



Mobile TERMINAL FLOWER1 determines seed size in *Arabidopsis*

Bin Zhang^{1,2}, Chengxiang Li^{1,2}, Yan Li¹ and Hao Yu^{1,2}✉

Seed size is a pivotal agronomic trait that links plant sexual reproduction and subsequent seedling establishment, and is affected by the timing of endosperm cellularization following endosperm proliferation after double fertilization. The molecular switch that controls the timing of endosperm cellularization has so far been largely unclear. Here, we report that the *Arabidopsis* TERMINAL FLOWER1 (TFL1) is a mobile regulator generated in the chalazal endosperm, and moves to the syncytial peripheral endosperm to mediate timely endosperm cellularization and seed size through stabilizing ABSCISIC ACID INSENSITIVE 5. We further show that Ras-related nuclear GTPases interact with TFL1 and regulate its trafficking to the syncytial peripheral endosperm. Our findings reveal TFL1 as an essential molecular switch for regulating endosperm cellularization and seed size. Generation of mobile TFL1 in the chalazal endosperm, which is close to maternal vascular tissues, could provide a hitherto-unknown means to control seed development by mother plants.

Seed development determines plant evolutionary fitness and crop yield in flowering plants. Among various growth parameters of seed development, seed size is closely associated with nutrient levels for seed germination and the capacity of seed tolerance to abiotic stress during seedling establishment¹. Thus, seed size is always an essential agronomic trait during domestication and breeding in many crops. By now, understanding of seed size control in the model plant *Arabidopsis thaliana* has provided important insights into molecular breeding for crop improvement.

Seed development in *Arabidopsis* starts with rapid growth of the endosperm and integument to form a large seed cavity, followed by replacement of the endosperm by the embryo^{2,3}. The volume of seed cavity is an essential factor that influences the final seed size by spatially confining the embryo growth^{2,4}. Integument elongation and endosperm expansion usually terminate after endosperm cellularization, which defines the available cavity space so that the seed maintains almost the same size in the course of subsequent embryo enlargement. Thus, the final seed size is tightly controlled by the timing of endosperm cellularization, and disruption of endosperm cellularization is usually associated with changes in seed size as exemplified by mutations in the regulators in the IKU pathway^{2,5–7}. In this pathway, loss of function of *HAIKU1* (*IKU1*), *HAIKU2* (*IKU2*) and *MINISEED3* (*MINI3*) shows precocious endosperm cellularization and consequently results in a small seed phenotype, whereas gain of function of *SHORT HYPOCOTYL UNDER BLUE1* (*SHB1*) causes delayed endosperm cellularization and big seeds^{2,8,9}. In addition, abscisic acid (ABA) has been found to influence the timing of endosperm cellularization and negatively mediate seed size through regulating *SHB1* transcription by ABA INSENSITIVE 5 (ABI5)¹⁰.

Following the fertilization of a sperm nucleus with the binucleate central cell, the resulting primary endosperm nucleus undergoes mitoses without cytokinesis, generating the syncytial endosperm that is spatially divided into three domains; micropylar, chalazal and peripheral endosperm³. The micropylar endosperm surrounds the embryo and suspensor, while the chalazal endosperm is at the chalazal pole located opposite the embryo. The peripheral endosperm spreads as a peripheral layer due to the force of a large central

vacuole at the initial stage of seed development. After eight rounds of syncytial mitoses, the micropylar and peripheral endosperm subsequently undergo cellularization, whereas the chalazal endosperm always keeps its syncytial state^{6,11–13}. Ultrastructural studies have indicated that the chalazal endosperm, which is close to the maternal vascular tissues, serves as an endospermic haustorium in the uptake, reprocessing and transport of nutrients or signals to developing seeds^{12–16}. Despite these studies, the molecular mechanisms underlying the role of the chalazal endosperm in seed development are still completely unknown.

In this study, we report that *TERMINAL FLOWER1* (*TFL1*), a phosphatidylethanolamine binding protein (PEBP) family member previously known as a shoot identity gene in *Arabidopsis*^{17–19}, plays an essential function in determining seed size by regulating appropriate timing of endosperm cellularization in *Arabidopsis*. We demonstrate that a group of small GTP-binding Ras-related nuclear (RAN) proteins²⁰ mediate the transport of TFL1 protein from the chalazal endosperm to the syncytial peripheral endosperm (SPE), where TFL1 stabilizes ABI5, thus mediating timely cellularization of the endosperm and seed size. These findings suggest that TFL1 serves as a hitherto unknown mobile signal that could respond to the uptake of maternal signals in the chalazal endosperm to determine the size of progeny seeds.

Results

***TFL1* loss-of-function mutants exhibit a large seed phenotype.** In the course of studying *TFL1* function in regulating inflorescence architecture in *Arabidopsis*¹⁹, we found that its loss-of-function mutants, *tfl1-1* (ref. ¹⁷) and *tfl1-20* (ref. ²¹), exhibited an obvious large seed phenotype (Fig. 1a,b), which was associated with an increased size of the mature embryo (Fig. 1c). Further examination of the seeds of the transfer-DNA (T-DNA) insertional mutant, *tfl1-20*, revealed that the main seed size parameters, including seed area, perimeter, length, width, length/width and seed mass, were notably increased in *tfl1-20* compared to wild-type seeds (Fig. 1d), whereas the seed number per siliques in *tfl1-20* was comparable to that of wild-type plants (Supplementary Fig. 1a,b).

¹Department of Biological Sciences, Faculty of Science, National University of Singapore, Singapore, Singapore. ²Temasek Life Sciences Laboratory, National University of Singapore, Singapore, Singapore. ✉e-mail: dbsyuha@nus.edu.sg

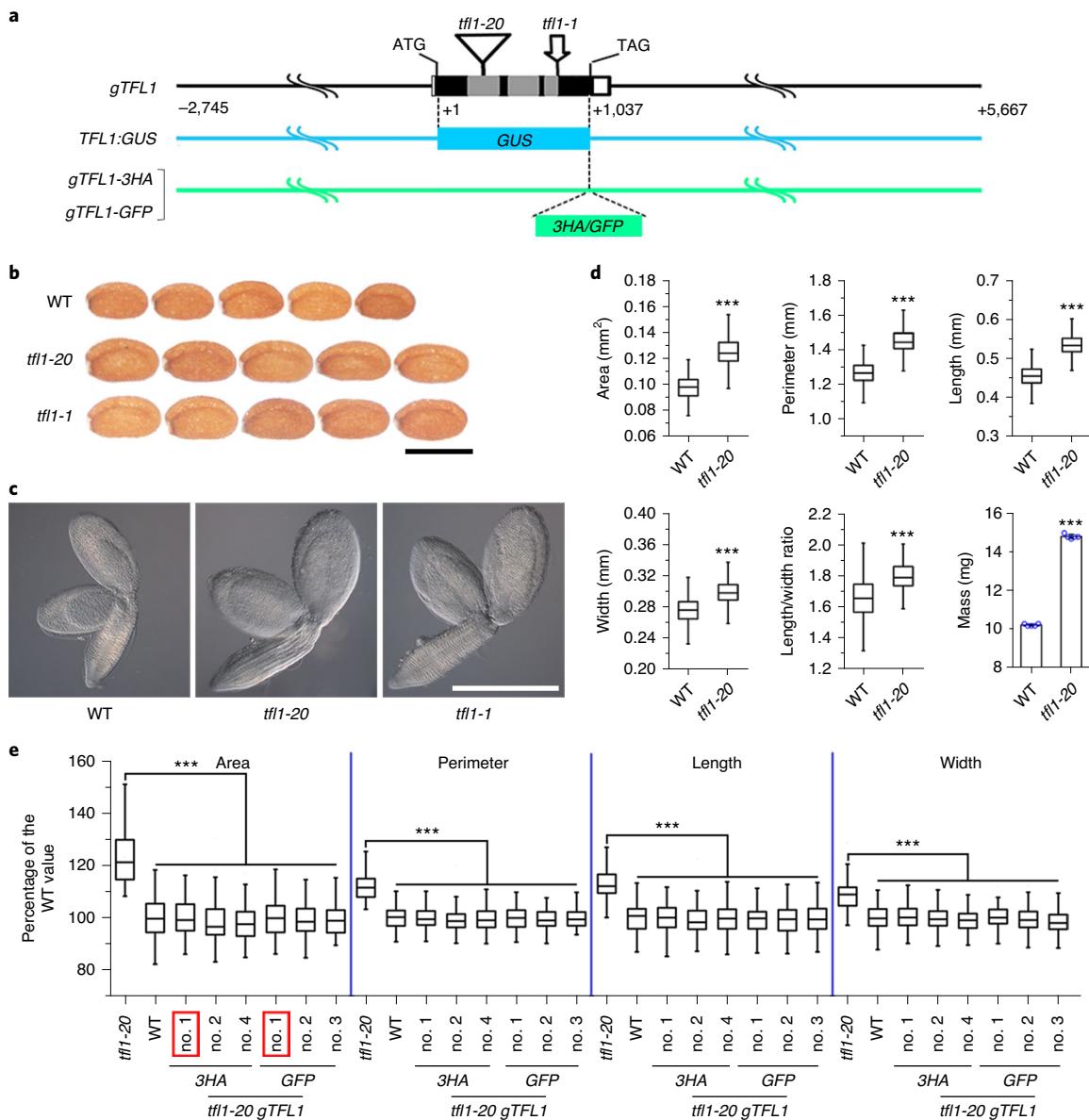


Fig. 1 | *TFL1* affects seed size. a, Schematic diagrams of *tfl1* mutants, *TFL1:GUS*, and various genomic *TFL1* (*gTFL1*) constructs. Inverted triangle or arrow indicates the T-DNA insertion site (*tfl1-20*) or the ethylmethanesulfonate-mutated site (*tfl1-1*), respectively. Exons, introns and untranslated regions are represented by black, grey and white boxes, respectively. Other genomic regions are represented by black lines. Nucleotide sequences are numbered relative to the first nucleotide of the translation start codon assigned as +1. **b**, Comparison of mature dry seeds of wild-type (WT), *tfl1-20* and *tfl1-1* plants. Scale bar, 500 µm. **c**, Comparison of mature embryos of WT, *tfl1-20* and *tfl1-1* plants. Scale bar, 500 µm. The experiments in **b,c** were repeated three times independently with similar results. **d**, Quantitative analysis of seed size and mass of WT and *tfl1-20* plants. Box plots show medians (lines), interquartile ranges (boxes) and whiskers (extending 1.5× the interquartile ranges) of seed size parameters (area, perimeter, length, width and length/width ratio) of 500 WT and *tfl1-20* seeds. Seed mass is the average value of five sample batches with each containing 500 seeds. **e**, Quantitative analysis of seed size parameters of representative *tfl1-20 gTFL1-3HA* and *tfl1-20 gTFL1-GFP* lines potentially containing the transgene at a single locus based on a 3:1 Mendelian segregation ratio. Box plots display medians (lines), interquartile ranges (boxes) and whiskers (extending 1.5× the interquartile ranges) of seed size parameters of seeds of different genotypes (*tfl1-20*, *n*=148; WT, *n*=106; *tfl1-20 gTFL1-3HA* nos. 1, 2, 4 were *n*=100, 102, 106, respectively; *tfl1-20 gTFL1-GFP* nos. 1, 2, 3 were *n*=108, 126, 131, respectively). The lines highlighted by the red frame were selected for further analysis. Percentage change (%) in a seed parameter of a specific genotype is shown relative to the mean value of WT plants set as 100% in box plots. Asterisks in **d,e** indicate significant differences between *tfl1-20* and other genotypes (two-tailed Student's *t*-test for seed mass comparison; two-tailed Mann-Whitney test for the other comparisons, $P < 0.0001$). In **d,e** all *P* values for seed size parameters are less than 1×10^{-15} , while *P* value for seed mass in **d** is 4.67×10^{-13} .

