ERFs*) genes, positive regulators of seed germination, and also ABA-inactivating genes, *CYP707A*s, via creating a transcription-inactive site on chromatin. This interaction may be related to the alteration in the ERF110 transcripts level in the non-dormant seed. This AP2 family member plays a mysterious role in dormancy induction and germination suppression through both GA suppression and ABA signalling regulation (Li *et al.* 2012). Taken together, these observations demonstrated that the chromatin remodelling functions, through the regulation of dormancy genes such as *DOG1* and *ABI3* as well as the differential expression of germination genes, result in dimorphic seeds of *X. strumarium*.

mRNA processing for ABA sensitivity and elongation control of DOG1

Our study revealed a sharp under-representation in some RNA splicing factors such as SUA, PWI, and FY in the non-dormant seed, which function in seed dormancy. *SUA* affects seed maturity through the regulation of ABI3 alternative splicing (Sugliani *et al.* 2010). Its activity results in two ABI3 transcript isoforms, *ABI3-α* and *ABI3-β* that encode full-length and truncated ABI3 proteins, respectively (Sugliani *et al.* 2010). Therefore, we hypothesised that low *SUA* expression in the non-dormant seed may result in the reduction of the full-length *ABI3* isoform level, which plays a critical role in dormancy regulation and ABA sensitivity. The role of PWI (RBM25) in RNA splicing and regulation of the ABA response in *Arabidopsis* mutants has previously been reported. Zhan *et al.* (2015) displayed a defect in alternative splicing of *PP2C* in *Arabidopsis* mutants of *rbm25*. The PP2Cs are negative regulators of the ABA signalling pathway through their interaction with SnRK2s (Rodrigues *et al.* 2013). This result indicates that variation in RNA splicing of genes like PP2C is of particular importance for *X. strumarium* seeds' response to ABA and that the splicing factor RBM25 has a critical role in this response. FY is also a component of the Cleavage and Polyadenylation Specificity Factor (CPSF) complex that acts to polyadenylation. Mutation in FY disrupts poly (A) signal recognition, which determines the cleavage site in the mRNA (Yu *et al.* 2019). Faster germination of freshly harvested seeds of *Arabidopsis fy-1* mutant while having a higher content of ABA suggests both ABA *fy-1* hyposensitivity and decreasing in the ABI5 level (Jiang *et al.* 2012). In addition, evidence is present in the literature that indicates FY is required for proper DOG1 protein expression and shows that the *fy-2* mutant is defective in DOG1 RNA processing, which leads to the suppression of translation (Cyrek *et al.* 2016). This finding led us to speculate that *FY* downregulation in the large seed of *X. strumarium* may be a potential regulatory mechanism to reduce the amount of shDOG1mRNA level in this seed, resulting in a weak dormancy level in the large seed compared to the small seed.

Hormonal regulatory circuitry in dimorphic seeds of *X. strumarium*

In this study, different expression levels of phytohormones biosynthesis and signalling genes such as *SAUL1*, *TOR*, *SnRK1*, *RNL*, *ARF10*, *ABCB19*, and *UGT74B1* point to the pivotal role of hormonal regulations of seed dormancy. *TOR*, which is a serine/threonine-protein kinase, is also over-represented in the non-dormant seed. In plants, *SnRK1* has been appeared to indirectly interact with *TOR* via regulatory-associated protein of TOR (RAPTOR) phosphorylation based on the cellular energy status (Broeckx *et al.* 2016). Despite the identified positive regulation of ABA biosynthesis, TOR interaction with ABA depends on the environmental growth conditions (Fu *et al.* 2020). When the plant is subjected to growth-promoting conditions, active TOR phosphorylates ABA receptors, PYR/PYLs, to inhibit ABA signalling and shift resources toward growth. On the contrary, under stressful conditions, RAPTOR will be phosphorylated by ABA-activated SnRK2s to suppress TOR activity and makes growth sacrifices for plant survival (Fu *et al.* 2020). Therefore, to identify the real function of TOR in two seeds of *X. strumarium*, cellular metabolism and carbohydrate availability should be considered. Over-representation of respiration-involving genes such as *ACO3*, *PEPC3*, *D2HGDH*, and *CI51* can be positively related to higher cellular metabolism and energy supply in the non-dormant seeds. This may affect the final seed weight or seed size and justify differential interactions between TOR and ABA in the two seeds of *X. strumarium*. Production of seeds with delayed germination has been reported in *raptor1b* mutant, which may be a result of higher ABA, auxin, and JA levels (Salem *et al.* 2017). Therefore, we theorise that seed size and germination in *X. strumarium* is influenced by respiratory factors, and TOR may be at the centre of a regulatory hub controlling seed metabolism, and germination through hormonal regulatory.

