Dorone & Boeynaems *et al.* (2020) Hydration-dependent phase separation of a prion-like protein regulates seed germination during water stress.

Supplemental Information:

STAR Methods
Figs. S1 to S7
References

STAR Methods

Resource Availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Seung Rhee (**srhee@carnegiescience.edu).**

Materials Availability

All unique materials (seeds, plasmids and strains) generated in this study will be made available on request from the Lead Contact upon completion of a Materials Transfer Agreement.

Data and Code Availability

The accession numbers for the RNA sequencing data reported in this paper are GEO: GSE167280 and GEO: GSE167291.
This study generated datasets provided in Tables S1 and S2.

Experimental Model and Subject Details

Plant growth conditions

Arabidopsis thaliana plants from which seeds were harvested for the experimental assays (except for those used in the seed aging experiment) were grown in soil (PRO-MIX® HP Mycorrhizae) inside growth cabinets (Percival) held at 22 ˚C and 55 % humidity with a 16/8 hour photoperiod (130 µmol.m^{-2}.s^{-1}). Seeds were stratified for 3 days at 4 ˚C in darkness to break dormancy. Plants from each genotype were randomly distributed and rotated every day until bolting to minimize environmental variations. When siliques began to mature, humidity was decreased to 45% as recommended by the Arabidopsis Biological Resource Center (ABRC) (ftp://ftp.arabidopsis.org/ABRC/abrc_plant_growth.pdf). Harvested seeds were air-dried for a week before being stored in Eppendorf tubes at 4 ˚C.

Nicotiana benthamiana (tobacco) plants were grown in soil (PRO-MIX® PGX) inside chambers held at 22 ˚C with a 16/8 hour photoperiod (130 µmol.m^{-2}.s^{-1}).

In this study, "MS medium" refers to Murashige and Skoog medium with the following recipe: 0.5X Murashige and Skoog basal salt mixture (PhytoTechnologies Laboratories) at a pH of 5.7 supplemented with 0.8 % agar (Difco) and 1 % sucrose (Sigma-Aldrich)).

Plant material

floe1-1 T-DNA mutant:

The mutant line *floe1-1* (SALK_048257C) was obtained from ABRC and genotyped using primers priFLOE1cds-FWD/REV and the Salk genotyping primer LBb1.3 (Table S3). It was confirmed to be a knockout mutant by RT-qPCR (Fig. S6A) as described in the *RT-qPCR analyses* section.

Transgenic Lines:

Transgenic plants were generated by *Agrobacterium tumefaciens*-mediated (GV3101 strain) transformation (Clough, 2005) of *floe1-1* with the constructs described in the *Plant plasmid construction* section, with the exception of the control transgenic line overexpressing YFP-FLAG used in Fig. S1C that was generated by introducing the transgene into Col-0. Transgenic seedlings (T_1) were selected with Basta and lines containing only one T-DNA construct were selected for further characterization by determining the Mendelian segregation ratio (3:1) of Basta-resistant seedlings in their T_2 progeny. Homozygote T_2 lines were then identified by verifying that T_3 seedlings were 100% Basta-resistant.

CRISPR lines:

FLOE1 CRISPR lines were generated using the *Staphylococcus aureus* CRISPR-Cas9 system (Steinert et al., 2015) and by following the protocol described in (https://www.botanik.kit.edu/molbio/940.php). A region within the QPS-rich region was identified as having an NNGGGT protospacer adjacent motif (PAM) downstream of a protospacer sequence (5' TTACAGCCCCCAGACTGGC 3') that did not have any significant similarities to other genomic regions. The corresponding guide RNA was inserted in the BbsI site of the pEn-Sa-Chimera vector through digestion-ligation following hybridization of the oligo duplex priCRISPR-FWD/REV (Table S3), to generate pEn-Sa-Chimera-FLOE1. The resulting sgRNA coding vector was then transferred to pDe-Sa-Cas9 through LR (Thermo Fisher Scientific) recombination, to generate pDe-Sa-Cas9-FLOE1. The final binary destination vector was then used to transform *Agrobacterium* (GV3101 strain), which was used to transform Col-0 plants

using the floral dip method (Clough, 2005). Seeds obtained from the T_0 parental lines were sown on MS medium supplemented with 30 mg/L Kanamycin (G-Biosciences) for selection of successfully transformed transgenics. Selected T_1 seedlings were then transferred to soil to mature. Genomic DNA was extracted from mature rosette leaves of each of these T_1 plants and the Cas9-recognition site within *FLOE1* was amplified through PCR with Phusion DNA polymerase (Thermo Fisher Scientific) using primers prigenoCRISPR-FWD/REV (Table S3). Sequencing (Sequetech Inc.) of the amplicons revealed that twelve plants demonstrated heterogenous sequences at the targeted region, which were subsequently selected for growing the T_2 generation. For each selected T_1 plant, eight T_2 progeny were grown, and PCR amplification followed by sequencing of the *FLOE1* amplicon was again performed on genomic DNA extracted from mature rosette leaves. Four individuals from the various T_2 populations (*floe1-2*, *floe1-3*, *floe1-4*, *floe1-5*) presented different homozygous mutations in the *FLOE1* amplicon (Table S3), leading to frameshift mutations and pre-mature stop codons in the QPS region, and were selected for further assays.

Method Details

Identification and analysis of the seed proteome

Arabidopsis thaliana genes were scored via the Expression Angler tool based on similarity to a "Developmental Map" expression pattern with "High Relative Expression" in "Dry Seed" and "Low Relative Expression" for all other tissues (https://bar.utoronto.ca/ExpressionAngler/) (Austin et al., 2016). The output was then normalized to Z-scores (Table S1) and genes were considered as seed-specific if they had a Z score of 3 or higher. The MobiDB-lite disorder scores of each gene in the "Z > 3" and "Z < 3" groups were retrieved from the MobiDB (version 3.1) *A. thaliana* dataset (https://mobidb.bio.unipd.it/dataset) (Piovesan et al., 2018), and their amino acid compositional profiles were obtained using the protr package (version 1.6-2) (Xiao et al., 2015) in RStudio (version 1.2.5033). Genes in the "Z > 3" group were then checked for the presence of a predicted prion-like domain based on a list of Arabidopsis prion-like proteins (Chakrabortee et al., 2016). For FLOE1's disorder prediction per residue, we used PONDR VSL2 (http://www.pondr.com) (Xue et al., 2010) and for identifying its prion-like domain we used PLAAC (http://plaac.wi.mit.edu) (Lancaster et al., 2014).