To confirm that the large seed phenotype of *tfl1-20* is attributed to the loss of *TFL1* function, we transformed *tfl1-20* with the genomic constructs of *gTFL1-3HA* and *gTFL1-GFP* harbouring a 8.41-kb *TFL1* genomic fragment including the 2.74-kb 5' upstream sequence, the 4.63-kb 3' downstream sequence, and the entire

1.04-kb coding region fused in frame with a 3HA or green fluorescent protein (GFP) tag before the stop codon TAG, respectively (Fig. 1a). Most of the *tfl1-20 gTFL1-3HA* and *tfl1-20 gTFL1-GFP* independent transformants exhibited seed phenotypes comparable to those of wild-type plants (Fig. 1e). In addition, *gTFL1-3HA* and *gTFL1-GFP*

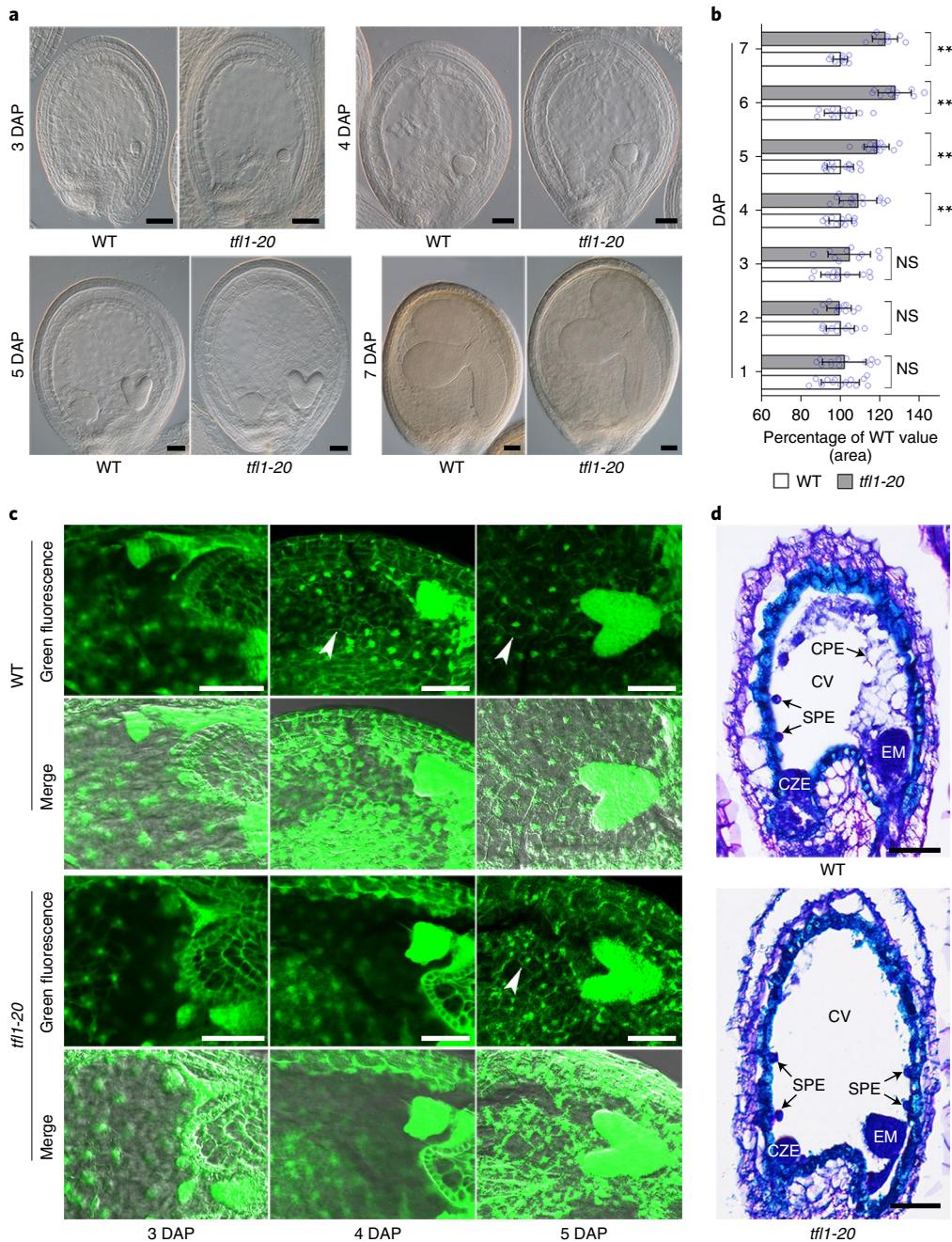


Fig. 2 | Endosperm cellularization is delayed in *tf1-20*. **a**, DIC microscopy of cleared whole-mount wild-type (WT) and *tf1-20* seeds at 3, 4, 5 and 7 DAP. Scale bars, 50 µm. **b**, Comparison of WT and *tf1-20* seed cavity area at 1–7 DAP. Randomly selected seeds were examined at each time point (WT seeds at 1–7 DAP, $n=15, 11, 12, 12, 14, 13$ and 9, respectively; *tf1-20* seeds at 1–7 DAP, $n=12, 12, 10, 14, 13, 11$ and 9, respectively). Values are mean \pm s.d. The WT mean value on each day is set to 100%. Asterisks indicate statistically significant differences (two-tailed Student's *t*-test, $P < 0.01$). NS, not significant. P values for 1–7 DAP are $0.628, 0.811, 0.313, 8.99 \times 10^{-3}, 1.11 \times 10^{-7}, 4.31 \times 10^{-8}$ and 7.75×10^{-8} , respectively. **c**, Comparison of the peripheral endosperm development of cleared whole-mount WT and *tf1-20* seeds at 3–5 DAP by confocal microscopy. The autofluorescence (green fluorescence) signal was generated by glutaraldehyde treatment. Arrowheads indicate new cell wall during endosperm cellularization. Merge, merge of green fluorescence and bright-field images. Scale bars, 50 µm. **d**, Histological analysis of WT and *tf1-20* seeds at 4 DAP. CPE, cellularized peripheral endosperm; CV, central vacuole; CZE, chalazal endosperm; EM, embryo. Scale bars, 50 µm. The experiments in **a,c,d** were repeated three times independently with similar results.

also fully rescued other developmental defects of *tf1-20*, including early flowering and terminal flower phenotypes (Extended Data Fig. 1a,b). These results indicate that *TFL1* is responsible for the seed size phenotype observed in *tf1-20*, and that the *gTFL1* genomic fragment contains the essential regulatory elements required for TFL1 activity. We further selected one representative line each for *tf1-20*

gTFL1-3HA and *tf1-20 gTFL1-GFP* (Fig. 1e), which potentially contained the transgene at a single locus based on a 3:1 Mendelian segregation ratio, for further investigation.

***TFL1* determines seed size by affecting endosperm cellularization.** Histological analysis of *tf1-20* and wild-type seeds from 1 to

7 d after pollination (DAP) revealed that progression of their embryo development was similar (Fig. 2a and Extended Data Fig. 2a), whereas the difference in their seed cavity became more obvious immediately after the globular stage at 3 DAP (Fig. 2a,b). From the globular-heart transition stage at 4 DAP, the seed cavity in *tfl1-20* was significantly larger than that in wild-type seeds (Fig. 2b), which is consistent with distinguishable size between *tfl1-20* and wild-type seeds (Fig. 2a and Extended Data Fig. 2a).

Since the enlarged seed cavity phenotype was detectable in *tfl1-20* from 4 DAP, we then performed confocal microscopy analysis of endosperm development of *tfl1-20* and wild-type seeds from 3 to 5 DAP through monitoring the autofluorescence generated by glutaraldehyde-fixed endosperm cells. Notably, the wild-type peripheral endosperm underwent cellularization at the transition stage (4 DAP), whereas endosperm cellularization was delayed to the heart stage (5 DAP) in *tfl1-20* seeds (Fig. 2c and Extended Data Fig. 2b). Furthermore, histological analysis of the paraffin-embedded seeds at the transition stage (4 DAP) also revealed partially cellularized peripheral endosperm only in wild-type, but not in *tfl1-20* seeds (Fig. 2d). These observations demonstrate that endosperm cellularization is delayed in *tfl1-20* relative to wild-type seeds.

TFL1 protein is a mobile signal in developing seeds. To understand how *TFL1* affects endosperm cellularization and seed size, we performed quantitative real-time PCR analysis of total RNA extracted from various tissues of wild-type plants. *TFL1* messenger RNA (mRNA) was ubiquitously expressed almost in all tissues examined with the lowest expression in leaves, while its expression in developing siliques decreased before the transition stage (4 DAP), but increased afterwards (Extended Data Fig. 3a). In situ analysis of developing seeds revealed that *TFL1* mRNA was expressed specifically in the chalazal endosperm, but not in the SPE at the transition stage (Fig. 3a and Extended Data Fig. 3b). We further created a *TFL1:β-glucuronidase (GUS)* reporter construct in which the *GUS* gene was fused between the same 5' upstream and 3' downstream sequences included in the *TFL1* genomic fragments used for the gene complementation test (Fig. 1a). Strong GUS signal was again detected in the chalazal endosperm in seeds at the transition stage (Fig. 3b). These results, together with the *TFL1* expression pattern analysed from the public microarray data in the *Arabidopsis* eFP browser (Fig. 3c), demonstrate the specific expression of *TFL1* mRNA in the chalazal endosperm.

To examine the TFL1 protein localization in developing seeds, we took advantage of the established *tfl1-20 gTFL1-GFP* transgenic line (Fig. 1a), in which TFL1-GFP was able to rescue the pleiotropic defects of *tfl1-20* (Fig. 1e and Extended Data Fig. 1). Confocal microscopy analysis detected TFL1-GFP signal in the inflorescence shoot apical meristem and root phloem tissues (Extended Data Fig. 3c,d) as shown in a previous study²², indicating that the trackable TFL1-GFP is biologically functional in vivo. TFL1-GFP protein

was detectable abundantly in the SPE, but only weakly in the chalazal endosperm in developing seeds at the transition stage (Fig. 3d and Extended Data Figs. 3e,f and 4a). TFL1 localization in the SPE was further substantiated by transmission electron microscopy of the *tfl1-20 gTFL1-3HA* transgenic line (Fig. 3e and Extended Data Fig. 4c), in which TFL1-3HA rescued the pleiotropic defects of *tfl1-20* (Fig. 1e and Extended Data Fig. 1) and was detectable by western blot analysis using anti-HA antibody (Extended Data Fig. 3g).

We also noted that TFL1-GFP localization in the syncytial or cellularized peripheral endosperm accumulated in the dense cytoplasm surrounding each endosperm nucleus in developing seeds (Fig. 3d and Extended Data Figs. 3e and 4a–c), which is consistent with its cytoplasmic localization observed in inflorescence meristem cells (Extended Data Fig. 4d)²². Further immunoblot analysis of TFL1-3HA in proteins extracted from *tfl1-20 gTFL1-3HA* siliques confirmed TFL1 localization specifically in the cytoplasm (Fig. 3f).

These expression analyses of developing seeds demonstrate that while *TFL1* mRNA was specifically expressed in the chalazal endosperm (Fig. 3a–c and Extended Data Fig. 3b), its protein was abundantly localized in the cytoplasm of the SPE (Fig. 3d,e and Extended Data Figs. 3e and 4a,c). Since *TFL1* mRNA was undetectable in the SPE (Fig. 3a–c, Extended Data Fig. 3b), the differential localization of *TFL1* mRNA and protein suggests that TFL1 protein probably moves from the chalazal endosperm to the SPE in developing seeds.

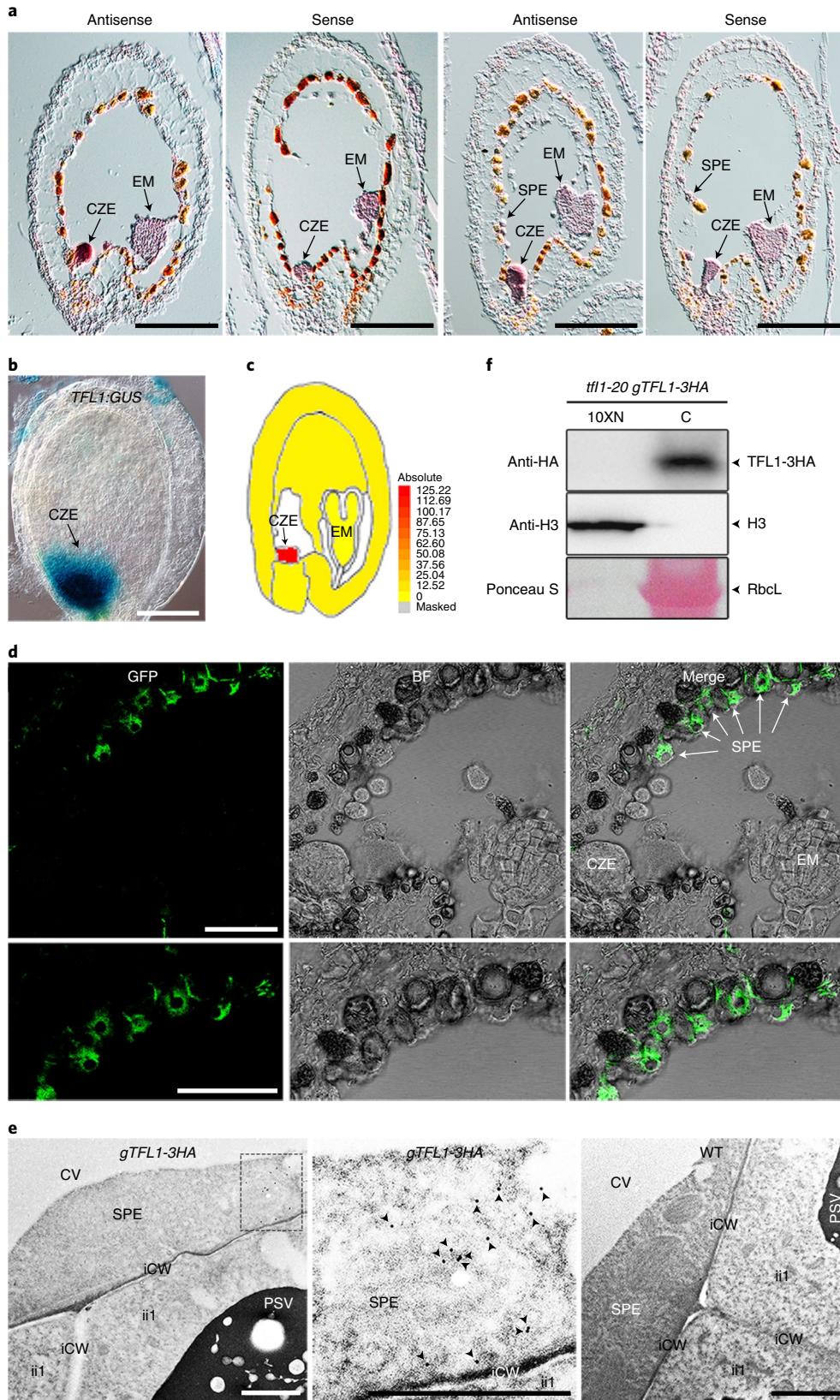
RAN proteins are potential mediators for TFL1. To gain insights into the molecular mechanism of TFL1 trafficking within the syncytial endosperm, we performed immunoprecipitation–mass spectrometry (IP–MS) and pull-down–MS (PD–MS) assays to identify TFL1 interacting partners in seeds. Total protein extracts from *tfl1-20 gTFL1-3HA* siliques were immunoprecipitated by anti-HA antibody for IP–MS, while the recombinant MBP-TFL1 protein was used to pull down its interacting proteins in total protein extracts from wild-type siliques for PD–MS. The copurified complexes from both approaches were then analysed by liquid chromatography coupled with tandem MS (LC–MS/MS). This identified two peptides (LVIVGDGGTGK and NLQYYEISAK) (Fig. 4a) that resided in two well-conserved regions with 100% identity in amino acid sequences among three putative Ran GTPases, RAN1–3 (Fig. 4b and Supplementary Fig. 2), which are involved in nucleocytoplasmic transport²⁰ and share high sequence similarity.