SAUL1 belongs to the E3 Plant U-box (PUB) family and plays important role in the regulation of cell death, developmental responses, and hormonal responses. The regulation of ABA biosynthesis through AAO3, SAUL1, and 26*S* proteasome degradation has previously been reported (Raab *et al.* 2009). AAO3 catalyses the abscisic aldehyde oxidation to ABA at the final step of ABA biosynthesis. Since *saul1* mutation affects ABA content, but not ABA signalling (Raab *et al.* 2009), SAUL1-mediated AAO3 degradation can be a mechanism for the ABA content and seed dormancy regulation. Therefore, over-expression of *SAUL1* in the non-dormant seed of *X. strumarium* may be related to the lower ABA content.

A reduced presence of *RNL* involved in tRNA splicing was also observed in the dormant seed. There is some evidence that auxin-related abnormalities associated with reduced abundance of PIN-formed (PIN) auxin transport proteins were observed after the downregulation of the RNA ligase RNL (Leitner *et al.* 2015). Furthermore, there is some evidence showing the role of ABCB19 and auxin distribution in cotyledon development. Analysis of gene expression pattern in *abcb19* revealed a significantly diminished auxin distribution as well as a growth rate reduction correlated with reduced auxin levels in cotyledons of *abcb19*, indicating that cotyledon expansion depends on ABCB19-mediated auxin import (Lewis *et al.* 2009). The role of UDP-glycosyltransferases in the modulation of auxin levels (Mateo-Bonmatí *et al.* 2021) and also *UGT74B1* in IAA homeostasis during *Arabidopsis* development (Grubb *et al.* 2004) has previously been revealed. Conjugation of IAA to low-weight molecules such

as UDP-glucose is an auxin inactivation mechanism (Grubb *et al.* 2004). We speculate the growth, and germination differences between two seeds of *X. strumarium* may have been linked to alterations in ABA content and auxin distribution and responses. While ABA produced by the embryo is fundamental for the promotion of seed dormancy (Sohindji *et al.* 2020), auxin transport defines local variations in hormone levels, which are perceived and transmitted to induce hormonally controlled adjustments in gene expression and activity (Leitner *et al.* 2015). In addition, an increase in CSN3/FUS11 transcripts was also observed in the large seeds that displayed significantly more germinability. Germination assays using wild-type and mutant seedlings of *csn5a-2* showed that removal of CSN function displayed deeper dormancy (Franciosini *et al.* 2015). The degradation of ABA receptors such as RGL2 and ABI5 by CULLIN4-RING E3 ubiquitin ligases (CRL4s) seems to be a mechanism by which CSN promotes seed germination (Irigoyen *et al.* 2014; Jin *et al.* 2018). It has also been reported that the *csn3–3* mutation confers several phenotypes indicative of impaired auxin signalling including auxin resistant root growth and diminished auxin-responsive gene expression (Huang *et al.* 2013), suggesting CSN3 functions in a distinct protein complex that is required for proper auxin signalling. All together, this indicates degradation of ABA biosynthesis and/or proper auxin signalling in non-dormant seeds of *X. strumarium* may control germinability.