Plant plasmid construction

In our study, we used the following naming convention: pPLASMID; PROMOTERp; *GENE*; *mutant*; and PROTEIN. Constructs were generated using the Gateway system (Thermo Fisher

Scientific) and vectors from the pGWB601-661 collection (https://shimane-u.org/nakagawa/pgwb-tables/4.htm) (Nakagawa et al., 2007) as follows:

Transgenes for Arabidopsis experiments:

FLOE1's genomic region from its promoter, as predicted by AGRIS (Davuluri et al., 2003), to its last coding codon was amplified by Phusion (Thermo Fisher Scientific) PCR from Col-0 DNA (extracted with DNeasy Plant Mini Kit (Qiagen)) using the prigFLOE1-FWD/REV primers (Table S3). The amplicon was first cloned into pDONR221 (Thermo Fisher Scientific) using BP Clonase II (Thermo Fisher Scientific) and then subcloned into pGWB604, pGWB610, and pGWB633 (Nakagawa et al., 2007) using LR Clonase II (Thermo Fisher Scientific) to generate pFLOE1p:FLOE1-GFP, pFLOE1p:FLOE1-FLAG, and pFLOE1p:FLOE1-GUS respectively.

pFLOE1p:FLOE1ΔDS-GFP, pFLOE1p:FLOE1ΔQPS-GFP, and pFLOE1p:FLOE1ΔDUF-GFP were obtained by modifying pFLOE1p:FLOE1-GFP using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) with primers priDSdeletion-FWD/REV, priQPSdeletion-FWD/REV, and priDUFdeletion-FWD/REV (Table S3) respectively.

An entry vector containing the *YFP* gene was donated by Dr. Zhiyong Wang (Carnegie Institution for Science, USA) and another one, G18395, containing *FLOE1*'s coding sequence was obtained from ABRC. The two genes were then transferred from the entry vector into the binary vector pB7HFC3_0 (Bossi et al., 2017) using Gateway cloning (Life Technologies), to create the vectors p35S:YFP-FLAG and p35S:FLOE1-FLAG.

Transgenes for tobacco (*Nicotiana benthamiana*) experiments:

Arabidopsis genes: The coding sequences of *FLOE1*'s isoforms, *FLOE1.1* and *FLOE1.2*, were amplified by Phusion (Thermo Fisher Scientific) PCR from the entry vector G18395 using priFLOE1.1-FWD/REV and priFLOE1.2-FWD/REV (Table S3) and then BP recombined into pDONR221 (Thermo Fisher Scientific) to generate pDONR221-FLOE1.1 and pDONR221-FLOE1.2. These were then transferred by LR recombination into pGWB605 (Nakagawa et al., 2007) to generate p35S:FLOE1.1-GFP and p35S:FLOE1.2-GFP. Similarly, p35S:FLOE1.1-RFP and p35S:FLOE1.2-RFP were generated by subcloning FLOE1.1 and FLOE1.2 into pGWB660 (Nakagawa et al., 2007). The N-terminal version p35S:GFP-FLOE1.1 was generated by LR recombination of G18395 into pGWB606 (Nakagawa et al., 2007). To generate p35S:FLOE2-GFP and p35S:FLOE3-GFP, the coding sequences of *FLOE2* and *FLOE3* were obtained from a cDNA library from 5-day old Col-0 seedlings by PCR amplification using Phusion DNA polymerase

(Thermo Fisher Scientific) and the primers priFLOE2-FWD/REV and priFLOE3-FWD/REV (Table S3). Total cDNA was obtained by reverse transcription using M-MLV Reverse Transcriptase (Thermo Fisher Scientific) from total RNA extracted with the RNeasy Plant Mini Kit (Qiagen). The *FLOE2* and *FLOE3* amplicons were then BP recombined into pDONR221 before being transferred by LR recombination into: 1) pGWB605 to generate p35S:FLOE2-GFP and p35S:FLOE3-GFP; and 2) pGWB660 to generate p35S:FLOE2-RFP and p35S:FLOE3-RFP.

Mutated FLOE1 versions: FLOE1wt, FLOE1Δnucl, FLOE1ΔCC, FLOE1ΔQPS, and FLOE1-QPS-15xY-S were amplified from the corresponding human expression vectors described in *Human plasmid construction* using prihFLOE1-FWD/REV (Table S3), and BP recombined into pDONR221 (Thermo Fisher Scientific) before being transferred by LR recombination into pGWB605 to generate p35S:FLOE1wt-GFP, p35S:FLOE1Δnucl-GFP, p35S:FLOE1ΔCC-GFP, p35S:FLOE1ΔQPS-GFP, and p35S:FLOE1-QPS-15xY-S-GFP. p35S:FLOE1ΔDS-GFP and p35S:FLOE1ΔDUF-GFP were obtained by the same process but with different primer pairs: prihFLOE1ΔDS-FWD/prihFLOE1-REV and prihFLOE1-FWD/prihFLOE1ΔDUF-REV (Table S3), respectively. p35S:FLOE1ΔDS-RFP was similarly generated by subcloning into pGWB660.

Non-*Arabidopsis* FLOE1 homologs: Protein sequences for all FLOE1 homologs shown in Fig. 7 and Fig. S6K were obtained from UniProt (UniProt, 2019) and Phytozome v13 (Goodstein et al., 2012). Their corresponding DNA sequences were generated with codon-optimization for *Nicotiana benthamiana* expression using IDT's codon optimization tool (https://www.idtdna.com/CodonOpt) (Table S3). The sequences were synthesized by GenScript Biotech Corporation (Piscataway, NJ) with flanking attB sites for subsequent BP cloning into pDONR221 (Thermo Fisher Scientific). They were then subcloned into pGWB605 by LR recombination to generate p35S:OptimizedHOMOLOG-GFP constructs (where HOMOLOG refers to the relevant FLOE1 homolog).

Analysis of FLOE1 homologs

Phylogenetic tree construction:

All *Viridiplantae* protein sequences containing the highly-conserved DUF1421 domain were retrieved from UniProt (UniProt, 2019). After removal of duplicates due to re-annotations, the remaining 791 sequences (Table S3) were submitted to the phylogenetic analysis tool NGPhylogeny.fr (Lemoine et al., 2019) with default settings. The FastME Output Tree was then uploaded to iTOL (version 5) (Letunic and Bork, 2019) for tree visualization.

Lengths of QPS and DS domains:

All monocot and eudicot sequences from the FLOE1-like and FLOE2-like clades were aligned using the msa package (version 1.20.0) in R (Bodenhofer et al., 2015). The DS and QPS regions of the homologs were defined as aligning to the DS and QPS regions of FLOE1. The lengths of these regions were used for subsequent analysis.

Tobacco transient assays

Tobacco infiltrations:

Agrobacterium cultures (GV3101 strain) carrying the relevant constructs were grown overnight at 28 ˚C in LB broth (Fisher BioReagents) containing 25 mg/L rifampicin (Fisher BioReagents), 50 mg/mL gentamicin (GoldBio) and 50 mg/L spectinomycin (GoldBio). Cultures were washed four times with infiltration buffer (10 mM $MgCl_2$ (omniPur, EMD), 10 mM MES (pH 5.6) (J. T. Baker), and 100 uM acetosyringone (Sigma-Aldrich)) and diluted to reach an OD_{600} of 0.8. Fully expanded 3^{rd}, 4^{th}, or 5^{th} leaves from 6-week-old tobacco plants were infiltrated with these diluted *Agrobacterium* cultures using Monoject 1mL Tuberculin Syringes (Covidien). For colocalization experiments, an equal amount of each culture was pre-mixed before infiltration. For each construct or combination of constructs, at least three individual tobacco plants were infiltrated.