To understand the biological functions of RAN genes, we obtained three T-DNA insertional lines, including *ran2-1* for *RAN2* as well as *ran3-1* and *ran3-2* for *RAN3* (Extended Data Fig. 5a), from the *Arabidopsis* Biological Resource Center. As *RAN2* or *RAN3* was barely detectable in *ran2-1* or *ran3-1* (Extended Data Fig. 5b), we used these two lines for further phenotypic analysis. Although two mutants, *ran1-1* and *ran1-2*, with the T-DNA insertional sites located at the 5' promoter region of *RAN1* have been reported²³, *RAN1* expression was only partially downregulated

Fig. 3 | *TFL1* mRNA and protein localization in developing seeds. **a**, In situ localization of *TFL1* expression in developing seeds at 4 (left two panels) and 4.5 (right two panels) DAP using the *TFL1* antisense or sense probe. Scale bars, 100 µm. **b**, GUS staining of a *TFL1:GUS* seed at 4 DAP. Scale bar, 100 µm. **c**, Expression pattern of *TFL1* in a seed from the public microarray data shown in the *Arabidopsis* eFP browser (<http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi?primaryGene=AT5G03840&dataSource=Seed&modelInput=Absolute>). The colour from yellow to red indicates the increased absolute signal values of gene expression retrieved from microarray data. **d**, Localization of TFL1-GFP in developing *tfl1-20 gTFL1-GFP* seeds at 4 DAP by cryosectioning. TFL1-GFP signal is observable in the SPE (upper panels). The lower panels are a close-up view of the SPE shown above. The corresponding images with enhanced fluorescence intensity are shown in Extended Data Fig. 3e. BF, bright-field image; Merge, merge of GFP and BF images. Scale bars, 50 µm. **e**, Analysis of TFL1-3HA localization by immunogold electron microscopy using anti-HA antibody in the peripheral endosperm of *tfl1-20 gTFL1-3HA* (left panel) versus wild-type (WT, right panel) seeds at 4 DAP. The middle panel is a higher magnification of the area within the box indicated in the left panel. Arrows indicate gold particles. Scale bars, 2 µm. **f**, TFL1 subcellular localization shown by cell-fractionation assay. Western blot analysis was conducted using anti-HA antibody to examine TFL1 protein in nuclear (N) or cytoplasmic (C) fractions extracted from *tfl1-20 gTFL1-3HA* young siliques. The nuclear fraction was tenfold in excess compared with the cytosol fraction. The RbcL stained with Ponceau S and immunoblot analysis using anti-histone 3 (H3) are used as the indicators for cytosol and nuclear fractions, respectively. iCW, cell wall of ii; ii, the innermost cell layer of the inner integument (endothelium cell); PSV, protein storage vacuoles. The experiments in **a,b,d–f** were repeated three times independently with similar results.

in these two mutants. Thus, we performed targeted mutagenesis of *RAN1* using the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated nuclease 9 (Cas9) technology²⁴. Since *RAN1* and *RAN2* were physically linked on

chromosome 5 (Extended Data Fig. 5a), we created *ran1-3* and *ran1-4* in the wild-type and *ran2-1* backgrounds, respectively. *ran1-3* and *ran1-4* contained a short deletion at the second exon and 1-bp thymine (T) insertion at the beginning of the fifth exon



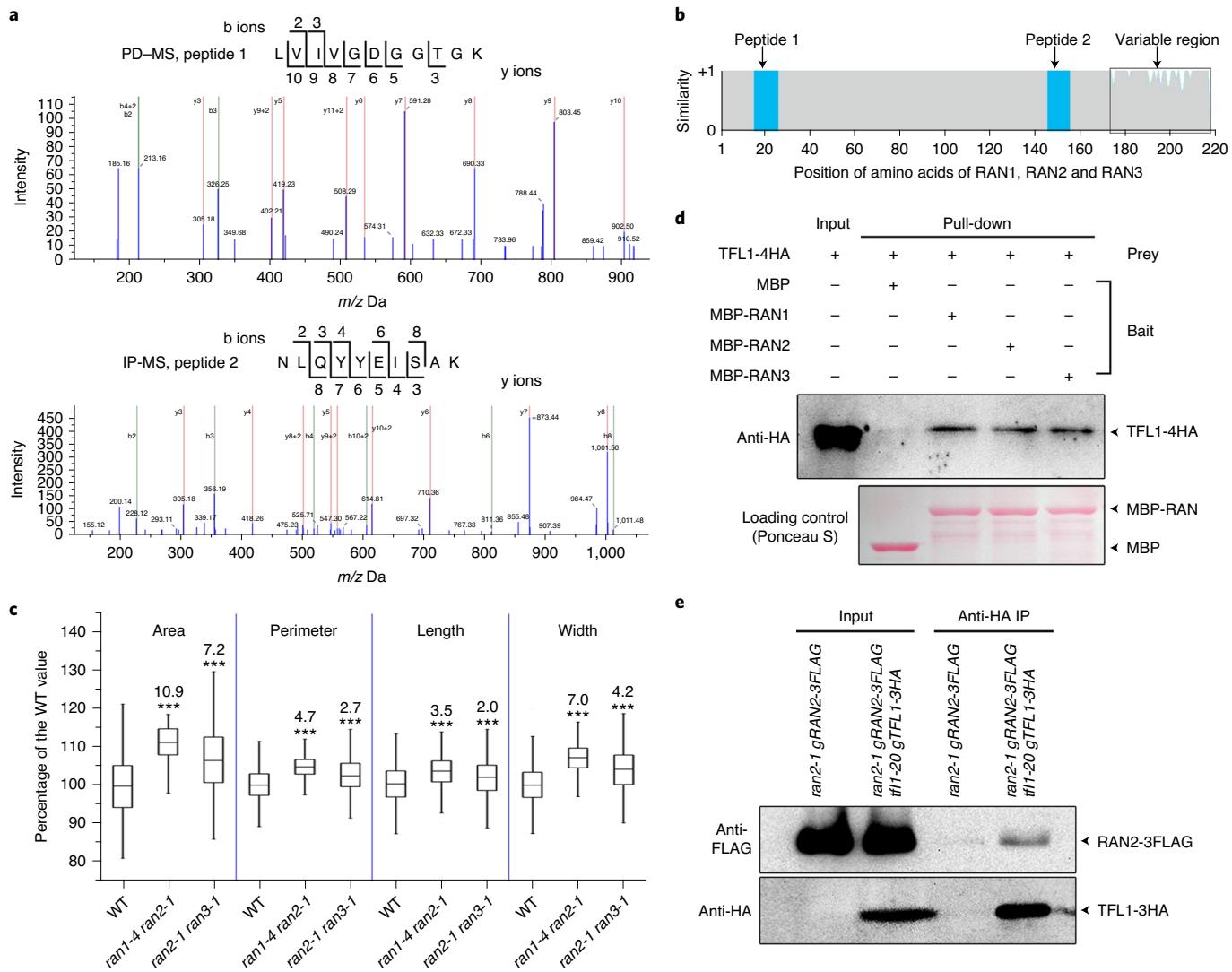


Fig. 4 | TFL1 interacts with RAN proteins. **a**, TFL1 interacting partners identified by a PD-MS or IP-MS approach. The identified RAN peptides with high confidence are shown above the corresponding mass spectra. **b**, Amino acid sequence similarity of RAN1, RAN2 and RAN3. The regions highlighted in blue are detectable by PD-MS (peptide 1) and IP-MS (peptide 2), as shown in **a**. The framed region indicates the variable C termini among RAN1, RAN2 and RAN3. The similarity level from 0 to +1 indicates the lowest to the highest sequence homology. **c**, Quantitative analysis of seed size parameters of *ran* double mutants. Box plots display median (lines), interquartile range (boxes) and whiskers (extending 1.5× the interquartile ranges) of seed size parameters of seeds of different genotypes (wild type (WT), $n=692$; *ran1-4 ran2-1*, $n=395$; *ran2-1 ran3-1*, $n=647$). Percentage change (%) in a seed parameter of double mutants is shown relative to the mean value of WT plants set as 100%. The increase in the percentage in a double mutant over WT is indicated above each box. Asterisks indicate significant differences between WT plants and other genotypes (two-tailed Mann-Whitney test, $P < 0.0001$). P value for seed length between WT and *ran2-1 ran3-1* is 4.44×10^{-11} , while the other P values are all less than 1×10^{-15} . **d**, Pull-down assay of the interaction between TFL1 and RAN proteins. MBP and MBP-RAN1/2/3 were used as baits, and the corresponding loading control was stained with Ponceau S (lower panel). The input of the prey protein TFL1-4HA extracted from 35S:TFL1-4HA siliques and its corresponding pulled-down signals were examined by immunoblot analysis using anti-HA antibody (upper panel). **e**, In vivo interaction between TFL1-3HA and RAN2-3FLAG in *Arabidopsis* siliques shown by CoIP. Total protein extracts from young siliques of *ran2-1 gRAN2-3FLAG* in WT or *tf1-20 gTFL1-3HA* background were immunoprecipitated by anti-HA antibody. The input and coimmunoprecipitated proteins were examined by immunoblot analysis using anti-HA (lower panel) or anti-FLAG (upper panel) antibody. The experiments in **d,e** were repeated three times independently with similar results.

in *RAN1*, respectively, resulting in disruption of the characteristic RAN domain²⁵ in *RAN1* protein (Extended Data Fig. 5a,c,d). *RAN2* and/or *RAN3* sequences corresponding to the CRISPR/Cas9 target sites in *RAN1* remained unaltered in *ran1-3* and *ran1-4 ran2-1* (Supplementary Fig. 3), indicating that CRISPR/Cas9-mediated target mutagenesis specifically modifies *RAN1* sequence in these mutants.

Among *ran1-3*, *ran2-1* and *ran3-1* single mutants, *ran2-1* showed weak changes in seed size parameters compared to wild-type seeds

(Extended Data Fig. 5e), while the seed size phenotypes of *ran2-1* were further enhanced in the background of *ran1-4* or *ran3-1* (Fig. 4c). Combination of these mutants failed to generate any homozygous triple mutants of *RAN1-3*, indicating their essential roles in either reproductive development or seed development. Furthermore, while *RAN1-3* were ubiquitously expressed in almost all tissues examined, *RAN2* expression levels were considerably higher than those of *RAN1* and *RAN3* (Extended Data Fig. 6a). These observations suggest that although *RAN1-3* plays a redundant

role in affecting seed size and possibly other developmental processes, *RAN2* could be a relatively important player. Thus, we chose *RAN2* as the representative *RAN* gene to study its role in mediating seed size in subsequent experiments. To understand *RAN2* protein expression in developing seeds, we created *GUS-gRAN2* reporter lines expressing a 4.9-kb *RAN2* genomic fragment containing the 2.2-kb 5' upstream sequence, the *GUS* reporter translationally fused to the 1.6-kb *RAN2* coding region and the 1.1-kb 3' downstream sequence. Most of the *GUS-gRAN2* reporter lines showed similar staining patterns in whole developing seeds (Extended Data Fig. 6b), including the SPE where TFL1 was localized.

RAN proteins mediate TFL1 transport to the SPE. We then performed a detailed analysis of the interaction between TFL1 and *RAN2* using various approaches. Pull-down assays demonstrated that TFL1-4HA extracted from 35S:TFL1-4HA siliques bound to in vitro-translated MBP-*RAN2* protein, but not MBP itself (Fig. 4d). The similar interaction was found between TFL1-4HA and MBP-*RAN1* or MBP-*RAN3* (Fig. 4d). Further pull-down assays of the interaction between TFL1 and various *RAN2* truncated proteins revealed that deletion of the effector-binding domain in *RAN2*, which is implicated in protein–protein interaction²⁵, partially compromised their interaction (Extended Data Fig. 7a,b). To examine the in vivo interaction between TFL1 and *RAN2*, we generated *ran2-1 gRAN2-3FLAG* transgenic plants in which 3FLAG fused with the same *RAN2* genomic fragment for *GUS-gRAN2* rescued the seed size phenotype of *ran2-1*. Coimmunoprecipitation (CoIP) analysis on total extracts from *ran2-1 gRAN2-3FLAG tfl1-20 gTFL1-3HA* siliques confirmed the in vivo interaction between *RAN2-3FLAG* and *TFL1-3HA* (Fig. 4e). These results indicate that TFL1 interacts with *RAN2* in seeds.