Control of proliferation, and endoreduplication in two seeds of *X. strumarium*

An over-represented factor in non-dormant seed was *YODA/YDA*, which encodes a ubiquitously expressed MAPKK Kinase and functions in the regulation of asymmetric divisions in the zygote and the formation of normal suspensor independently from auxin (Lukowitz *et al.* 2004). In loss-of-function mutants, the zygote does not elongate properly, and the cells of the basal lineage are eventually incorporated into the embryo instead of differentiating the extra-embryonic suspensor (Lukowitz *et al.* 2004). Therefore, less *YDA* expression in the dormant seed may indicate that there is some difference in the embryonic patterning processes between two dimorphic seeds. We also revealed an over-representation of both *ERMO2/SEC24a* and *DEL1/E2Fe* in the non-dormant seed. *ERMO2* encodes the COPII protein Sec24a that participates in ER-to-Golgi anterograde transport and acts in setting up cell proliferation, cell size determining via endoreduplication suppression, and mediating the membrane fusion (Ebine *et al.* 2008). Moreover, *Arabidopsis* DEL1 encodes an E2F-DP-like DNA binding protein that was previously shown to be mostly expressed in dividing cells and to inhibit endoreduplication (Vlieghe *et al.* 2005). We suggest that the suppression of endoreduplication in the small seed of *X. strumarium* may be a mechanism for cell differentiation and cell size determination. However, future investigations should compare this result through microscopic and molecular studies with a focus on the early developmental stage.

Conclusion

Our research is the first study on the molecular mechanisms underlying seed dormancy in dimorphic seeds of *X. strumarium*. It confirmed that a compilation of different mechanisms (dormancy-specific genes activation, epigenetic and hormonal regulations, cell wall modifications, and mRNA processing) were differentially activated between two seeds. We revealed the positive regulation of seed dormancy by *DOG1*, the main dormancy gene in *Arabidopsis* leading to the formation of two seeds with different cell wall properties. This study also demonstrated that the chromatin remodelling functions in the regulation of dormancy genes such as *DOG1* as well as activation of germination-associated genes through DNA methylation and histone demethylation, respectively. We also indicated that variation in RNA splicing for ABA sensitivity and transcriptional elongation control of *DOG1* is of particular importance for *X. strumarium* seeds. Our results also showed that seed size and germination may be influenced by respiratory factors, and alterations in ABA content and auxin distribution and responses. TOR probably is at the centre of a regulatory hub controlling seed metabolism, maturation, and germination. We also theorised that the suppression of endoreduplication in the small seed of *X. strumarium* may be a mechanism for cell differentiation and cell size determination. However, future investigations should compare these results through molecular studies with a focus on the different developmental stages.

References

Arc E, Sechet J, Corbineau F, Rajjou L, Marion-Poll A (2013) ABA crosstalk with ethylene and nitric oxide in seed dormancy and germination. *Frontiers in Plant Science* **4**, 63. doi:10.3389/fpls.2013.00063

Baena-González E, Rolland F, Thevelein JM, Sheen J (2007) A central integrator of transcription networks in plant stress and energy signalling. *Nature* **448**, 938–942. doi:10.1038/nature06069

Bentsink L, Jowett J, Hanhart CJ, Koornneef M (2006) Cloning of *DOG1*, a quantitative trait locus controlling seed dormancy in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **103**(45), 17042–17047. doi:10.1073/pnas.0607877103

Broeckx T, Hulsmans S, Rolland F (2016) The plant energy sensor: evolutionary conservation and divergence of SnRK1 structure, regulation, and function. *Journal of Experimental Botany* **67**(22), 6215–6252. doi:10.1093/jxb/erw416

Chen H, Tong J, Fu W, Liang Z, Ruan J, Yu Y, *et al.* (2020) The H3K27me3 demethylase RELATIVE OF EARLY FLOWERING6 suppresses seed dormancy by inducing abscisic acid catabolism. *Plant Physiology* **184**(4), 1969–1978. doi:10.1104/pp.20.01255

Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, *et al.* (2002) DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell* **110**, 33–42. doi:10.1016/S0092-8674(02)00807-3

Cyrek M, Fedak H, Ciesielski A, Guo Y, Sliwa A, Brzezniak L, *et al.* (2016) Seed dormancy in Arabidopsis is controlled by alternative polyadenylation of *DOG1*. *Plant Physiology* **170**, 947–955. doi:10.1104/pp.15.01483