Plasmolysis:

Square (~1cm) leaf segments were cut from tobacco leaves transiently expressing 35S:FLOE1-GFP or 35S:YFP-FLAG, and incubated for 15 minutes in either 1 M NaCl to induce plasmolysis (Zhou et al., 2019) or water (control). Cell walls were stained by supplementing the solutions with propidium iodide (Thermo Fisher Scientific) at a final concentration of 10 µg/mL. Samples were then imaged by confocal microscopy (see *Plant live imaging microscopy and image analysis* for details).

Germination experiments

Seeds were first sterilized by vortexing in 70 % ethanol for 5 minutes after which the solution was removed and replaced with 100 % ethanol. Seeds were then placed on pre-sterilized filter papers (Grade 410, VWR) and left to dry in a laminar flow hood. Sterilized seeds were then sown on square petri dishes (120 x 120 wide x 15 mm high (VWR)) containing 40 mL of MS medium (See *Plant growth conditions* section) supplemented with NaCl (Sigma-Aldrich), mannitol (Sigma-Aldrich), abscisic acid (MilliporeSigma), and paclobutrazol (MilliporeSigma) at the concentrations indicated below. Plates were then sealed with micropore surgical tape (3M)

and covered in aluminum foil before being placed at 4˚C. After exactly 120 h (5 days) of stratification to break seed dormancy, plates were transferred to a 24 h light (130 µmol.m^{-2}.s^{-1}), 22 ˚C growth cabinet (Percival). Germination (identified by radicle protrusion) was counted under a dissecting microscope.

Germination experiments were performed on seeds from two independent batches of plants (A and B) grown as described in the *Plant growth conditions* section. Germination data displayed in this publication are representative of 2-3 independent experiments as indicated in the relevant legends.

Batch A (Fig. 2, Fig. 5, Fig. S2E-H, Fig. S6A-G): 40 Col-0 and 40 *floe1-1* plants were grown alongside 10 plants of each of the following lines: four independent CRISPR lines (*floe1-2, floe1-3, floe1-4, floe1-5*), five independent FLOE1p:FLOE1-GFP lines, two independent FLOE1p:FLOE1-FLAG lines, one FLOE1p:FLOE1-GUS line, three independent FLOE1p:FLOE1ΔDS-GFP lines, four independent FLOE1p:FLOE1ΔQPS-GFP lines, and three independent FLOE1p:FLOE1ΔDUF-GFP lines. For each line, seeds from five plants were randomly pooled together which resulted in 2 biological replicates of each CRISPR and complemented line, and 8 biological replicates of Col-0 and *floe1-1*. For each biological replicate and each germination condition (0, 80 mM, 100 mM, 120 mM, 140 mM, 160 mM, 180 mM, 195 mM, 200 mM, 210 mM, 220 mM, 230 mM and 240 mM NaCl), three technical replicates (random sampling from a biological replicate) were conducted. At the end of the 230 mM NaCl germination experiment (day 15), seeds that did not germinate were rinsed in sterile double distilled water and sown on normal MS medium. Two days later, germination was scored to test whether they maintained their germination potential. The 80mM, 100mM, 120mM, 140mM, 160mM, 180mM, 200mM, 220mM, and 240mM NaCl experiments were conducted independently from the 195mM, 210mM, and 230mM NaCl experiments. For the paclobutrazol (PAC) and abscisic acid (ABA) experiments, seeds were pooled from all Col-0 or all *floe1-1* biological replicates, totaling 40 plants for each genotype, and sown on MS medium supplemented with 2µM PAC (MilliporeSigma) and 5µM ABA (MilliporeSigma), respectively. Following 5 days of stratification, germination was counted daily for the ABA experiment and after 5 days for the PAC experiment. For each condition and genotype, three technical replicates were conducted.

Batch B (Fig. S2A-D): 14 Col-0 and 27 *floe1-1* plants were grown alongside 6 plants of each of the following lines: three independent FLOE1p:FLOE1-GFP lines, two independent FLOE1p:FLOE1-FLAG lines, one FLOE1p:FLOE1-GUS line, and two independent 35S:FLOE1-FLAG lines. The 35S:FLOE1-FLAG lines failed to express *FLOE1* as revealed by RT-qPCR (Fig. S2D) and

were therefore chosen as transgenic controls. Seeds from each individual plant were sown on MS medium (See *Plant growth conditions* section) supplemented with either mannitol (400 mM) or NaCl (190 mM, 205 mM, and 220 mM). For each biological replicate and each germination condition, three technical replicates were conducted.

Root assays

Col-0 and *floe1-1* seeds were sown on MS medium (See *Germination experiments* section) and stratified for 5 days. They were then placed vertically in a 16 h light (130 µmol.m^{-2}.s^{-1}), 22 ˚C growth cabinet (Percival). 4 days later, seedlings were transferred to plates containing MS medium supplemented with NaCl at the following concentrations: 40 mM, 80 mM, 120 mM, 160 mM, and 200 mM. 15 days later, plates were imaged and roots were quantified using Fiji (Schindelin et al., 2012).

Embryo dissection and *in vivo* FLOE1 visualization experiments:

Salt, mannitol, sorbitol, cycloheximide, and water assays (Fig. 1G-J, Fig. 5B-C, Fig. 5E, Fig. S1D-E, Fig. S1E-I)

Seeds of the relevant GFP-tagged lines were submerged in either glycerin or in solutions of NaCl (Sigma-Aldrich), mannitol (Sigma-Aldrich), sorbitol (Sigma-Aldrich), cycloheximide (GoldBio) or double distilled water at the following concentrations for 15-30 min (NaCl: 0, 0.2 M, 0.4 M, 0.6 M, 0.8 M, 1 M, 1.2 M, 1.4 M, 1.6 M, 1.8 M, 2 M; mannitol: 0, 0.95 M; sorbitol: 0, 0.725 M, 1.45 M; cycloheximide: 1 g/L). They were then dissected to remove the seed coat and imaged by confocal microscopy (see *Plant live imaging microscopy and image analysis* for details). As controls, 35S:GFP (Bossi et al., 2017) seeds were dissected in water to verify that GFP alone could not induce condensate formation, and Col-0 seeds were dissected in either water or 2 M NaCl to assess the level of autofluorescence of the protein storage vacuoles in the absence of GFP in these conditions.

Salt assays (Fig. S3):

FLOE1p:FLOE1-GFP seeds were sown on MS medium (See *Plant growth conditions* section) containing NaCl at the following concentrations: 80 mM, 120 mM, 160 mM, 200 mM, 240 mM, and 280 mM. After 5 days of stratification, plates were transferred to a 24 h light (130 µmol.m^{-2}.s^{-1}), 22 ˚C growth cabinet (Percival). After 24 h, seeds were transferred to glycerin and embryos were dissected out. Embryos were then imaged by confocal microscopy (see *Plant live imaging microscopy and image analysis* for details).