Compared to *tfl1-20* (Fig. 1b–d), single and double mutants of *RAN1-3* exhibited relatively weak seed size phenotypes (Fig. 4c and Extended Data Fig. 5e), while their homozygous triple mutants failed to survive. To investigate the possible redundant effects of *RAN* proteins on TFL1, we created *gDN-ran2-3FLAG* transgenic lines in which the threonine (T) at the residue 27 of *RAN2* was mutated into asparagine (N) in the 4.9-kb *RAN2* genomic fragment translationally fused with 3FLAG. This produced a dominant-negative version of *RAN2* (DN-*ran2*) (Extended Data Fig. 8a) that blocked the normal function of wild-type *RAN* proteins in binding to GTP as previously described^{25–27}. In addition, we also generated

TFL1:3FLAG-DN-ran2 transgenic plants expressing 3FLAG-tagged *DN-ran2* (Extended Data Fig. 8b) driven by the *TFL1* regulatory sequences used in *TFL1:GUS* (Fig. 1a) to test the in situ effect of *RAN* proteins on TFL1.

Notably, both *gDN-ran2-3FLAG* and *TFL1:3FLAG-DN-ran2* transgenic lines phenocopied *tfl1-20* pertaining to the main seed size parameters, producing much larger seeds than single and double mutants of *RAN1-3* (Figs. 4c and 5a and Extended Data Figs. 5e and 8c). Like *tfl1-20*, the enlarged seed cavity phenotype was also detectable in *gDN-ran2-3FLAG* and *TFL1:3FLAG-DN-ran2* from 4 DAP (Fig. 5b,c). Furthermore, compared to wild-type seeds, endosperm cellularization did not occur in most of the *gDN-ran2-3FLAG* and *TFL1:3FLAG-DN-ran2* seeds at 4 DAP (Extended Data Fig. 8d,e), demonstrating that the negative effects of *RANs* also influence endosperm cellularization, thus producing large seeds. It is noteworthy that although *gTFL1-3HA* and *gTFL1-GFP* fully rescued the seed size phenotype of *tfl1-20* (Fig. 1e), the rescue was completely abolished in the background of *TFL1:3FLAG-DN-ran2* as *tfl1-20 gTFL1-3HA TFL1:3FLAG-DN-ran2* and *tfl1-20 gTFL1-GFP TFL1:3FLAG-DN-ran2* still produced large seeds such as *tfl1-20* (Extended Data Fig. 9a,b). These results indicate that TFL1 function in controlling endosperm cellularization and seed size is dependent on the activity of *RAN* proteins.

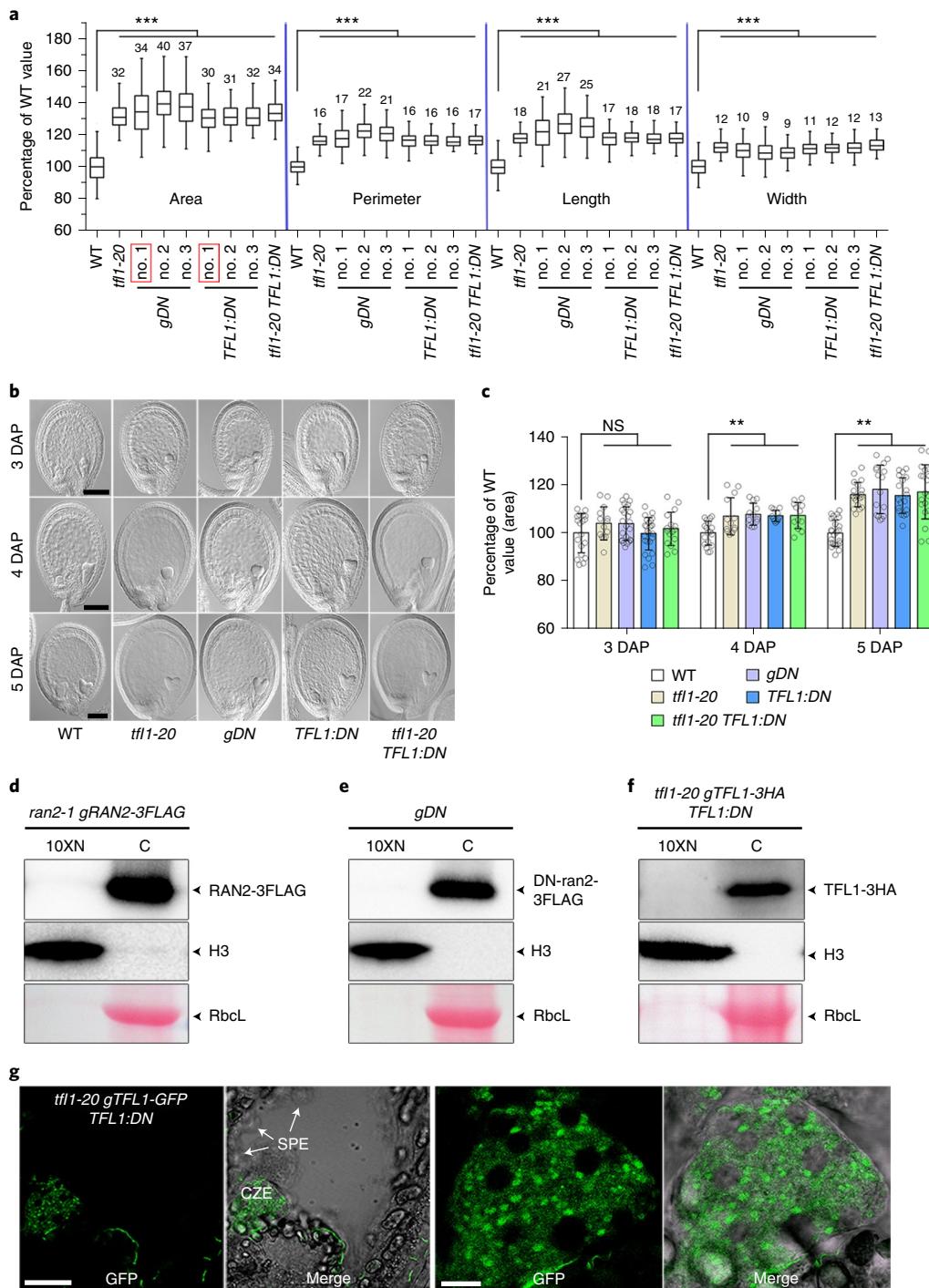
To investigate how *RAN* proteins influence TFL1 role in seed development, we performed cell-fractionation assays of *ran2-1 gRAN2-3FLAG* and *gDN-ran2-3FLAG* siliques, and found that both *RAN2-3FLAG* and its dominant-negative *DN-ran2* version were localized in the cytoplasm (Fig. 5d,e). *TFL1-3HA* remained exclusively in the cytoplasm regardless of the presence of the wild-type *RAN2* in *tfl1-20 gTFL1-3HA* (Fig. 3f) or the dominant-negative *DN-ran2* in *tfl1-20 gTFL1-3HA TFL1:3FLAG-DN-ran2* (Fig. 5f), suggesting that *RAN* proteins unlikely influence nucleocytoplasmic transport of TFL1. In contrast, confocal analysis of *tfl1-20 gTFL1-GFP TFL1:3FLAG-DN-ran2* revealed that *TFL1-GFP* was confined in the cytoplasm of the chalazal endosperm at the transition stage (Fig. 5g and Extended Data Fig. 9c), but not in the SPE as shown in *tfl1-20 gTFL1-GFP* (Fig. 3d). This demonstrates that dominant-negative *DN-ran2* suppresses TFL1 transport from the chalazal endosperm to the SPE. These observations overall indicate that *RAN* proteins interact with TFL1 in the chalazal endosperm and determine its transport to the SPE, thus affecting timely endosperm cellularization during seed development.

Fig. 5 | Dominant-negative *RAN2* (DN-*ran2*) phenocopies loss of function of TFL1 in seed development. **a**, Quantitative analysis of seed size parameters of wild-type (WT), *tfl1-20*, *gDN-ran2-3FLAG* (*gDN*), *TFL1:3FLAG-DN-ran2* (*TFL1:DN*) and *tfl1-20 TFL1:3FLAG-DN-ran2* (*tfl1-20 TFL1:DN*). Box plots display medians (lines), interquartile ranges (boxes) and whiskers (extending 1.5× the interquartile ranges) of seed size parameters (area, perimeter, length and width) of seeds of different genotypes (WT, $n=334$; *tfl1-20*, $n=313$; *gDN* nos. 1, 2, 3, $n=309, 245, 279$, respectively; *TFL1:DN* nos. 1, 2, 3, $n=279, 270, 251$, respectively; *tfl1-20 TFL1:DN*, $n=203$). Percentage change (%) in a seed parameter of a specific genotype is shown relative to the mean value of WT plants set as 100%. The increase in the percentage in a specific genotype over WT is indicated above each box. Asterisks indicate significant differences between WT plants and other genotypes (two-tailed Mann–Whitney test, $P < 0.0001$). P values are all less than 1×10^{-15} . The lines highlighted by the red rectangles were selected for further analysis. **b**, DIC microscopy of cleared whole-mount seeds of WT, *tfl1-20*, *gDN*, *TFL1:DN* and *tfl1-20 TFL1:DN* from 3 to 5 DAP. Scale bars, 100 μm . **c**, Comparison of seed cavity areas of WT, *tfl1-20*, *gDN*, *TFL1:DN* and *tfl1-20 TFL1:DN* at 3–5 DAP. Randomly selected seeds were examined at each time point (WT seeds at 3–5 DAP, $n=19, 18$ and 24, respectively; *tfl1-20* seeds at 3–5 DAP, $n=14, 13$ and 20, respectively; *gDN* seeds at 3–5 DAP, $n=24, 12$ and 16, respectively; *TFL1:DN* seeds at 3–5 DAP, $n=25, 11$ and 19, respectively; *tfl1-20 TFL1:DN* seeds at 3–5 DAP, $n=16, 12$ and 19, respectively). Values are mean \pm s.d. The WT mean value on each day is set as 100%. Asterisks indicate statistically significant differences between WT plants and other genotypes (two-tailed Student's *t*-test, $P < 0.01$). P values for WT versus *tfl1-20*, *gDN*, *TFL1:DN* and *tfl1-20 TFL1:DN* are 0.141, 0.0937, 0.916 and 0.503, respectively, at 3 DAP; 4.12×10^{-3} , 1.57×10^{-4} , 1.27×10^{-4} and 7.07×10^{-4} , respectively, at 4 DAP; and 9.33×10^{-13} , 7.27×10^{-9} , 8.16×10^{-10} and 7.05×10^{-8} , respectively, at 5 DAP. **d–f**, Subcellular localization of *RAN2* (**d**), *DN-ran2* (**e**) and *TFL1* (**f**) shown by cell-fractionation assays of the siliques at 4 DAP from specified genotypes. Western blot analysis was conducted using anti-FLAG or anti-HA antibody, respectively, to examine *RAN2*, *DN-ran2* and *TFL1* protein in nuclear (N) or cytoplasmic (C) fractions from these siliques. The nuclear fraction was tenfold in excess compared with the cytosol fraction. The *RbcL* stained with Ponceau S and immunoblot analysis using anti-histone 3 (H3) are used as the indicators for cytosol and nuclear fractions, respectively. **g**, Localization of *TFL1-GFP* in a *tfl1-20 gTFL1-GFP TFL1:DN* seed at 4 DAP by cryosectioning. *TFL1-GFP* signal is observable in the chalazal endosperm (CZE; left two panels) and its close-up view (right two panels). Note that *TFL1-GFP* is absent in the SPE. Merge, merge of GFP and bright-field images. Scale bars, 40 μm (left) and 10 μm (right). The experiments in **b,d–g** were repeated three times independently with similar results.

TFL1 interacts with ABI5 and affects ABI5 stability during seed development. To further elucidate how TFL1 affects endosperm cellularization, we sought to assess whether TFL1 function is relevant to the regulators in the IKU pathway^{2,6,7}. Previous studies have suggested that ABI5 directly represses *SHB1* expression, thus affecting endosperm cellularization via *SHB1-MINI3-IKU2* pathway^{8,10}. Loss of function of *ABI5* displayed correspondingly a large seed phenotype (Extended Data Fig. 10a). *ABI5* was ubiquitously expressed in all tissues examined, including the siliques from 2 to 7 DAP (Extended Data Fig. 10b), while the *ABI5-GUS* fusion protein was expressed in the whole seed at the transition stage at 4 DAP (Extended Data Fig. 10c). In addition,

TFL1 has been reported to interact with FD protein, which is the homologue of ABI5 in the A group of basic leucine zipper (bZIP) transcription factors^{28,29} (Extended Data Fig. 10d). These results prompted us to investigate whether TFL1 interacts with ABI5 to regulate seed size.