Dekkers BJW, He H, Hanson J, Willems LAJ, Jamar DCL, Cueff G, *et al.* (2016) The Arabidopsis *DELAY Of GERMINATION 1* gene affects *ABSCISIC ACID INSENSITIVE 5* (*ABI5*) expression and genetically interacts with *ABI3* during Arabidopsis seed development. *The Plant Journal* **85**(4), 451–465. doi:10.1111/tpj.13118

Ebine K, Okatani Y, Uemura T, Goh T, Shoda K, Niihama M, *et al.* (2008) A SNARE complex unique to seed plants is required for protein storage vacuole biogenesis and seed development of *Arabidopsis thaliana*. *The Plant Cell* **20**, 3006–3021. doi:10.1105/tpc.107.057711

Ferraz dos Santos L, Santana Silva RJ, Oliveira Jordão do Amaral D, Barbosa de Paula MF, Falcão Ludke L, Legavre T, *et al.* (2016) Selection of reference genes for expression study in pulp and seeds of *Theobroma Grandiflorum* (Willd. Ex Spreng.) Schum. *PLoS ONE* **11**, e0160646. doi:10.1371/journal.pone.0160646

Franciosini A, Moubayidin L, Du K, Matari NH, Boccaccini A, Butera S, *et al.* (2015) The COP9 SIGNALOSOME is required for postembryonic meristem maintenance in *Arabidopsis thaliana*. *Molecular Plant* **8**, 1623–1634. doi:10.1016/j.molp.2015.08.003

Fu L, Wang P, Xiong Y (2020) Target of rapamycin signaling in plant stress responses. *Plant Physiology* **182**(4), 1613–1623. doi:10.1104/PP.19.01214

Ghelis T, Bolbach G, Clodic G, Habricot Y, Miginiac E, Sotta B, Jeannette E (2008) Protein tyrosine kinases and protein tyrosine phosphatases are involved in abscisic acid-dependent processes in Arabidopsis seeds and suspension cells. *Plant Physiology* **148**(3), 1668–1680. doi:10.1104/pp.108.124594

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**(7), 644–652. doi:10.1038/nbt.1883

Graeber K, Nakabayashi K, Miatton E, Leubner-Metzger G, Soppe WJJ (2012) Molecular mechanisms of seed dormancy. *Plant, Cell & Environment* **35**(10), 1769–1786. doi:10.1111/j.1365-3040.2012.02542.x

Graeber K, Linkies A, Steinbrecher T, Mummenhoff K, Tarkowská D, Turečková V, *et al.* (2014) *DELAY OF GERMINATION 1* mediates a conserved coat-dormancy mechanism for the temperature- and gibberellin-dependent control of seed germination. *Proceedings of the National Academy of Sciences of the United States of America* **111**(34), E3571–E3580. doi:10.1073/pnas.1403851111

Grubb CD, Zipp BJ, Ludwig-Muller J, Masuno MN, Molinski TF, Abel S (2004) Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. *The Plant Journal* **40**, 893–908. doi:10.1111/j.1365-313X.2004.02261.x

Heyman J, Canher B, Bisht A, Christiaens F, De Veylder L (2018) Emerging role of the plant erf transcription factors in coordinating wound defense responses and repair. *Journal of Cell Science* **131**, jcs208215. doi:10.1242/jcs.208215

Hu Y, Yu D (2014) BRASSINOSTEROID INSENSITIVE2 interacts with ABSCISIC ACID INSENSITIVE5 to mediate the antagonism of brassinosteroids to abscisic acid during seed germination in Arabidopsis. *The Plant Cell* **26**(11), 4394–4408. doi:10.1105/tpc.114.130849

Huang H, Quint M, Gray WM (2013) The *eta7/csn3-3* auxin response mutant of Arabidopsis defines a novel function for the CSN3 subunit of the COP9 signalosome. *PLoS ONE* **8**(6), e66578. doi:10.1371/journal.pone.0066578