Condensate reversibility assays:

Three different types of FLOE1 condensate reversibility assays were performed: 1) Embryos from dry seeds were first dissected in glycerin as described above, and after imaging, glycerin was washed off from the embryos with water and the same embryos were imaged immediately in water (less than 5 minutes); 2) Seeds were submerged in water for 1 h before being transferred to 2 M NaCl for 10 min and imaged and *vice versa* (1 h in 2 M NaCl followed by 10 min in water); and 3) Seeds were submerged in water overnight and then left to dry for an additional day. Seeds were then either dissected in glycerin to obtain the condensate state of the dry seeds or in water to assess their ability to re-form condensates.

End of germination experiment analysis:

At the end of the 230 mM NaCl germination experiment described in the *Germination experiments* section (5 days of stratification followed by 15 days in light on MS medium supplemented with 230 mM NaCl), seeds that did not germinate were either: 1) dissected directly in glycerin to leave the hydration state of the seed unaltered; or 2) transferred first to standard MS medium and dissected in glycerin two hours later. Dissected embryos were then imaged by confocal microscopy to obtain a snapshot of their final condensate state (see *Plant live imaging microscopy and image analysis*).

Developmental stages of the embryos:

FLOE1p:FLOE1-GFP and 35S:YFP-FLAG flower buds were manually self-pollinated 11, 8, 6 and 4 days before dissection to obtain developing siliques carrying embryos at mature, torpedo, heart and globular stages, respectively. Manual self-pollination entailed removing all floral parts except the pistil (female organ harboring ovules) from unopened floral buds where anthers have not shed pollen. Stamen (male organ containing pollen) from fully opened flowers from the same plant were used to pollinate the exposed pistils. Seeds from the various developmental stages were dissected either in glycerin or water and imaged by confocal microscopy (see *Plant live imaging microscopy and image analysis*).

Nuclear staining and colocalization:

Embryos from FLOE1p:FLOE1-GFP seeds were dissected in 1.6M NaCl, and 4',6-diamidino-2-phenylindole (DAPI) (Thermo Fisher) was added at a final concentration of 0.2 µg/mL. After 10 min, embryos were imaged by confocal microscopy (see *Plant live imaging microscopy and image analysis* for details).

GUS staining

FLOE1p:FLOE1-GUS seeds carrying embryos at different developmental stages were incubated at 37 ˚C overnight in GUS staining solution (Jefferson et al., 1987). In the case of dry seeds, seed coats were first removed (as they were impermeable to the staining solution) and embryos were incubated at 37 ˚C for one hour in GUS staining solution. Following the incubation, samples were destained in 70 % ethanol at room temperature for 24 hours and embryos were dissected out (in the case of developing siliques) before imaging. Pictures were taken with a compound microscope (Nikon) and a dissecting microscope (Leica MZ6 microscope).

Plant live imaging microscopy and image analysis

Image acquisition:

Embryos and tobacco leaves were imaged at room temperature on a LEICA TCS SP8 laser scanning confocal microscope in resonant scanning mode using the Leica Application Suite X software. All samples were imaged with a HC PL APO CS2 63X/1.20 water objective with the exception of embryos submerged in glycerin that were imaged with a 63X/1.30 glycerin objective and of embryos of early developmental stages that were imaged with a HC PL APO CS2 20x/0.75 dry objective. GFP, RFP, YFP, DAPI and propidium iodide fluorescence signals were detected by exciting with a white light laser at 488 nm, 561 nm, 514 nm, 405 nm, and 536 nm, respectively, and by collecting emission from 500-550 nm, 591-637 nm, 524-574 nm, 425-475 nm, and 600-650 nm, respectively, on a HyD SMD hybrid detector (Leica) with a lifetime gate filter of 1-6 ns to reduce background autofluorescence due to chlorophyll (tobacco) or protein storage vacuoles (embryos). Z-stacks were collected with a bidirectional 96-line averaging while single-frame images (tobacco) were collected with a bidirectional 1024-line averaging. For colocalization experiments, samples were imaged sequentially between each line to ensure that the colocalization signals were not due to bleed-throughs. Images shown were representative of at least three biological replicates for each construct (tobacco) or line (Arabidopsis). All samples that were compared in the publication were imaged with the same magnification and laser intensity.

Intracellular heterogeneity analysis:

For each embryo radicle and experimental condition, maximum projection images of their corresponding Z-stacks were obtained using the Leica Application Suite X software. ROIs were then manually drawn around each individual cell to obtain their standard deviation (RMS) and mean intensity levels. Heterogeneity scores were obtained by dividing the standard deviation

by the mean (coefficient of variation). Between 363 and 461 cells were measured per embryo with a total of 3 embryos per condition. Cells were characterized as exhibiting FLOE1 condensates if their heterogeneity score was higher than the top 5 percentile of the 2 M NaCl condition (heterogeneity cut-off = 0.3 arbitrary units (a.u.)).

FLOE1 condensate size determination:

For *A. thaliana* embryo radicles and tobacco leaves, individual slices of a Z-stack were analyzed using Fiji (Schindelin et al., 2012). Individual condensates were identified using a threshold, and subsequently measured for their area. A total of 3-4 embryos per condition were analyzed. For tobacco leaves data: isoform and paralog comparisons were performed on granules with a minimum size of 0.2 μm^2 and 0.5 μm^2, respectively.

Seed phenotyping

Seed weight:

Seeds from twelve *floe1-1* and fourteen Col-0 biological replicates (individual plants from Batch B described in the *Germination experiments* section) were used for the seed weight analysis. Seeds were weighed on a Sartorius M2P scale in batches of nine to twenty seeds and the process was replicated three times per biological replicate. The average weight per seed was calculated and used for subsequent statistical analysis.

Seed size and aspect ratio:

Seeds from fourteen and sixteen biological replicates of *floe1-1* and Col-0, respectively, were used for the seed size and aspect ratio analyses. Seed images were scanned using a Canon CanoScan LiDE 700 F (Canon Inc). All images were scanned at 600 dpi and, for ease of collection, the seeds were placed in transparent bags before scanning. The number of seeds per image varied, but ten seeds per sample were randomly selected and analyzed for area quantification and aspect ratio using Fiji (Schneider et al., 2012). This process was replicated ten times per biological replicate to obtain a total of one hundred seeds per biological replicate.