Pull-down assays showed His-ABI5 binding to in vitro-translated MBP-TFL1, but not MBP itself (Fig. 6a). To examine the interaction between TFL1 and ABI5 in *Arabidopsis*, we generated *gABI5-3FLAG* transgenic plants in which a 4.1-kb *ABI5* genomic fragment, including the 2.3-kb 5' upstream sequence, the 1.6-kb *ABI5* coding region translationally fused with *3FLAG*, and the 216-bp 3' downstream sequence, rescued the large seed phenotype of *abi5-1*. CoIP



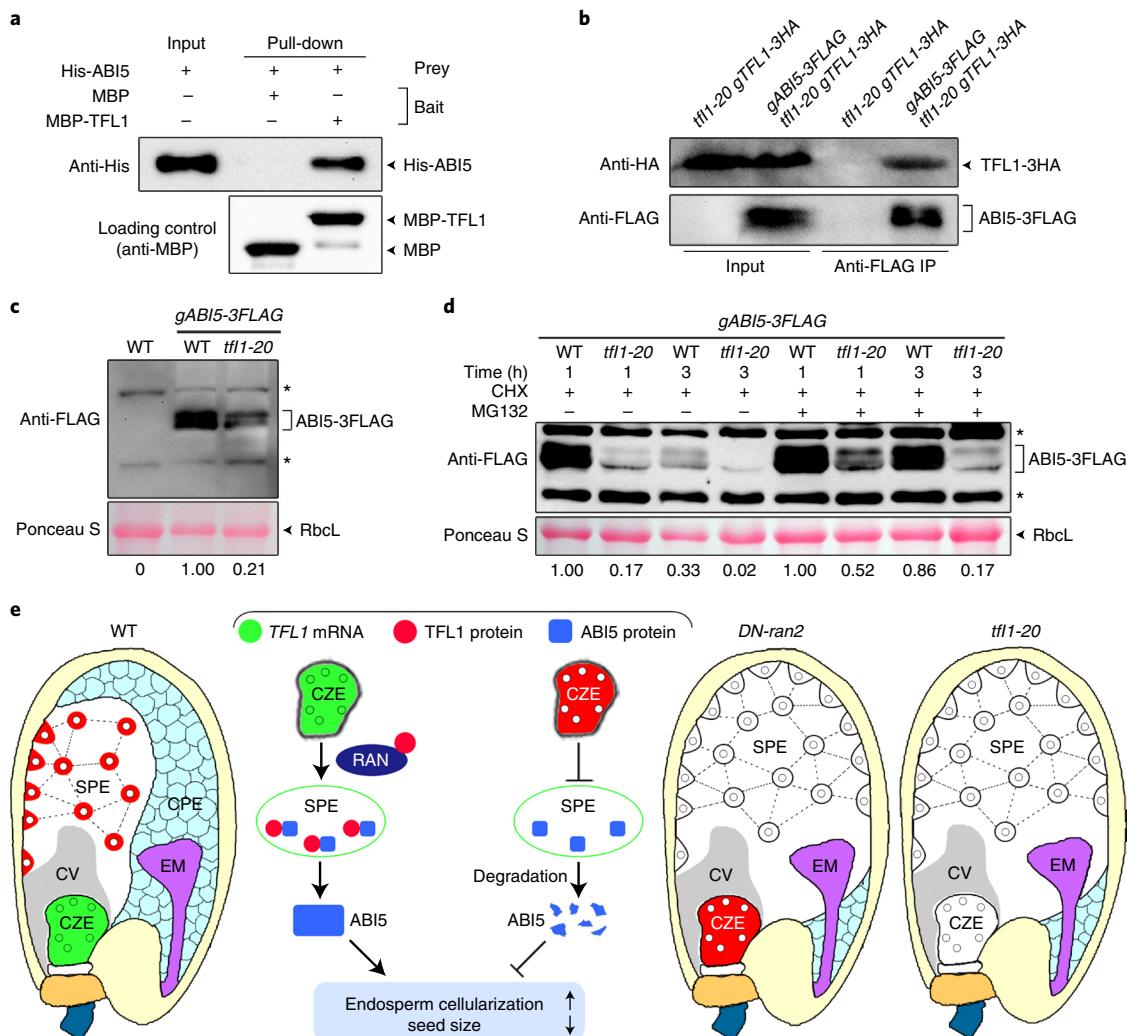


Fig. 6 | TFL1 regulates seed size partially through affecting ABI5 stability. **a**, In vitro pull-down assay of the interaction between TFL1 and ABI5. MBP and MBP-TFL1 were used as baits, and the corresponding loading control was examined by immunoblot analysis using anti-MBP antibody (lower panel). The input of the prey protein His-ABI5 and its corresponding pull-down signals were examined by immunoblot analysis using anti-His antibody (upper panel). **b**, In vivo interaction between TFL1-3HA and ABI5-3FLAG in *Arabidopsis* siliques shown by CoIP. Total protein extracts from young siliques of *tfl1-20 gTFL1-3HA* or *gABI5-3FLAG tfl1-20 gTFL1-3HA* were immunoprecipitated by anti-FLAG antibody. The input and coimmunoprecipitated proteins were examined by immunoblot analysis using anti-HA (upper panel) or anti-FLAG (lower panel) antibody. **c**, In vivo assay of ABI5 degradation in *gABI5-3FLAG* in wild-type (WT) or *tfl1-20* background. Total proteins extracted from the siliques at 4 DAP were examined by immunoblot analysis using anti-FLAG antibody. The number below each lane indicates the ABI5 protein level relative to the value in *gABI5-3FLAG* (WT) set as 1.00, which were all normalized against the RbcL stained with Ponceau S. Asterisks indicate non-specific bands. **d**, In vivo assay of ABI5 degradation in *gABI5-3FLAG* in wild-type (WT) or *tfl1-20* background under the treatment of MG132 or cycloheximide (CHX). The siliques at 4 DAP were treated without or with 50 µM of MG132 in the presence of 50 µM of CHX plus 0.012% Silwet L-77, and collected at the indicated time points. The number below each lane indicates the ABI5 protein level relative to the value in *gABI5-3FLAG* (WT) for 1 h of treatment (without or with MG132) set as 1.00, which were all normalized against RbcL stained with Ponceau S. Asterisks indicate non-specific bands. The experiments in **a-d** were repeated three times independently with similar results. **e**, A model depicting the regulation of seed size by the mobile TFL1 within the SPE. In wild-type (WT) seeds, RAN proteins interact with TFL1 and are required for TFL1 transport from the chalazal endosperm (CZE) to SPE, where TFL1 stabilizes ABI5, thus mediating timely cellularization of the endosperm and seed size. In dominant-negative *RAN2* (*DN-ran2*) or *TFL1* loss-of-function mutant (*tfl1-20*) seeds, lack of TFL1 in the SPE promotes ABI5 degradation, thus delaying endosperm cellularization and generating large seeds. Green or red CZE represents *TFL1* mRNA or protein expression in the CZE, respectively. In seeds, small circles in CZE or SPE represent nuclei.

analysis on total protein extracts from the siliques of *gABI5-3FLAG tfl1-20 gTFL1-3HA*, which was generated through genetic crossing of *gABI5-3FLAG* and *tfl1-20 gTFL1-3HA*, revealed the in vivo interaction of ABI5-3FLAG with TFL1-3HA (Fig. 6b), indicating that TFL1 interacts with ABI5 in developing seeds. Moreover, *tfl1-20 abi5-1* double mutants displayed a large seed phenotype such as *tfl1-20* (Extended Data Fig. 10a), supporting that TFL1 and ABI5 act in the same genetic pathway to control seed size.

Notably, we found that ABI5-3FLAG protein levels were much reduced in *tfl1-20* siliques compared to wild-type ones at 4 DAP (Fig. 6c), which is consistent with weaker ABI5-GUS signal observed in *tfl1-20* versus wild-type seeds (Extended Data Fig. 10c). As ABI5 protein is degraded in plant cells via the 26S proteasome pathway³⁰, we then tested whether the interaction between TFL1 and ABI5 may affect ABI5-3FLAG stability in *gABI5-3FLAG* and *tfl1-20 gABI5-3FLAG* siliques treated with cycloheximide (a protein

synthesis inhibitor) in the absence or presence of MG132 (a 26S proteasome inhibitor). Combination of cycloheximide and MG132 treatment for 3 h greatly increased the stability of ABI5-3FLAG in both wild-type and *tfl1-20* backgrounds, and evidently resulted in a less dramatic decline in ABI5-3FLAG levels in *tfl1-20* versus wild-type siliques than the cycloheximide treatment alone (Fig. 6d). These observations indicate that the interaction between TFL1 and ABI5 at least affects ABI5 stability mediated by the 26S proteasome pathway.

Discussion

Seed size is a critical agronomic trait that influences crop yield and seedling establishment, and is determined by coordinated growth of the endosperm, embryo and seed coat upon perceiving multiple developmental and environmental signals. After double fertilization in *Arabidopsis*, rapid proliferation of the syncytial endosperm through mitoses governs the increase in seed volume until endosperm cellularization occurs. Thus, the timing of endosperm cellularization determines the final seed size^{3,6}. Precocious cellularization results in reduced endosperm proliferation and small seed size, whereas delayed cellularization causes endosperm overproliferation and produces larger seeds^{6,7}.

The findings in this study provide several pieces of evidence that support an essential role of TFL1 as a mobile signal in regulating the timing of endosperm cellularization and seed size in *Arabidopsis*. First, compared to wild-type plants, loss of TFL1 function in *tfl1-20* produces larger developing seeds from the globular-heart transition stage, when endosperm cellularization occurs in the wild-type background, but is delayed in *tfl1-20* until the heart stage. These observations suggest that TFL1 plays an important role in mediating an appropriate timing of endosperm cellularization, which has a causal link with seed size as previously reported^{3,6,7}. Second, *TFL1* mRNA is present specifically in the chalazal endosperm, but not in the SPE at the transition stage, whereas its protein is mainly localized in the peripheral endosperm, indicating that TFL1 transport from the chalazal endosperm to the SPE is associated with its role in regulating endosperm cellularization. Third, TFL1 function in seed development is dependent on the activity of its interacting partners, which are small GTP-binding RAN proteins. The dominant-negative version of RAN2 protein interacts with TFL1 and suppresses TFL1 transport from the chalazal endosperm to the SPE, resulting in a large seed phenotype as shown in *tfl1-20*, suggesting that TFL1 movement to the SPE is indeed important for its function in mediating endosperm cellularization and seed size.

In eukaryotic cells, small GTP-binding proteins, which contain the motifs for binding to GDP/GTP and GTPase activity, are structurally divided into five subfamilies, including the Ran subfamily³¹. Previous studies have indicated the characteristic function of Ran GTPases in mediating nucleocytoplasmic transport and spindle bipolarization during mitosis in animals^{32,33}. In plants, although several studies have implicated the involvement of RANs in nucleocytoplasmic transport^{33,34,35}, how this is relevant to in vivo functions of RANs in plant development is still largely unknown. Our findings reveal that both RAN2 and TFL1 localize in the cytoplasm, and that their interaction is required for TFL1 transport within the multinucleate syncytial endosperm specifically from the chalazal to peripheral endosperm. These results indicate that mechanisms underlying RAN2-mediated TFL1 trafficking in the cytoplasmic network could be different from those known for RAN function in nucleocytoplasmic trafficking. Further identification of RAN2 interacting partners in seeds could help to elucidate TFL1 trafficking machinery in the syncytial endosperm.

Notably, compared to other known regulators involved in endosperm cellularization⁷, TFL1 exhibits a unique mRNA expression pattern only in the chalazal endosperm, but its protein is required for stabilizing ABI5 in the SPE to confer an appropriate timing of

endosperm cellularization. The chalazal endosperm, which is adjacent to the chalazal proliferating tissue of seed integument and links with the rest of the syncytial endosperm, could serve as a key passage to mediate the flux of signals or nutrients from the maternal tissues to the whole endosperm and embryo^{12–14,36–38}. Thus, generation of TFL1 in the chalazal endosperm and its mobile protein may allow mother plants to mediate seed development in response to the uptake of maternal signals, which possibly include ABA, flavonoids and sugar produced in maternal tissues^{7,39,40}. Indeed, ABA has been suggested to promote endosperm cellularization and regulate the final seed size via ABI5-related bZIP transcription factors¹⁰. Our finding of the effect of TFL1 on stabilizing ABI5, together with a maternal effect of TFL1 on seed size (data not shown), strongly indicates that the mobile TFL1 protein could be another hitherto unknown maternal signal that regulates endosperm cellularization to determine seed size according to the status of mother plants.

Methods

Plant materials and growth conditions. *A. thaliana* plants were grown under long day conditions (16 h light/8 h dark) at 22 °C. The mutants *ran2-1* (SALK_124236), *ran3-1* (SALK_125153), *ran3-2* (SALK_078740C), *tfl1-1* and *tfl1-20* (GABI_221D07) are in the Columbia (Col-0) background. The Col-0 near-isogenic *abi5-1* line was obtained by three backcrosses of *abi5-1* in the Ws background into the Col-0 wild-type plants. All transgenic plants were generated in the Col-0 background through *Agrobacterium tumefaciens*-mediated transformation.