Huo H, Wei S, Bradford KJ (2016) *DELAY OF GERMINATION1* (*DOG1*) regulates both seed dormancy and flowering time through microRNA pathways. *Proceedings of the National Academy of Sciences of the United States of America* **113**(15), E2199–E2206. doi:10.1073/pnas.1600558113

Irigoyen ML, Iniesto E, Rodriguez L, Puga MI, Yanagawa Y, Pick E, *et al.* (2014) Targeted degradation of abscisic acid receptors is mediated by the ubiquitin ligase substrate adaptor DDA1 in *Arabidopsis*. *The Plant Cell* **26**, 712–728. doi:10.1105/tpc.113.122234

Jaiswal PS, Kaur N, Randhawa GS (2019) Identification of reference genes for QRT-PCR gene expression studies during seed development and under abiotic stresses in *Cyamopsis tetragonoloba*. *Crop Science* **59**, 252–265. doi:10.2135/cropsci2018.05.0313

Jiang S, Kumar S, Eu Y-J, Jami SK, Stasolla C, Hill RD (2012) The *Arabidopsis* mutant, *fy-1*, has an ABA-insensitive germination phenotype. *Journal of Experimental Botany* **63**(7), 2693–2703. doi:10.1093/jxb/err452

Jin D, Wu M, Li B, Bucker B, Keli P, Zhang S, *et al.* (2018) The COP9 Signalosome regulates seed germination by facilitating protein degradation of RGL2 and ABI5. *PLOS Genetics* **14**, e1007237. doi:10.1371/journal.pgen.1007237

Katoh H, Esashi Y (1975) Dormancy and impotency of cocklebur seeds I. CO_2, $C_2H_4O_2$, and high temperature. *Plant and Cell Physiology* **16**, 687–696. doi:10.1093/oxfordjournals.pcp.a075189

Kim DH, Sung S (2014) Polycomb-mediated gene silencing in *Arabidopsis thaliana*. *Molecular Cells* **37**, 841–850. doi:10.14348/molcells.2014.0249

Kim SY, Zhu T, Renee Sung Z (2010) Epigenetic regulation of gene programs by EMF1 and EMF2 in Arabidopsis. *Plant Physiology* **152**, 516–528. doi:10.1104/pp.109.143495

Kim S, Park JS, Lee J, Lee KK, Park OS, Choi HS, Seo PJ, Cho HT, Frost JM, Fischer RL, Choi Y (2021) The DME demethylase regulates sporophyte gene expression, cell proliferation, differentiation, and meristem resurrection. *Proceedings of the National Academy of Sciences of the United States of America* **118**(29), e2026806118. doi:10.1073/pnas.2026806118

Klupczyńska EA, Pawłowski TA (2021) Regulation of seed dormancy and germination mechanisms in a changing environment. *International Journal of Molecular Sciences* **22**, 1357. doi:10.3390/ijms22031357

Köhler G, Hennig L, Spillane C, Pien S, Gruissem W, Grussiklaus U (2003) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. *Genes and Development* **17**, 1540–1553. doi:10.1101/gad.257403

Köster T, Marondedze C, Meyer K, Staiger D (2017) RNA-binding proteins revisited – the emerging Arabidopsis MRNA interactome. *Trends in Plant Science* **22**, 512–526. doi:10.1016/j.tplants.2017.03.009

Leitner J, Retzer K, Malenica N, Bartkeviciute R, Lucyshyn D, Jäger G, *et al.* (2015) Meta-regulation of Arabidopsis auxin responses depends on TRNA maturation. *Cell Reports* **11**, 516–526. doi:10.1016/j.celrep.2015.03.054

Lewis DR, Wu G, Ljung K, Spalding EP (2009) Auxin Transport into Cotyledons and Cotyledon Growth Depend Similarly on the ABCB19 Multidrug Resistance-like Transporter. *The Plant Journal* **60**, 91–101. doi:10.1111/j.1365-313X.2009.03941.x

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659. doi:10.1093/bioinformatics/btl158

Li Y, Shu Y, Peng C, Zhu L, Guo G, Li N (2012) Absolute quantitation of isoforms of post-translationally modified proteins in transgenic organism. *Molecular & Cellular Proteomics* **11**(8), 272–285. doi:10.1074/mcp.M111.016568