RNA extraction from seeds and siliques

DNA-free total RNA was extracted from seeds and siliques using the protocol described in (Meng and Feldman, 2010) as summarized in this section. For each extraction, 50-100 mg of seeds or siliques were ground in liquid nitrogen using a mortar and pestle. While still frozen, 1.2 mL of extraction buffer (100 mM Tris-HCl, pH 9.5, 150 mM NaCl, 1% sarkosyl, 0.5 % β-mercaptoethanol) was added and the solution was transferred to an RNase-free Eppendorf

tube. After vortexing for 5 min, the solution was centrifuged for 5 min at 11,000 g, and the supernatant was transferred to a new tube. 500 µL chloroform were then added and the mixture was vortexed for 2 min. 500 µL acid phenol (water saturated) were added and the tube was vortexed for an additional 2 min. After centrifugation at 11,000 g for 15 min, the upper aqueous phase was transferred to a new tube containing 90 µL 3 M sodium acetate (pH 5.2). After mixing, 600 µL isopropanol were added and mixed. The solution was then incubated for 10 min and centrifuged for 10 min at 11,000 g. After removal of the supernatant, 1 mL 75% ethanol was added to the pellet, and the tube was centrifuged at 11,000 g for 10 min. The supernatant was then discarded, and the tube was briefly spun down so that any leftover supernatant could be removed with a 10 µL pipette. The pellet was left to dry for 5 min. After this step, the dried pellet was resuspended in 1 mL TRIzol Reagent (Invitrogen) by vortexing until the pellet was fully dissolved. 200 µL chloroform were then added and mixed by vortexing for 15 s. The tube was then incubated for 2 min and centrifuged at 11,000 g at 4 ˚C for 15 min. The upper aqueous phase was transferred to a new tube containing 500 µL isopropanol. After mixing, the tube was incubated for 10 min. It was then centrifuged at 11,000 g for 15 min at 4 ˚C. The pellet was washed with 1.2 mL 75% ethanol and centrifuged at 11,000 g for 10 min at 4 ˚C. The supernatant was removed, and the tube was then briefly spun down so that any leftover supernatant could be removed with a 10 µL pipette. The pellet was left to dry for 5 min. It was then dissolved in 100 µL water and mixed with 10 µL 3 M sodium acetate (pH 5.2) and 250 µL 100% ethanol. After incubation for 20 min, the tube was centrifuged at 11,000 g for 15 min at 4 ˚C. The pellet was then washed with 1.2 mL 75% ethanol and centrifuged at 11,000 g for 10 min at 4 ˚C. The supernatant was removed, and the tube was then briefly spun down so that any leftover supernatant could be removed with a 10 µL pipette. The pellet was left to dry for 5 min, and then resuspended in 30 µL water. RNA quantity and purity were assessed using a NanoDrop Spectrophotometer (Thermo Fisher Scientific).

Unless otherwise indicated, all steps were performed at room temperature in a laminar flow hood cleaned with RNaseOUT (G-Biosciences). All solutions were made using UltraPure DNase/RNase-Free Distilled Water (Invitrogen).

RT-qPCR analyses

cDNA was synthesized from 1 µg of extracted RNA using M-MLV Reverse Transcriptase (Thermo Fisher Scientific), per manufacturer's protocol. qPCR was performed using the SensiFAST SYBR No-ROX Kit (Bioline). Primers used to quantify *FLOE1* expression were priqPCRFLOE1set1-FWD/REV (Table S3), with the exception of the qPCRs conducted on the CRISPR lines (Table S3) as well as on siliques and seeds from different developmental stages (Fig. S1A) where

priqPCRFLOE1set2-FWD/REV were used (Table S3). The different developmental stages of siliques were defined based on color: dark green, light green, and yellow, which roughly correspond to 4-7, 8-10, and 11-13 days post-anthesis, respectively (Mizzotti et al., 2018). The reference gene that was used to normalize gene expression levels, At5G25760 (*PEX4*), was chosen for consistent expression in seeds (Dekkers et al., 2012). We used a primer pair, priAT5G25760-FWD/REV (Table S3), which was reported in (Czechowski et al., 2005). Reactions were run on 96-well plates in the LightCycler® 480 Instrument II system and were repeated three times.

ELISA assays

ELISA assays were performed using the GFP ELISA Kit from Abcam. ~300 seeds from five independent FLOE1p:FLOE1-GFP lines, three independent FLOE1p:FLOE1ΔDS-GFP lines, four independent FLOE1p:FLOE1ΔQPS-GFP lines, three independent FLOE1p:FLOE1ΔDUF-GFP lines, and from Col-0 (WT) were ground in liquid nitrogen and mixed with 1 mL of "1X Cell Extraction Buffer PTR" provided in the kit and supplemented with cOmplete EDTA-free Protease Inhibitor Cocktail (MilliporeSigma). Samples were diluted in a 1:20 ratio and then quantified following the manufacturer's instructions using a Tecan INFINITE M1000 PRO plate reader.

Estimation of FLOE1 concentration in *A. thaliana* embryos

Estimation of the cytoplasmic content of *A. thaliana* embryos

To calculate the proportion of cellular volume occupied by the cytosol in embryo cells of dry seeds, we first calculated the surface area represented by the cytoplasm excluding nuclei and protein storage vacuoles using transmission electron microscopy and electron tomography (see *Transmission electron microscopy and electron tomography of Arabidopsis seeds* section below). Based on the analysis of 43 cells from embryos sectioned in random directions, we estimated that surface area to be approximately 68% (±7% SD) of the total cellular area, as seen in our micrographs. To determine how much of that 68% of cytoplasmic content is cytosol, we calculated seven electron tomograms of randomly chosen and sectioned areas of dry embryos, segmented all membrane bound organelles, except nuclei and vacuoles since they have already been accounted for. We then calculated the volume of these organelles (mainly lipid droplets, plastids, and Golgi stacks) and subtracted them from the tomographic cellular volume to estimate the proportion of cytosol, which was approximately 24.6% (±6.5% SD). Of note, very few ribosomes are detected in dry seeds, so we did not consider them in our volumetric analysis. We then calculated that 24.6% (cytosol) of 68% (cytoplasm excluding nuclei and

vacuoles) is 16.7%. Therefore, the average proportion of cytosol in an embryo cell in the dry seed is 16.7% of the total cellular volume.

To determine the cytosol volume of an entire embryo in order to estimate the concentration of FLOE1 *in vivo*, we set out to calculate the number and volume of all cells in the embryo. More specifically, we imaged full embryos to count the number of cells, and performed 3D analysis to quantify cellular volume (see *3D Cellular Resolution imaging of the mature Arabidopsis embryo* section below). We estimated the cell number at 14,000 cells per embryo and the average cellular volume at 1,210 μm^3. As such, we estimate the total cellular volume of an *Arabidopsis thaliana* embryo to be ~17,000,000 μm^3. Representative micrographs and tomographic reconstructions of embryo cells are depicted in Fig. S5B-C, Fig. S5E-F and Movie S1.

Taken together, we concluded that the cytosol occupies approximately 2,840,000 μm^3 of the whole embryo. This number was subsequently used to calculate the cytoplasmic FLOE1 concentration in dry seeds (see Table S3). Based on ELISA assays on FLOE1-GFP tagged complemented lines, our estimate of the physiological FLOE1 concentration is 0.37 µM ± 0.18 µM. However, we would like to stress that this is a lower bound estimate of the physiological concentration as we need to account for potential loss of protein during sample prep (e.g., insoluble fraction) or potential effects of protein condensation on ELISA epitope availability.