Plasmid construction. To construct *gTFL1*, the 8.41-kb *TFL1* genomic fragment was amplified and inserted into the pENTR entry vector. On the basis of this construct, *Xma*I site was introduced before the stop codon using a site-directed mutagenesis approach⁴¹. The coding sequences of 3HA and GFP were amplified and cloned into *gTFL1* to generate *gTFL1-3HA* and *gTFL1-GFP*, respectively. To construct *TFL1:GUS* and *TFL1:3FLAG-DN-ran2*, the 5' and 3' promoter regions of *TFL1* (−2,745 to −1 and +1,037 to +5,667, respectively) were successively inserted into the pENTR entry vector. This backbone was further modified with the incorporation of *GUS* and the dominant-negative version of *RAN2* (*3FLAG-DN-ran2*), which was generated by the site-directed mutagenesis approach⁴¹ on the basis of the *3FLAG-RAN2* coding sequence. To construct *3S:TFL1-4HA*, the coding sequence of *TFL1* was cloned into the 35S-pENTR-4HA vector. To construct *GUS-gRAN2* or *gRAN2-3FLAG*, the 4.9-kb *RAN2* genomic fragment fused to either *GUS* or *3FLAG* gene was cloned into pENTR, respectively. *gDN-ran2-3FLAG* was generated based on *gRAN2-3FLAG* by the site-directed mutagenesis approach⁴¹. To construct *gABI5-GUS* and *gABI5-3FLAG*, the 4.1 kb *ABI5* genomic fragment fused to *GUS* or *3FLAG* was cloned into pENTR, respectively. The created entry constructs were then introduced into the binary vector pBGW through the Gateway LR recombination reaction (Invitrogen). The primers for creating the above constructs for plant transformation are listed in Supplementary Table 1.

Seed size measurement. Mature dry seeds were photographed under a stereomicroscope (Nikon). The seed area, perimeter, length and width were measured using the ImageJ 1.49g software. Box plots were prepared using the OriginPro 8 software, while overlaid dot plots in bar graphs were prepared using the GraphPad Prism 7 software. Two-tailed Student's *t*-test and two-tailed Mann-Whitney test were performed using the GraphPad Prism 7 software.

Differential interference contrast (DIC) microscopy. To observe whole-mount seeds, whole siliques were cut at both ends and fixed in ethanol:acetic acid (9:1) solution under vacuum for 10 min followed by overnight incubation at 4 °C. The samples were then washed with 90 and 70% ethanol for 30 min each. The chloral hydrate/glycerol/water solution (2 ml of glycerol, 4 ml of water and 8 g of chloral hydrate) was used to clear the samples before being visualized by DIC microscopy. To observe mature embryos, dry seeds were imbibed at room temperature for 1 h and dissected to isolate mature embryos, which were subsequently incubated in a solution (10 mM of EDTA, 1% Triton X-100, 1% dimethyl sulfoxide and 50 mM of sodium phosphate, pH 7) at 37 °C overnight. The embryos were fixed with 5% acetic acid, 10% formalin, 0.01% Triton X-100 and 45% ethanol for 50 min, followed by washes with a graded series of ethanol (30, 40, 50, 60, 70 and 100%). Embryos were cleared and visualized by DIC microscopy.

Confocal microscopy of whole-mount endosperm cellularization. Developing seeds were fixed in 4% glutaraldehyde in 12.5 mM of dimethyl arsenite (pH 6.9) under vacuum conditions for 30 min, then incubated overnight at 4 °C. After dehydration with a graded series of ethanol (30, 40, 50, 60, 70 and 100%) for 30 min each, the samples were then cleared with benzoate and benzyl alcohol (v/v 2:1) for 2 h, and observed under a confocal microscope (Zeiss) with an excitation wavelength at 488 nm.

Histological analysis. Tissues were fixed with 4% paraformaldehyde (w/v) overnight. The fixed materials were dehydrated and embedded in paraffin. Microtome sections (8 µm thickness) were washed thrice with sterile water for 20 s each, immersed in 0.05% toluidine blue O in 0.1 M of Na₂PO₄ (pH 6.8) for 10 s, and followed by two washes with sterile water. Subsequently, the sections were differentiated in 95% ethanol for several seconds. The slides were then dehydrated by washing once in 70% ethanol and twice in 100% ethanol. The air-dried slides were mounted with 50% glycerol and visualized by DIC microscopy.

Gene expression analysis. Total RNA from various *Arabidopsis* tissues was isolated using FavorPrep Plant Total RNA Mini Kit (Favorgen) and was reverse transcribed with M-MLV Reverse Transcriptase (Promega) according to the manufacturer's instructions. Quantitative real-time PCR reactions were done on the CFX384 real-time PCR detection system (Bio-Rad) with three biological replicates. The expression of *U-BOX* (*AT5G15400*) was used as an internal control⁴². The primers for quantitative real-time PCR are summarized in Supplementary Table 1.

GUS staining. GUS staining was performed as previously reported⁴². Siliques were cut at both ends and fixed in 90% acetone solution for 30 min on ice, then submerged in GUS staining solution under vacuum conditions for 30 min. The samples were incubated in the staining solution at 37°C, after which they were immersed in ethanol to remove chlorophyll. Finally, the stained tissues were placed in the clear solution and visualized by DIC microscopy.

RNA in situ hybridization. Non-radioactive in situ hybridization was performed as previously reported⁴³. RNA probe for *TFL1* was generated as described previously¹⁹. Briefly, the gene-specific region for *TFL1* was amplified and cloned into the pGEM-T Easy vector (Promega) to produce templates for in vitro transcription by DIG RNA Labeling Kit (Roche).

Cryosectioning. Siliques were fixed with ice-cold PFA (4% paraformaldehyde in PBS, pH 7.0) under vacuum for 1 h. Samples were then incubated in freshly prepared 4% paraformaldehyde overnight at 4°C, after which they were gently shaken at room temperature for 1 h, followed by washing with PBS buffer. The samples were subsequently incubated in 30% sucrose in PBS overnight with gentle shaking before being immobilized onto the microtome tissue holder using Tissue Tek OCT compound and immediately frozen with liquid nitrogen. Cryosections of 20 µm thick were made as previously described⁴⁴ with a Leica Sliding Microtome (Leica CM 3050 S) at -25°C and dried on a 42°C heating plate for 2 h. The slides were mounted with PBS solution and examined under confocal microscope (Zeiss).

Immunogold transmission electron microscopy. Immunoelectron microscopy was carried out as previously described⁴⁵. The grids containing ultra-thin sections (85 nm) were incubated with anti-HA antibody at 1:5 (v/v) for 1 h at room temperature. After two washes with tris-buffered saline and polysorbate 20 (20 mM of Tris, 500 mM of NaCl and 0.05% Tween 20 (pH 7.5)), the grids were incubated with 15 nm gold-conjugated goat anti-mouse antibody (25132, EMS, 1:20 dilution) for 40 min, followed by two washes with tris-buffered saline and polysorbate 20 and water, respectively. The samples were observed under transmission electron microscope (Jeol JEM-1230).

Pull-down assay. For LC-MS/MS analysis, the total protein extract from Col-0 wild-type siliques was incubated with the recombinant MBP-TFL1 immobilized onto amylose resin for 4 h at 4°C. After washing, the pulled-down mixture was eluted and analysed by LC-MS/MS using a TripleTOF 5600 System (AB Sciex). For pull-down assay, the coding sequences of *TFL1*, *RAN1*, *RAN2* and *RAN3* were amplified and cloned into the pMAL-c2x vector to generate MBP-tagged proteins. The *ABI5* coding sequence was amplified and cloned into the pET-28b vector to generate the His-tagged protein. The primers used for these constructs are listed in Supplementary Table 1. The resulting constructs and the empty pMAL-c2x and pET-28b vectors were transformed into *E. coli* strain Rosetta (DE3). The soluble fusion proteins were induced, extracted and immobilized on amylose beads (NEB) for subsequent pull-down assays. *TFL1-4HA* fusion protein was extracted from 35S:*TFL1-4HA* siliques. The HA-tagged protein was then incubated with immobilized MBP and MBP fusion proteins. Proteins retained on the beads were subsequently resolved by SDS-PAGE and detected with various antibodies (anti-MBP-HRP: E8038S, NEB, 1:5,000 dilution; anti-HIS: 28004, NewEast Biosciences, 1:5,000; anti-HA: sc-7392, Santa Cruz, 1:1,000 dilution).

CoIP assay. Total proteins were extracted from siliques in various genetic backgrounds by extraction buffer (50 mM of Tris-HCl (pH 7.5), 150 mM of NaCl, 1 mM of EDTA, 5% glycerol, 0.5% Triton X-100, 1 mM of PMSF, proteinase inhibitor cocktail). The protein extracts were then incubated with anti-HA agarose beads (A2095, Sigma) or anti-FLAG magnetic beads (M8823, Sigma) at 4°C for 4 h, and washed eight times with extraction buffer. The immunoprecipitated proteins and total protein extracts as inputs were resolved by SDS-PAGE and detected by anti-HA HRP (anti-HA: sc-7392 HRP, Santa Cruz, 1:1,000 dilution) or anti-FLAG antibody (anti-FLAG M2: F3165, Sigma, 1:5,000 dilution) followed by anti-mouse IgG-HRP (SC-516102, Santa Cruz, 1:5,000 dilution).

Cell-fractionation assay. Siliques were ground to a fine powder in liquid nitrogen and lysed with fractionation buffer (20 mM of Tris-HCl (pH 7.5), 250 mM of sucrose, 25% glycerol, 20 mM of KCl, 2 mM of EDTA, 2.5 mM of MgCl₂, 30 mM of beta-mercaptoethanol, protease inhibitor cocktail and 0.7% Triton X-100). The resulting extract was filtered with a 100-µm cell strainer to remove tissue debris before being centrifuged at 3,000g for 10 min. The supernatant was kept as the cytosolic fraction, while the pellet was washed with resuspension buffer (20 mM of Tris-HCl (pH 7.5), 20 mM of KCl, 2 mM of EDTA, 2.5 mM of MgCl₂, 30 mM of beta-mercaptoethanol, protease inhibitor cocktail) until the pellet was no longer green. The white pellet was resuspended in the resuspension buffer to obtain the nuclear fraction. The rubisco large subunit (RbcL) stained with Ponceau S and immunoblot analysis using anti-histone 3 (AB1791, Abcam, 1:3,000 dilution) followed by mouse anti-rabbit IgG-HRP (SC-2357, Santa Cruz, 1:5,000 dilution) were used as the indicators for cytosol and nuclear fractions, respectively.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

We declare that all data supporting the findings of this study are available within the article and its supplementary information files or from the corresponding author upon reasonable request. Source data are provided with this paper.

Received: 4 January 2020; Accepted: 21 July 2020;

Published online: 24 August 2020

References

- Moles, A. T. et al. A brief history of seed size. *Science* **307**, 576–580 (2005).
- Zhou, Y. et al. SHORT HYPOCOTYL UNDER BLUE1 associates with *MINISEED3* and *HAIKU2* promoters in vivo to regulate *Arabidopsis* seed development. *Plant Cell* **21**, 106–117 (2009).
- Boisnard-Lorig, C. et al. Dynamic analyses of the expression of the HISTONE::YFP fusion protein in *Arabidopsis* show that syncytial endosperm is divided in mitotic domains. *Plant Cell* **13**, 495–509 (2001).
- Borisjuk, L. et al. Seed architecture shapes embryo metabolism in oilseed rape. *Plant Cell* **25**, 1625–1640 (2013).
- Berger, F., Grini, P. E. & Schnittger, A. Endosperm: an integrator of seed growth and development. *Curr. Opin. Plant Biol.* **9**, 664–670 (2006).
- Sun, X., Shantharaj, D., Kang, X. & Ni, M. Transcriptional and hormonal signaling control of *Arabidopsis* seed development. *Curr. Opin. Plant Biol.* **13**, 611–620 (2010).
- Orozco-Arroyo, G., Paolo, D., Ezquer, I. & Colombo, L. Networks controlling seed size in *Arabidopsis*. *Plant Reprod.* **28**, 17–32 (2015).
- Kang, X., Li, W., Zhou, Y. & Ni, M. A WRKY transcription factor recruits the SYG1-like protein SHB1 to activate gene expression and seed cavity enlargement. *PLoS Genet.* **9**, e1003347 (2013).
- Luo, M., Dennis, E. S., Berger, F., Peacock, W. J. & Chaudhury, A. *MINISEED3* (*MINI3*), a WRKY family gene, and *HAIKU2* (*IKU2*), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **102**, 17531–17536 (2005).
- Cheng, Z. J. et al. Abscisic acid regulates early seed development in *Arabidopsis* by *ABI5*-mediated transcription of SHORT HYPOCOTYL UNDER BLUE1. *Plant Cell* **26**, 1053–1068 (2014).
- Olsen, O.-A. Nuclear endosperm development in cereals and *Arabidopsis thaliana*. *Plant Cell* **16**, S214–S227 (2004).
- Otegui, M. S., Capp, R. & Staehelin, L. A. Developing seeds of *Arabidopsis* store different minerals in two types of vacuoles and in the endoplasmic reticulum. *Plant Cell* **14**, 1311–1327 (2002).
- Nguyen, H., Brown, R. & Lemmon, B. The specialized chalazal endosperm in *Arabidopsis thaliana* and *Lepidium virginicum* (Brassicaceae). *Protoplasma* **212**, 99–110 (2000).
- Brown, R., Lemmon, B. E., Nguyen, H. & Olsen, O.-A. Development of endosperm in *Arabidopsis thaliana*. *Sex. Plant Reprod.* **12**, 32–42 (1999).
- Baroux, C., Fransz, P. & Grossniklaus, U. Nuclear fusions contribute to polyploidization of the gigantic nuclei in the chalazal endosperm of *Arabidopsis*. *Planta* **220**, 38–46 (2004).
- Brown, R. C., Lemmon, B. E. & Nguyen, H. Comparative anatomy of the chalazal endosperm cyst in seeds of the Brassicaceae. *Bot. J. Linn. Soc.* **144**, 375–394 (2004).
- Bradley, D., Ratcliffe, O., Vincent, C., Carpenter, R. & Coen, E. Inflorescence commitment and architecture in *Arabidopsis*. *Sci.* **275**, 80–83 (1997).
- Ratcliffe, O. J. et al. A common mechanism controls the life cycle and architecture of plants. *Development* **125**, 1609–1615 (1998).
- Liu, C. et al. A conserved genetic pathway determines inflorescence architecture in *Arabidopsis* and rice. *Dev. Cell* **24**, 612–622 (2013).
- Vernoud, V., Horton, A. C., Yang, Z. & Nielsen, E. Analysis of the small GTPase gene superfamily of *Arabidopsis*. *Plant Physiol.* **131**, 1191–1208 (2003).