Liu Y, Schmidt B (2012) Long read alignment based on maximal exact match seeds. *Bioinformatics* **28**, i318–i324. doi:10.1093/bioinformatics/bts414

Liu F, Guo DD, Tu YH, Xue YR, Gao Y, Guo ML (2016) Identification of reference genes for gene expression normalization in safflower (*Carthamus tinctorius*). *Revista Brasileira de Farmacognosia* **26**, 564–570. doi:10.1016/j.bjp.2016.05.006

Liu Y, Geyer R, van Zanten M, Carles A, Li Y, Hörold A, van Nocker S, Soppe WJJ (2011) Identification of the *Arabidopsis REDUCED DORMANCY 2* gene uncovers a role for the polymerase associated factor 1 complex in seed dormancy. *PLoS ONE* **6**(7), e22241. doi:10.1371/journal.pone.0022241

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2-\Delta\Delta CT$ method. *Methods* **25**, 402–408. doi:10.1006/meth.2001.1262

Lukowitz W, Roeder A, Parmenter D, Somerville V (2004) A MAPKK kinase gene regulates extra-embryonic cell fate in Arabidopsis. *Cell* **116**, 109–119. doi:10.1016/S0092-8674(03)01067-5

Mateo-Bonmatí E, Casanova-Sáez R, Šimura J, Ljung K (2021) Broadening the roles of UDP-glycosyltransferases in auxin homeostasis and plant development. *New Phytologist* **232**, 642–654. doi:10.1111/nph.17633

Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marín MI, Martínez-Macías MI, Ariza RR, Roldán-Arjona T (2006) *DEMETER* and *REPRESSOR OF SILENCING 1* encode 5-methylcytosine DNA

glycosylases. *Proceedings of the National Academy of Sciences of the United States of America* **103**(18), 6853–6858. doi:10.1073/pnas.0601109103

Nakabayashi K, Bartsch M, Xiang Y, Miatton E, Pellengahr S, Yano R, *et al.* (2012) The time required for dormancy release in *Arabidopsis* is determined by DELAY OF GERMINATION1 protein levels in freshly harvested seeds. *The Plant Cell* **24**(7), 2826–2838. doi:10.1105/tpc.112.100214

Née G, Kramer K, Nakabayashi K, Yuan B, Xiang Y, Miatton E, *et al.* (2017) DELAY Of GERMINATION1 requires PP2C phosphatases of the ABA signalling pathway to control seed dormancy. *Nature Communications* **8**(1), 72. doi:10.1038/s41467-017-00113-6

Nishimura N, Tsuchiya W, Moresco JJ, Hayashi Y, Satoh K, Kaiwa N, *et al.* (2018) Control of seed dormancy and germination by DOG1-AHG1 PP2C phosphatase complex *via* binding to heme. *Nature Communications* **9**, 2132. doi:10.1038/s41467-018-04437-9

Nonogaki H (2014) Seed dormancy and germination—emerging mechanisms and new hypotheses. *Frontiers in Plant Science* **5**, 233. doi:10.3389/fpls.2014.00233

Penfield S, MacGregor DR (2017) Effects of environmental variation during seed production on seed dormancy and germination. *Journal of Experimental Botany* **68**(4), 819–825. doi:10.1093/jxb/erw436

Pereira WJ, Bassinello PZ, Brondani C, Vianello RP (2017) An improved method for RNA extraction from common bean seeds and validation of reference genes for qPCR. *Crop Breeding and Applied Biotechnology* **17**, 150–158. doi:10.1590/1984-70332017v17n2a22

Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ (2006) Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* **443**(7), 1008–1012. doi:10.1038/nature05198

Qian S, Wang Y, Ma H, Zhang L (2015) Expansion and functional divergence of Jumonji C-containing histone demethylases: significance of duplications in ancestral angiosperms and vertebrates. *Plant Physiology* **168**(4), 1321–1337. doi:10.1104/pp.15.00520