We performed a dilution series of FLOE1 *in vitro* and found that it could form visible condensates up to 1 µM (Fig. S4D). While this concentration is higher than our lower bound estimate of the *in vivo* FLOE1 concentration, we want to stress that we did not make use of any molecular crowders in our test tube phase separation assays. Molecular crowders (e.g., PEG, dextrans, etc.) are commonly used to mimic the crowding of the cytoplasmic environment in phase separation assays. Yet, given that these molecules can have pleiotropic effects and can artificially drive condensation of proteins, we instead tested whether FLOE1 had the ability to phase separate in isolation without the need of such polymers and their potential confounding effects. Our results indicate that FLOE1 can phase separate at concentrations in the range of our best estimate of its physiological concentration.

Transmission electron microscopy and electron tomography of Arabidopsis seeds

Dry seeds of *Arabidopsis thaliana* Col-0 were dissected and frozen in a Leica ICE high-pressure freezer. Samples were freeze-substituted in 2% osmium tetroxide (OsO_4) in acetone for 12 h at -80 °C followed by infiltration in Epon resin (Electron Microscopy Sciences) at room temperature. Thin (70 nm) and semi-thin (200 nm) thick sections were prepared in a Leica EM UC7 ultramicrotome, mounted in single slot grids coated with carbon-formvar, and stained with 2% uranyl acetate in 70% methanol and Reynold's lead citrate (2.6% lead nitrate and 3.5% sodium citrate, pH 12). Colloidal gold particles of 10 nm in diameter were applied to both

surfaces of the semi-thick sections to be used as fiducial markers during the alignment of the series of tilted images. Sections were imaged in a Thermo Fisher Talos F200C operated at 200 kV. For electron tomography, images were taken from +60° to –60° at 1.0° intervals using the SerialEM software (Mastronarde, 2005) and collected in a Ceta 16-megapixel CMOS camera at a pixel size of 1 nm. Tomograms were calculated using Simultaneous Iterative Reconstruction Technique (SIRT) algorithm and segmented using the IMOD package (Kremer et al., 1996). The thinning factor for each tomogram was calculated and corrected for in the models.
Fiji (Schindelin et al., 2012) was used for calculating the area of large organelles such as vacuoles and nuclei in the thin sections. Imodinfo within IMOD was employed to calculate the volume of other organelles and that of the cytoplasm.

3D Cellular Resolution imaging of the mature Arabidopsis embryo

Wild-type Col-0 seeds were imbibed for 3 h on MS plates, then embryos were dissected and fixed in 70% ethanol overnight. Samples were then rinsed 3 times in water and stained using the mPA-PI method as previously reported (Bassel et al., 2014; Truernit et al., 2008). Embryos where imaged using a Zeiss LSM710 and cells were segmented in 3D using MorphoGraphX (Barbier de Reuille et al., 2015). Each cell number and cell volume were calculated using the Heat Map function of this software.

RNA-seq experimental conditions and analysis

Experimental design:

Two RNA-seq experiments were conducted: the first on seeds from *floe1-1* and Col-0, and the second on seeds from FLOE1p:FLOE1ΔDS-GFP and FLOE1p:FLOE1-GFP. Seeds were subjected to three conditions: 1) dry; 2) imbibed in MS medium; and 3) imbibed in MS with 220 mM NaCl. Three biological replicates corresponding to pooled seeds from different plants were performed per condition, with 50 mg of mature seeds used per biological replicate. For condition (1), RNA was extracted directly from dry seeds using the protocol described in the *RNA extraction from seeds* section. For condition (2), dry seeds were sown onto agar plates of MS medium conditions and cold-stratified for 5 days at 4˚C in the dark. Plates were subsequently transferred to and held in a growth cabinet (Percival) for 4 hours under light (130 µmol.m−2.s−1) and 22˚C. After the 4-hour incubation, imbibed seeds were scraped from each plate and transferred to a clean mortar and pestle and ground in liquid nitrogen. RNA was then extracted as described in the *RNA extraction from seeds* section. Condition (3) was conducted in parallel and using the same experimental settings with the MS medium supplemented with 220 mM NaCl.

Samples from all biological replicates were first sent to the Stanford University Protein and Nucleic Acid Facility (Stanford, CA) for quantification and quality analysis using a 2100 Bioanalyzer (Agilent). Samples were then sent to Novogene Corporation Inc. (Sacramento, CA) for RNA-seq library preparation (250-300 bp insert cDNA library) and sequencing (2x150 bp paired-end reads on an Illumina Platform).

Analysis:
Reads were mapped with HISAT2 to the *Arabidopsis thaliana* TAIR10 reference genome using the Galaxy (Version 2.1.0+galaxy5) web platform (https://usegalaxy.eu) (Afgan et al., 2018). The resulting BAM files were then analyzed on RStudio (version 1.2.5033) using the DESeq2 (Love et al., 2014) and TxDB.Athaliana.BioMart.plantsmart28 (Bioconductor) packages. Genes with adjusted p-value < 0.05 were considered differentially expressed. Gene Ontology and KEGG enrichment of the differentially expressed genes was obtained using g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) (Raudvere et al., 2019).

All-atom modeling and structural prediction
To obtain structural insight into the previously uncharacterized nucleation domain of FLOE1, we combined a set of integrative homology modelling approaches with all-atom simulations (Fig. S4E). We first leveraged a number of homology-modelling based tools to perform *de novo* structural prediction, including SWISS-MODEL and Phyre2 (Kelley et al., 2015; Waterhouse et al., 2018). SWISS-MODEL strongly predicted the majority of the nucleation domain to consist of a trimeric coiled-coil, while in agreement Phyre2 predicted a single long helix with the appropriate exposed interfacial residues necessary for trimeric coiled-coil formation. The secondary structure prediction server PSIPRED identified a number of overlapping residues predicted to be helical (McGuffin et al., 2000). Taken together, all structural modelling tools predict this region to be highly helical and to form higher-order oligomeric assemblies, providing a molecular explanation for its observed cellular function.

To provide further support for the predicted helical structure of the nucleation domain, we turned to all-atom simulations. Simulations were performed with the ABSINTH implicit solvent model and CAMPARI Monte Carlo simulation engine (Vitalis and Pappu, 2009). We first generated an initial helical starting structure by applying a bias force to drive the protein into a single alpha helix. From this helical starting state, production simulations were run in which no bias was applied, and the system was able to evolve. Five independent simulations were

performed with the reported helicity being the average of all five runs. Simulations were run using the ABSINTH implicit solvent model and abs_3.2_opls.prm parameters. Each simulation was run for 6.8 x 10^7 Monte Carlo steps with the first 6.0 x 10^6 steps discarded as equilibration. These simulations revealed a remarkably good agreement with the predictions from homology modelling tools (Fig. S4E) including a break in helicity identified in the homology modeling tools between residues 130 and 140. Taken together, distinct structural and biophysical approaches converge on a model in which the nucleation domain drives higher-order assembly as a multivalent trimeric coiled-coil domain.