21. Rosso, M. G. et al. An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol. Biol.* **53**, 247–259 (2003).
22. Conti, L. & Bradley, D. TERMINAL FLOWER1 is a mobile signal controlling *Arabidopsis* architecture. *Plant Cell* **19**, 767–778 (2007).
23. Liu, P. et al. *Arabidopsis* RAN1 mediates seed development through its parental ratio by affecting the onset of endosperm cellularization. *Mol. Plant* **7**, 1316–1328 (2014).
24. Feng, Z. et al. Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **111**, 4632–4637 (2014).
25. Haizel, T., Merkle, T., Pay, A., Fejes, E. & Nagy, F. Characterization of proteins that interact with the GTP-bound form of the regulatory GTPase Ran in *Arabidopsis*. *Plant J.* **11**, 93–103 (1997).
26. Kim, H.-J., Yano, A., Wada, Y. & Sano, H. Properties of a tobacco DNA methyltransferase, NtMET1 and its involvement in chromatin movement during cell division. *Ann. Bot.* **99**, 845–856 (2007).
27. Kornbluth, S., Dasso, M. & Newport, J. Evidence for a dual role for TC4 protein in regulating nuclear structure and cell cycle progression. *J. Cell Biol.* **125**, 705–719 (1994).
28. Jakoby, M. et al. bZIP transcription factors in *Arabidopsis*. *Trends Plant Sci.* **7**, 106–111 (2002).
29. Abe, M. et al. FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* **309**, 1052–1056 (2005).
30. Yu, F., Wu, Y. & Xie, Q. Precise protein post-translational modifications modulate ABI5 activity. *Trends Plant Sci.* **20**, 569–575 (2015).
31. Takai, Y., Sasaki, T. & Matozaki, T. Small GTP-binding proteins. *Physiol. Rev.* **81**, 153–208 (2001).
32. Moore, M. S. & Blobel, G. The GTP-binding protein Ran/TC4 is required for protein import into the nucleus. *Nature* **365**, 661–663 (1993).
33. Clarke, P. R. & Zhang, C. Ran GTPase: a master regulator of nuclear structure and function during the eukaryotic cell division cycle? *Trends Cell Biol.* **11**, 366–371 (2001).
34. Wang, X. et al. Overexpression of *RAN1* in rice and *Arabidopsis* alters primordial meristem, mitotic progress, and sensitivity to auxin. *Plant Physiol.* **140**, 91–101 (2006).
35. Zang, A., Xu, X., Neill, S. & Cai, W. Overexpression of *OsRAN2* in rice and *Arabidopsis* renders transgenic plants hypersensitive to salinity and osmotic stress. *J. Exp. Bot.* **61**, 777–789 (2009).
36. Garcia, D. et al. *Arabidopsis haiku* mutants reveal new controls of seed size by endosperm. *Plant Physiol.* **131**, 1661–1670 (2003).
37. Karmann, J., Müller, B. & Hammes, U. Z. The long and winding road: transport pathways for amino acids in *Arabidopsis* seeds. *Plant Reprod.* **31**, 253–261 (2018).
38. Maitz, M. et al. *rgf1*, a mutation reducing grain filling in maize through effects on basal endosperm and pedicel development. *Plant J.* **23**, 29–42 (2000).
39. Frey, A., Godin, B., Bonnet, M., Sotta, B. & Marion-Poll, A. Maternal synthesis of abscisic acid controls seed development and yield in *Nicotiana plumbaginifolia*. *Planta* **218**, 958–964 (2004).
40. Antoni, R., Rodriguez, L., Gonzalez-Guzman, M., Pizzio, G. A. & Rodriguez, P. L. News on ABA transport, protein degradation, and ABFs/WRKYs in ABA signaling. *Curr. Opin. Plant Biol.* **14**, 547–553 (2011).
41. Chen, L., Wang, F., Wang, X. & Liu, Y.-G. Robust one-tube Ω-PCR strategy accelerates precise sequence modification of plasmids for functional genomics. *Plant Cell Physiol.* **54**, 634–642 (2013).
42. Li, C., Zhang, B., Chen, B., Ji, L. & Yu, H. Site-specific phosphorylation of TRANSPARENT TESTA GLABRA1 mediates carbon partitioning in *Arabidopsis* seeds. *Nat. Commun.* **9**, 571 (2018).
43. Liu, C. et al. Specification of *Arabidopsis* floral meristem identity by repression of flowering time genes. *Development* **134**, 1901–1910 (2007).
44. Liu, L., Li, C., Teo, Z. W. N., Zhang, B. & Yu, H. The MCTP-SNARE complex regulates florigen transport in *Arabidopsis*. *Plant Cell* **31**, 2475–2490 (2019).
45. Liu, L. et al. FTIP1 is an essential regulator required for florigen transport. *PLoS Biol.* **10**, e1001313 (2012).

Acknowledgements

We thank the Arabidopsis Biological Resource Centre for providing various mutants, and the members of the Yu laboratory for discussion and comments on the manuscript. This work was supported by the Singapore National Research Foundation Investigatorship Programme (NRFNRFI2016-02), the Agency for Science, Technology and Research (A*STAR) under its Industry Alignment Fund—Pre-Positioning (IAF-PP) (A19D9a0096), and intramural research support from the National University of Singapore and Temasek Life Sciences Laboratory.

Author contributions

B.Z. and H.Y. conceived and designed the project. B.Z., C.L. and Y.L. performed the experiments. B.Z. and H.Y. analysed the data and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41477-020-0749-5>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41477-020-0749-5>.

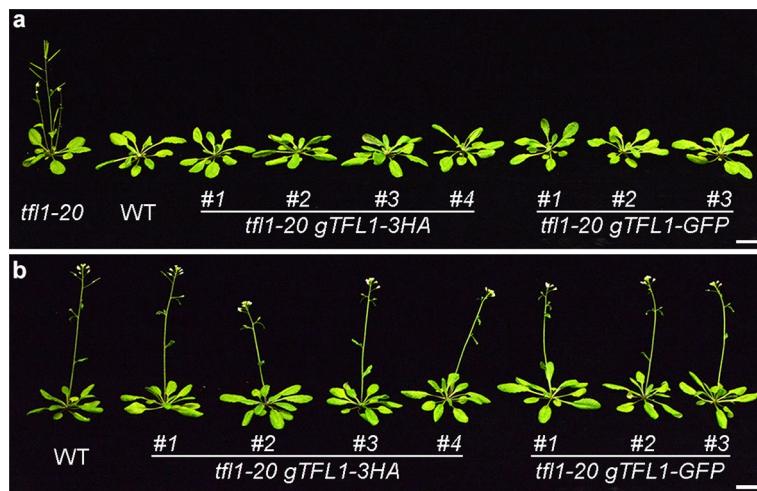
Correspondence and requests for materials should be addressed to H.Y.

Peer review information *Nature Plants* thanks Steven Penfield and Xian Sheng Zhang for their contribution to the peer review of this work.

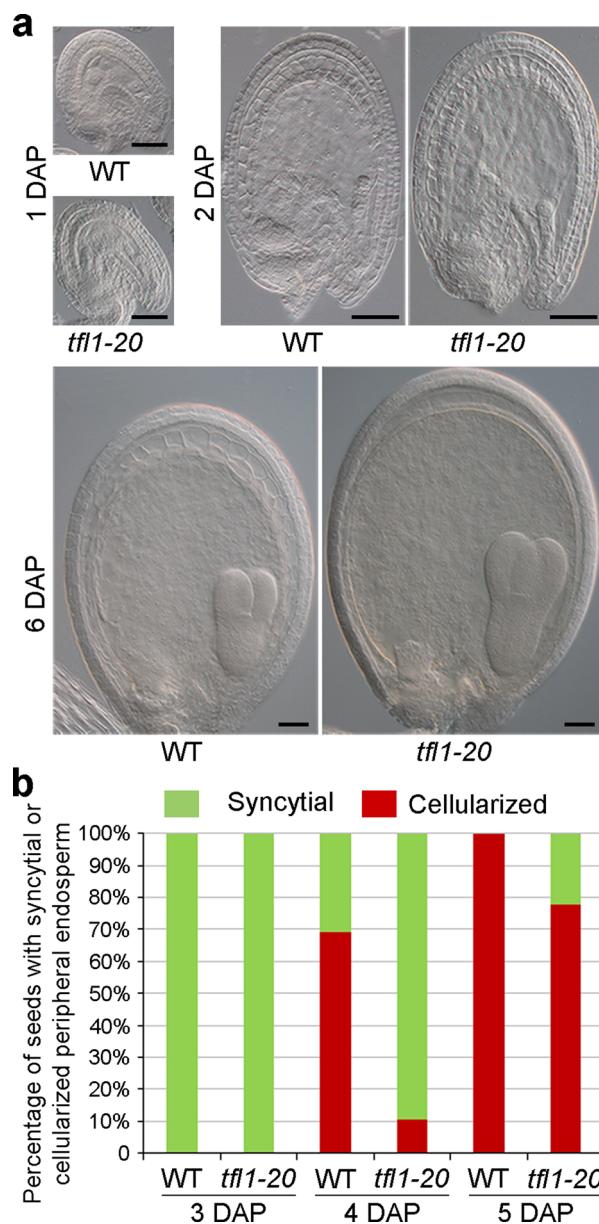
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

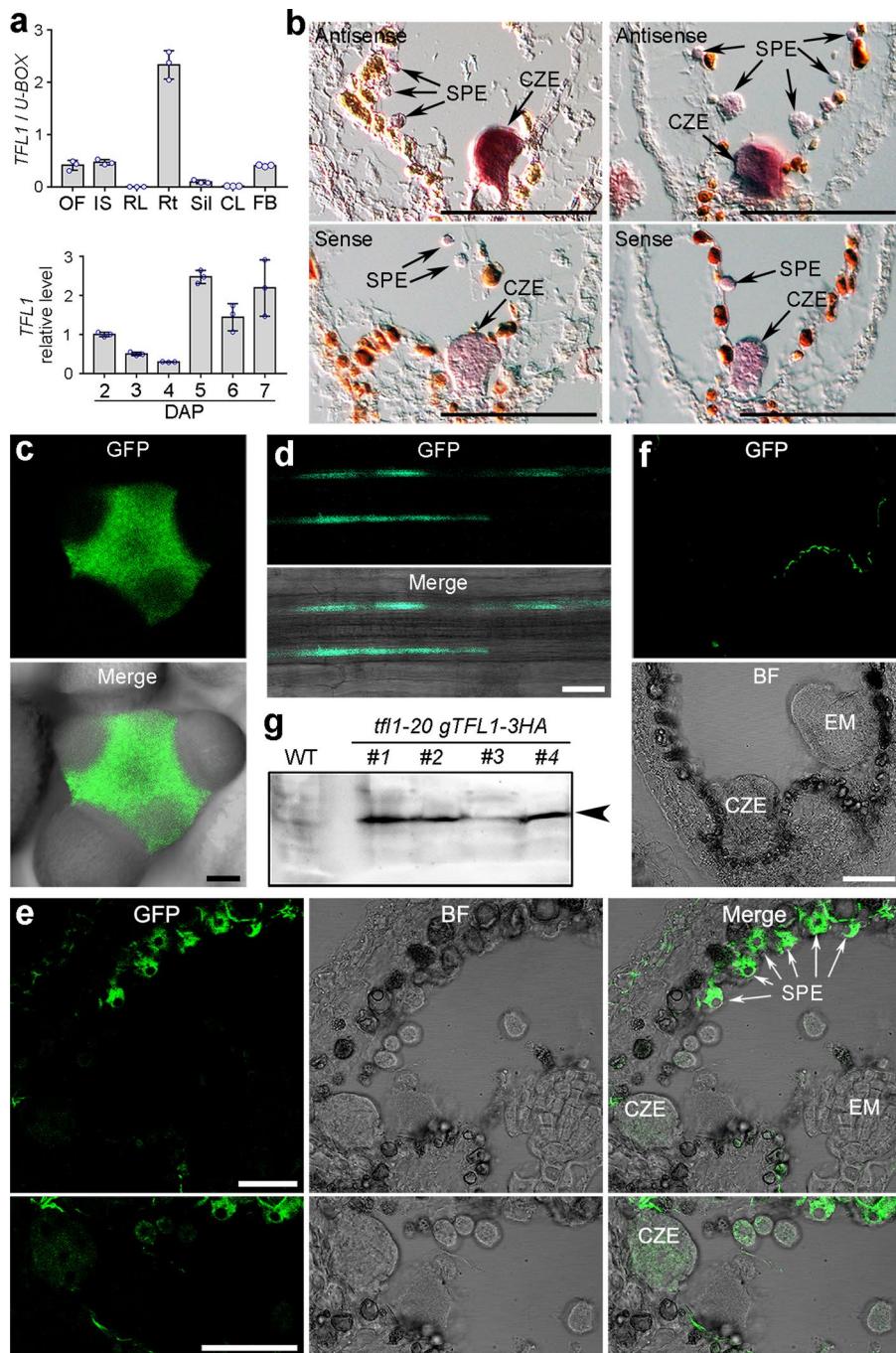
© The Author(s), under exclusive licence to Springer Nature Limited 2020