Raab S, Drechsel G, Zarepour M, Hartung W, Koshiba T, Bittner F, Hoth S (2009) Identification of a novel E3 ubiquitin ligase that is required for suppression of premature senescence in Arabidopsis. *The Plant Journal* **59**, 39–51. doi:10.1111/j.1365-313X.2009.03846.x

Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, *et al.* (2019) Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols* **14**, 482–517. doi:10.1038/s41596-018-0103-9

Rodrigues A, Adamo M, Crozet P, Margalha L, Confraria A, Martinho C, *et al.* (2013) ABI1 and PP2CA phosphatases are negative regulators of Snf1-related protein Kinase1 signaling in *Arabidopsis*. *The Plant Cell* **25**, 3871–3884. doi:10.1105/tpc.113.114066

Rubio de Casas R, Willis CG, Pearse WD, Baskin CC, Baskin JM, Cavender-Bares J (2017) Global biogeography of seed dormancy is determined by seasonality and seed size: a case study in the legumes. *New Phytologist* **214**, 1527–1536. doi:10.1111/nph.14498

Salem MA, Li Y, Wiszniewski A, Giavalisco P (2017) Regulatory-associated protein of TOR (RAPTOR) alters the hormonal and metabolic composition of Arabidopsis seeds, controlling seed morphology, viability and germination potential. *The Plant Journal* **92**(4), 525–545. doi:10.1111/tpj.13667

Schoft VK, Chumak N, Choi Y, Hannon M, Garcia-Aguilar M, Machlicova A, *et al.* (2011) Function of the DEMETER DNA glycosylase in the *Arabidopsis thaliana* male gametophyte. *Proceedings of the National Academy of Sciences of the United States of America* **108**(19), 8042–8047. doi:10.1073/pnas.1105117108

Shu K, Liu X-d, Xie Q, He Z-h (2016) Two faces of one seed: hormonal regulation of dormancy and germination. *Molecular Plant* **9**(1), 34–45. doi:10.1016/j.molp.2015.08.010

Singh M, Singh J (2012) Seed development-related expression of ARGONAUTE4_9 class of genes in barley: possible role in seed dormancy. *Euphytica* **188**, 123–129. doi:10.1007/s10681-012-0624-1

Sohindji FS, Sogbohossou DEO, Zohoungbogbo HPF, Houdegbe CA, Achigan-Dako EG (2020) Understanding molecular mechanisms of seed dormancy for improved germination in traditional leafy vegetables: an overview. *Agronomy* **10**, 57. doi:10.3390/agronomy10010057

Sridhar VV, Kapoor A, Zhang K, Zhu J, Zhou T, Hasegawa PM, *et al.* (2007) Control of DNA methylation and heterochromatic silencing by histone H2B deubiquitination. *Nature* **447**, 735–738. doi:10.1038/nature05864

Sugliani M, Brambilla V, Clerkx EJM, Koornneef M, Soppe WJJ (2010) The conserved splicing factor SUA controls alternative splicing of the developmental regulator *ABI3* in *Arabidopsis*. *The Plant Cell* **22**(6), 1936–1946. doi:10.1105/tpc.110.074674

Tamada Y, Yun JY, Woo Sc, Amasino RM (2009) *ARABIDOPSIS TRITHORAX-RELATED7* is required for methylation of lysine 4 of histone H3 and for transcriptional activation of *FLOWERING LOCUS C*. *The Plant Cell* **21**(10), 3257–3269. doi:10.1105/tpc.109.070060

Vaishnav K, George LB, Highland HN (2015) Antitumour activity of *Xanthium strumarium* L. on human cervical cancer HeLa cells. *Journal of Cancer and Tumor International* **2**(1), 1–13. doi:10.9734/JCTI/2015/15519

Vaistij FE, Gan Y, Penfield S, Gilday AD, Dave A, He Z, *et al.* (2013) Differential control of seed primary dormancy in *Arabidopsis* ecotypes by the transcription factor SPATULA. *Proceedings of the National Academy of Sciences of the United States of America* **110**(26), 10866–10871. doi:10.1073/pnas.1301647110