Recombinant FLOE1 cloning, expression, purification, and phase separation

FLOE1's coding sequence was transferred from G18395 (see *Plant vector construction* section) into pDEST-HisMBP (Addgene plasmid # 11085) (Nallamsetty et al., 2005) using LR Clonase II (Thermo Fisher Scientific). A sequence encoding the Tobacco Etch Virus (TEV) protease recognition sequence, ENLYFQ, was then added between the MBP and FLOE1 coding sequences using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) with primers priTEV-FWD/REV (Table S3) to generate pDEST-HisMBP-ENLYFQ-FLOE1. A sequence encoding ENLYFQ followed by FLOE1ΔDS was synthesized by GenScript Biotech Corporation (Piscataway, NJ) with flanking attB sites for subsequent BP cloning into pDONR221 (Thermo Fisher Scientific). The resulting vector, pDONR221-ENLYFQ-FLOE1ΔDS, was then subcloned into pDEST-HisMBP (Addgene plasmid # 11085) (Nallamsetty et al., 2005) to generate pDEST-HisMBP-ENLYFQ-FLOE1ΔDS.

BL-21(DE3) competent cells (Agilent) were transformed with the plasmids following supplier protocol and plated on LB agar plated with ampicillin selection overnight at 220 RPM and 37 °C. Transformed cells were expressed in 6 L of LB at pH 7.4 with ampicillin selection. Expression was induced at OD_{600} of 0.6 using 0.5 mM IPTG, and cells were left under shaking at 220 RPM and 16 °C for 19 h prior to collection. Collected cells were spun down at 4 °C for 8 min and the supernatant was discarded. The pellet was resuspended with 20 mL lysis buffer (50 mM NaH_2PO_4, 0.5 M NaCl, pH 8) and 1 cOmplete Mini Protease Inhibitor Cocktail tablet (Roche) per 1 L of expression, and the resuspended cells were lysed via homogenization for 8 min (Emulsiflex homogenizer) or via sonication (Qsonica Q700, 0.5 inch tip). The resulting cell lysate was spun down for 50 min at 19,500 g, and the supernatant was collected. Ni-NTA nickel beads (Qiagen) were equilibrated with lysis buffer (50 mM NaH_2PO_4, 0.5 M NaCl, pH 8), then loaded with the lysate and washed with 50 mL of wash buffer (50 mM NaH_2PO_4, 0.5 M NaCl, 20 mM Imidazole, pH 8) followed by 8 mL of Elution Buffer (50 mM NaH_2PO_4, 0.5 M NaCl, 250 mM

Imidazole, pH 8) all done at 4 °C. The eluent was collected, spun down to remove aggregates, and further purified with size exclusion chromatography (SEC) using a Superdex 200 16/60 column (GE). 5 mL of spun-down eluent was injected onto the column and ran at 0.5 mL/min at room temperature. Fractions were collected and the presence of the protein was verified using SDS-PAGE (Fig. S4A). Differential interference contrast (DIC) microscopy images were taken from each peak immediately after SEC. Peak #1 showed many small aggregates (not shown) whereas Peak #2 was largely devoid of aggregates (Fig. S4B).

Four successive rounds of SEC showed almost identical results, indicating that the protein remained intact for at least 24 hours. SDS-PAGE also showed almost identical bands from corresponding collected fractions. Of note is the distinction in properties between Peak #1 and Peak #2. Peak #1 elutes at the void volume of the column, meaning it is larger than what can be resolved with Superdex 200 (> 200 kDa) (Fig. S4A) and results in amorphous aggregation in both the pre- and post-cleaved FLOE1. Peak #2 elutes afterwards, implying it is closer to a monomeric state. It shows very little aggregation before cleavage and displays the spherical droplets indicative of phase separation after removing the MBP with TEV (Fig. S4B-C). We therefore believe that peak #1 contains aggregates of FLOE1-MBP, while peak #2 contains the monomeric protein, and have chosen this peak for our cleavage reaction. This behavior has been observed for several other LLPS-forming proteins, such as FUS (Burke et al., 2015).

TEV Cleavage

All proteins were thawed and spun at 4 °C, 20,800 g for 10 min to remove aggregates. TEV was added to 10 μM protein unless stated otherwise in a 1:50 w/w ratio in 20 mM TRIS and 150 mM NaCl, pH 8. Samples were incubated overnight at 4 °C and rotated using the Thermo Fisher tube revolver (speed 10). The cleavage reaction was confirmed with SDS-PAGE (Fig. S4C). DIC images of the mixture following cleavage confirmed the appearance of round droplets (Fig. S4B). For conditions using NaCl, the reaction was started with protein and TEV alone. A highly concentrated NaCl solution was added to the final required concentration after cleavage, and vortexed for 5 seconds immediately before imaging.

DIC Microscopy

8 well silicon gaskets (Grace Bio labs) were used as chambers and placed on a Fisherbrand glass microscopy slide. 21 μL of sample were placed in each well and sealed with a #1.5 coverslip. DIC images were taken on a Zeiss Observer 3 inverted microscope using a 40x 0.9 NA dry objective

18-22 h after TEV addition, and after 10-15 minutes in the imaging chamber. Images were taken using a Hamamatsu Orca Flash v3.0 camera with an exposure time of 100 ms (Fig. S4B,D,F).

Human plasmid construction

FLOE1 and derived mutant constructs for expression in human cells were optimized for human expression (Table S3) and generated through custom synthesis and subcloning into the pcDNA3.1+N-eGFP backbone by Genscript (Piscataway, USA). Constructs to test the nucleation domain were synthesized and subcloned by Genscript into a custom pcDNA3.1+ backbone containing mCherry.

Human cell culture and microscopy

U2OS cells (ATCC, HTB-96) were grown at 37 °C in a humidified atmosphere with 5 % CO_2 for 24 h in Dulbecco's Modified Eagle's Medium (DMEM), high glucose, GlutaMAX + 10 % Fetal Bovine Serum (FBS) and pen/strep (Thermo Fisher Scientific). Cells were transiently transfected using Lipofectamine 3000 (Thermo Fisher Scientific) according to manufacturer's instructions. Cells grown on cover slips were fixed for 24 h after transfection in 4 % formaldehyde in PBS. Slides were mounted using ProLong Gold antifade reagent (Life Technologies). Confocal images were obtained using a Zeiss LSM 710 confocal microscope. Images were processed using Fiji (Schindelin et al., 2012).

FRAP measurements in human cells

U2OS cells were cultured in glass bottom dishes (Ibidi) and transfected with GFP-FLOE1 constructs as described above. After 24 h, GFP-FLOE1 condensates were bleached and fluorescence recovery after bleaching was monitored using ZEN software on a Zeiss LSM 710 confocal microscope with incubation chamber at 37 °C and 5 % CO_2. Data were analyzed as described previously (Boeynaems et al., 2017). In brief, raw data were background subtracted and normalized using Excel, and plotted using GraphPad Prism 8.4.1 software.

Standard transmission electron microscopy (TEM) in human cells

U2OS cells were transfected at 80% confluency in Petri dishes, transfected for 24 h (see above) and fixed with 2.5 % glutaraldehyde in 0.1 M Na-cacodylate buffer (pH 7.2). After washing in the same buffer three times, the cells were scraped and centrifuged at 200 x g. The pellet was resuspended in 1.5 % low melting point agarose (type VIIa, Sigma-Aldrich) in 0.1 M Na-cacodylate buffer and pelleted for 1 min at 1000x g. After solidification on ice, the agarose-embedded cell pellet was cut in small blocks, incubated in 1 % osmium tetroxide for 2 h at room

temperature, washed 3x in milli-Q water and dehydrated in a graded ethanol series until 100 % ethanol in steps of 5 minutes. Finally, following 2 washes with propylene oxide, the cells were infiltrated with epoxy resin (Agar 100, EMS, Hatfield, PA, USA) embedded in BEEM-capsules and cured for 48 h at 60°. Ultrathin sections (70 nm) were cut from the polymerized samples with a Leica ultracut UCT ultramicrotome (Leica, Vienna, AU), and post-stained with 4% uranyl acetate (SPI Supplies, Westchester, PA, USA) for 8 mins and Reynold's lead citrate for 3 mins. Finally, the cells were observed and imaged at an acceleration voltage of 80 kV with a JEOL JEM1400 (Tokyo, JP) transmission electron microscope equipped with an 11 Mpx EMSIS Quemesa camera (EMSIS GmbH, Muenster, DE).

Correlative Light and Electron Microscopy (CLEM) in human cells

For CLEM, plated cells were lightly fixed with 2% PFA and washed 3x with PBS. The cells were scraped and pelleted at 200x g after which they were resuspended in 20 % BSA and pelleted again at 200x g. The loosely packed cells in BSA were high-pressure frozen in a Leica Empact 2 high-pressure freezer (Leica, Vienna, Au), and submitted to a quick freeze-substitution protocol that preserves fluorescence (McDonald and Webb, 2011; Peddie et al., 2014). Briefly, frozen samples were freeze-substituted in acetone containing 0.2 % uranyl acetate and 5 % H_2O in a Styrofoam box on a rotating platform while temperature was allowed to rise to -50°C at which moment they were transferred to the Leica AFS2 automatic freeze-substitution apparatus equipped with a Leica Freeze Substitution Processor (FSP) processing robot (Leica, Vienna, AU). After total time elapsed between -80°C and -50°C amounted 1.5 h, samples were washed in acetone and infiltrated in Lowicryl HM20 resin (EMS, Hatfield, PA, USA), and finally polymerized at -50 °C by UV-light. Sections of 100 to 200 nm were cut with a Leica ultracut UCT ultramicrotome and deposited on an indium-tin oxide coverslip, and fluorescent images were taken with a Zeiss Axioplan light microscope (Zeiss, Oberkochen, DE). After this, the ITO-coverslip with sections was mounted to a support stub, placed inside the specimen chamber of a Zeiss Sigma scanning electron microscope (Zeiss, Oberkochen, DE), and imaged at 1.25 -1.5 kV acceleration voltage with a Gatan backscattered electron detector. For correlation, fluorescent and electron images were overlaid using GIMP (GNU Image Manipulation Program).

Electron tomography in human cells

For tomography, cells were prepared until embedded in agarose blocks as described above. Then, the blocks were incubated in 1% osmium tetroxide (EMS, US) and 1.5% Ferrocyanide in milli-Q water for 1 hour in the dark on ice, after which they were washed in milli-Q and incubated in 0.2% tannic acid in milli-Q for 30 mins. After washing in milli-Q the cells were

osmicated a second time in 1% osmium tetroxide in milli-Q for 30 mins in the dark and washed in milli-Q. Then, samples were stained *en bloc* by incubation in 2% uranyl acetate (EMS, US) in milli-Q overnight at 4°C, followed by incubation in Walton's lead aspartate for 30 mins at 60°C in the dark (Walton, 1979). After washing in milli-Q, dehydration in a graded ethanol series and propylene oxide, the cells were infiltrated with epoxy resin (Agar 100), embedded in BEEM-capsules and cured for 2 days at 60°C. Finally, 200 nm sections were cut on a Leica ultracut S, and put on carbon-formvar coated tri-slot grids (Ted Pella, US). Image stacks from tilt series (-60° to 60°) were taken from selected regions with condensates with the same TEM as described above, but operated at 120 kV, using TEMography software (System In Frontiers, JP). Tomograms were made from the image stacks with IMOD (Kremer et al., 1996), using Simultaneous Iterative Reconstruction Technique algorithm and Non-linear Anisotropic Diffusion filtering. Segmentation was done with MIB (Fisher, NL) by thresholding after additional denoising with DenoisEM, a plug-in in Fiji (Roels et al., 2020).

Natural variation transcript expression analysis

RNA-seq data from leaf tissue samples of 478 *Arabidopsis thaliana* worldwide ecotypes from the 1001 transcriptomes project (Kawakatsu et al., 2016) (Table S4, sheet2), for which reliable *FLOE1* expression was obtained, were mapped to the Araport11 reference using kallisto (Bray et al., 2016) to estimate transcript abundance. Seed germination and dormancy phenotypes were obtained from multiple previously published studies (Table S4, sheet 1, n=45) and phenotype measurements from each study were matched to each ecotype's corresponding RNA-Seq data. This yielded variable overlap of *A. thaliana* samples, between 14-260, depending on the ecotypes used in each previously published study (Atwell et al., 2010; Martinez-Berdeja et al., 2020). We looked for correlation between these 45 phenotypes and *FLOE1* expression at both gene-level and transcript-level for each of the two *FLOE1* isoforms (AT4G28300.1, AT4G28300.2). For transcript-level analysis, estimated TPM (transcripts per million) values were used directly. For gene-level expression analysis, estimated TPM of the two *FLOE1* isoforms were summed, $\log_2$ transformed, and mean centered. Spearman's rank correlation coefficients were calculated between *FLOE1* expression (each isoform transcript and gene level) and the corresponding germination phenotypes. All expression-phenotypes correlations are listed in Table S4 sheet 1. In associated supplemental figures, TPM were $\log_2$ transformed and mean centered prior to plotting for better visualization. Climate data from the original collection sites of 478 *A. thaliana* populations (Table S4) was extracted from either 1960-1990 average climate maps of Europe (monthly precipitation, minimum and maximum temperature (Hijmans et al.,

2015)) or a time series between 1958 to 2017 (monthly temperature and precipitation values (Abatzoglou et al., 2018). We then conducted correlations with climate averages or coefficient of variation of time series data with *FLOE1* expression. All correlations were computed as Spearman's rank correlation, which are more robust to outliers and data distributions than linear Pearson's correlation.

Quantification and Statistical Analysis

All data was analyzed using RStudio (version 1.2.5033), GraphPad Prism (version 8.4.1) and Excel. Statistical tests, p values, number of biological and technical replicates, and number of independent experiments are indicated in the figure legends. R packages used are described in the STAR Methods.