Vlieghe K, Boudolf V, Beemster GTS, Maes S, Magyar Z, Atanassova A, *et al.* (2005) The DP-E2F-like gene DEL1 controls the endocycle in *Arabidopsis Thaliana*. *Current Biology* **15**, 59–63. doi:10.1016/j.cub.2004.12.038

Wang Z, Cao H, Sun Y, Li X, Chen F, Carles A, *et al.* (2013) *Arabidopsis* paired amphipathic helix proteins SNL1 and SNL2 redundantly regulate primary seed dormancy *via* abscisic acid–ethylene antagonism mediated by histone deacetylation. *The Plant Cell* **25**(1), 149–166. doi:10.1105/tpc.112.108191

Wei L, Miao H, Zhao R, Han X, Zhang T, Zhang H (2013) Identification and testing of reference genes for sesame gene expression analysis by quantitative real-time PCR. *Planta* **237**, 873–889. doi:10.1007/s00425-012-1805-9

Weiste C, Dröge-Laser W (2014) The *Arabidopsis* transcription factor bZIP11 activates auxin-mediated transcription by recruiting the histone acetylation machinery. *Nature Communications* **5**, 3883. doi:10.1038/ncomms4883

Xu F, Kuo T, Rosli Y, Liu MS, Wu L, Chen L-FO, *et al.* (2018) Trithorax group proteins act together with a polycomb group protein to maintain chromatin integrity for epigenetic silencing during seed germination in Arabidopsis. *Molecular Plant* **11**, 659–677. doi:10.1016/j.molp.2018.01.010

Yu Z, Lin J, Li QQ (2019) Transcriptome analyses of FY mutants reveal its role in mRNA alternative polyadenylation. *The Plant Cell* **31**(10), 2332–2352. doi:10.1105/tpc.18.00545

Zhan X, Qian B, Cao F, Wu W, Yang L, Guan Q, *et al.* (2015) An *Arabidopsis* PWI and RRM motif-containing protein is critical for pre-mRNA splicing and ABA responses. *Nature Communications* **6**, 8139. doi:10.1038/ncomms9139

Zheng J, Chen F, Wang Z, Cao H, Li X, Deng X, *et al.* (2012) A novel role for histone methyltransferase KYP/SUVH4 in the control of *Arabidopsis* primary seed dormancy. *New Phytologist* **193**, 605–616. doi:10.1111/j.1469-8137.2011.03969.x

Zhou W, Wang S, Yang L, Sun Y, Zhang Q, Li B, *et al.* (2019) Reference genes for QRT-PCR normalisation in different tissues, developmental stages, and stress conditions of hypericum perforatum. *PeerJ* **7**, e7133. doi:10.7717/peerj.6218

Data availability. The datasets generated for this study can be accessed from the National Centre for Biotechnology Information (NCBI) under accession number PRJNA809034.

Conflicts of interest statement. The authors declare no conflicts of interest.

Declaration of funding. This study was funded by the Agricultural Biotechnology Research Institute of Iran (ABRII) under the transcriptomic project's fund.

Ethics approval and consent to participate. This study complies with relevant institutional, national, and international guidelines and legislation for using plant material.

Acknowledgements. We thank Dr. Reza Ghaffari and Dr. Javad Gharechahi for invaluable discussions and feedback that help shape the content of this work.

Author contributions. G.HS. and I.N. conceptualised the study and worked on the methodology. I.N. and S.GH performed the investigation and prepared samples. I.N. designed and completed the bioinformatics analysis and wrote the manuscript. G.HS. and M.S reviewed and commented on the manuscript. All authors read and approved the final version of the manuscript.

Author affiliations

[A]Department of Plant Production and Genetics Engineering, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran.

[B]Department of Molecular Sciences, Macquarie University, Sydney, NSW 2109, Australia.

[C]Department of Systems and Synthetic Biology, Agricultural Biotechnology Research Institute of Iran (ABRII), Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran.

[D]Department of Agronomy, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.

[E]Agricultural and Natural Resources College, University of Tehran, Tehran, Iran.

[F]Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz, Iran.