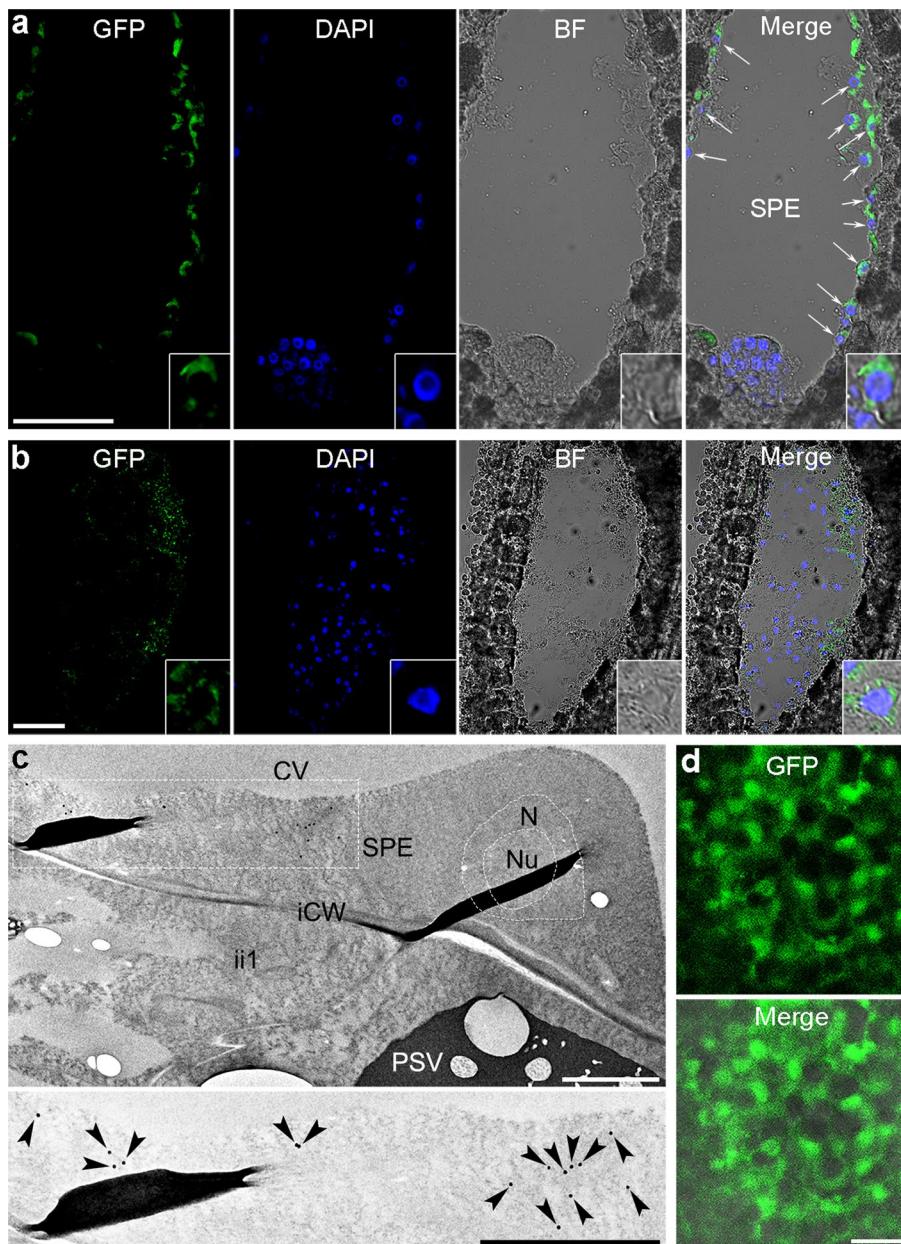
Extended Data Fig. 1 | *gTFL1-3HA* and *gTFL1-GFP* rescue the early flowering and terminal flower phenotypes of *tfi1-20*. **a**, *tfi1-20* flowers much earlier than a wild-type (WT) plant and independent *tfi1-20 gTFL1-3HA* or *tfi1-20 gTFL1-GFP* transgenic lines. **b**, Independent *tfi1-20 gTFL1-3HA* and *tfi1-20 gTFL1-GFP* transgenic lines develop normal inflorescence apices like a WT plant, whereas *tfi1-20* develops terminal flowers (**a**). Scale bars, 2 cm.



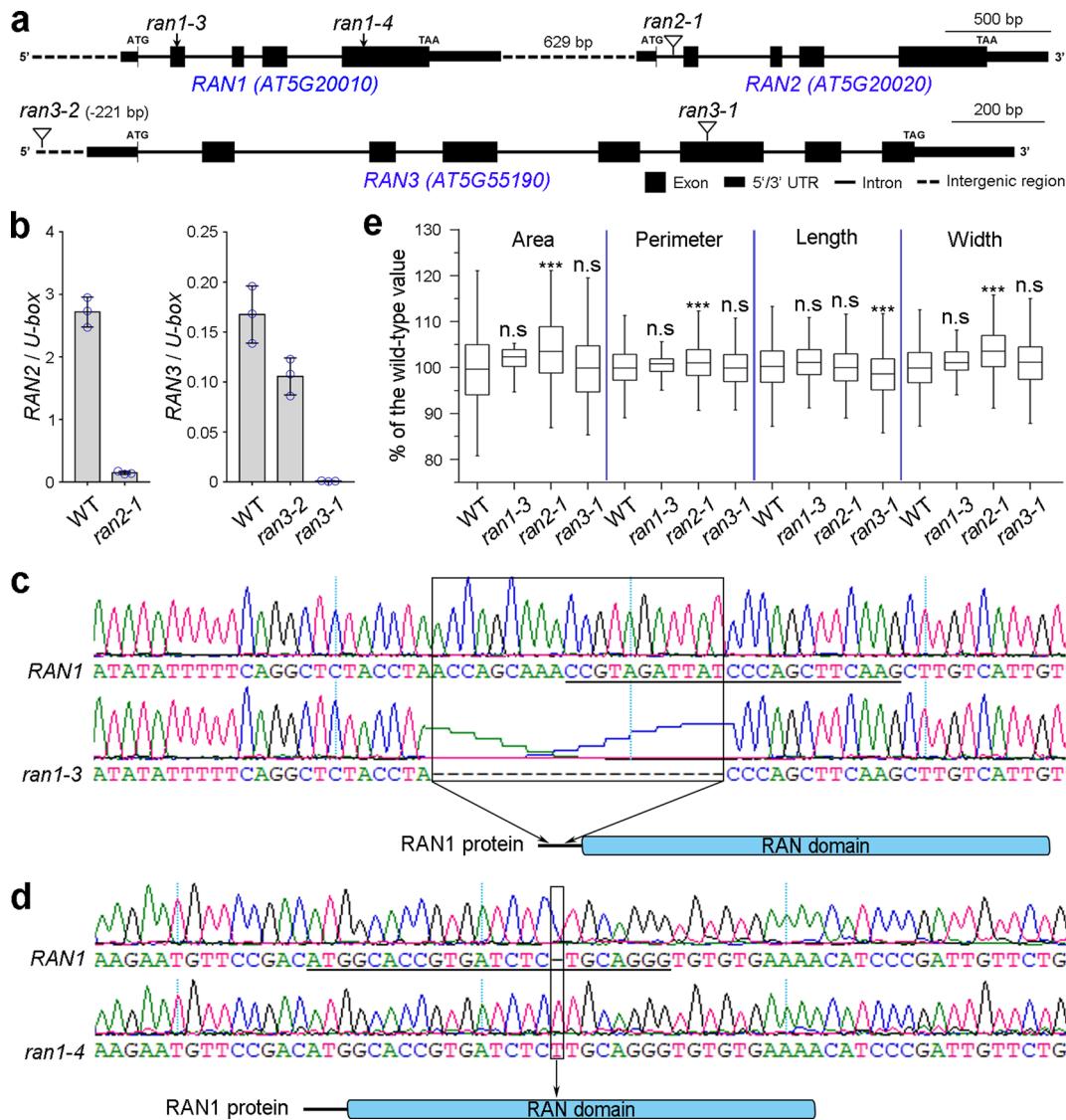
Extended Data Fig. 2 | TFL1 affects endosperm cellularization. **a**, Differential interference contrast (DIC) microscopy of cleared whole-mount seeds of wild-type (WT) and *tfi1-20* plants at 1, 2, and 6 days after pollination (DAP). Scale bars, 50 µm. The experiment was repeated three times independently with similar results. **b**, Percentage of seeds with syncytial or cellularized peripheral endosperm in WT and *tfi1-20* at 3 to 5 DAP examined by confocal microscopy. Randomly selected seeds from more than 10 siliques for each genotype were examined at each time point.



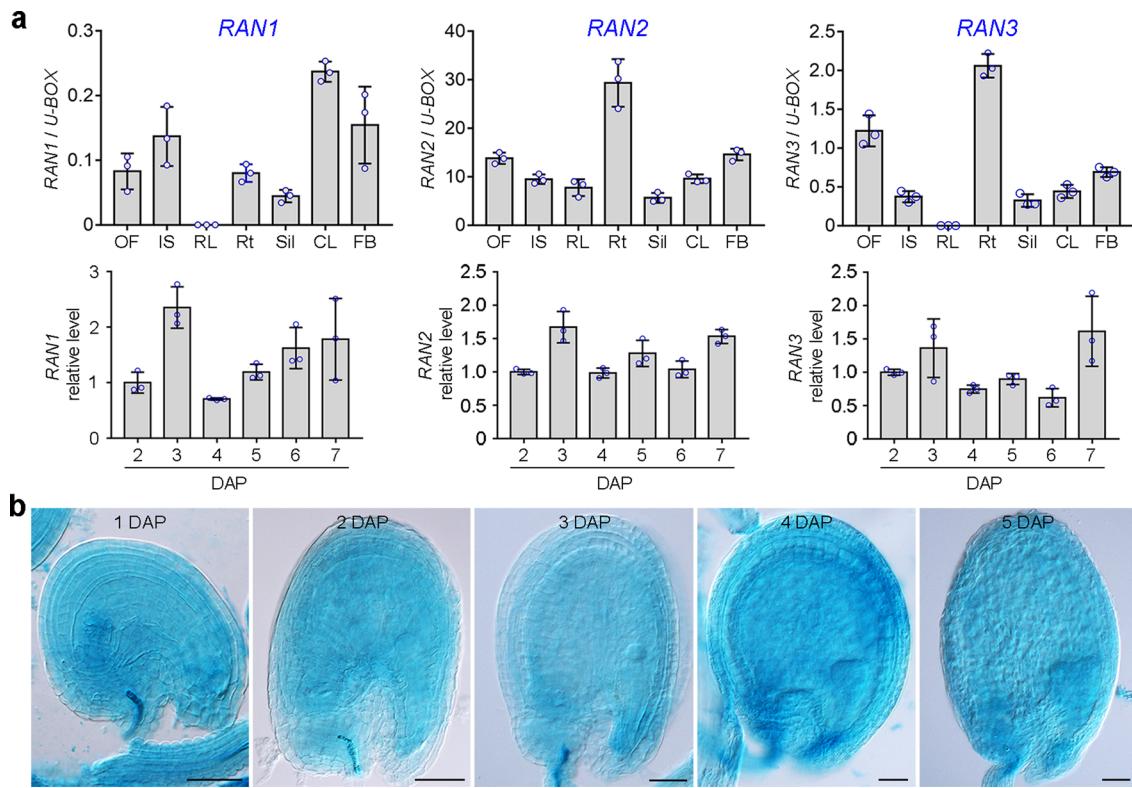
Extended Data Fig. 3 | *TFL1* mRNA and protein expression in *Arabidopsis*. **a**, Quantitative real-time PCR analysis of *TFL1* mRNA expression in various tissues (upper panel) or developing siliques at 2 to 7 days after pollination (DAP; lower panel) in *Arabidopsis*. OF, open flower; IS, inflorescence stem; RL, rosette leaf; Rt, root; Sil, siliques; CL, cauline leaf; FB, flower bud. Results were normalized against the expression levels of the *U-BOX* gene as an internal control. Expression levels in the lower panel are shown as relative values to the 2 DAP level set as 1. Values are mean \pm SD of three biological replicates. **b**, *In situ* localization of *TFL1* expression in developing seeds at 4 DAP using the *TFL1* antisense or sense probe. SPE, syncytial peripheral endosperm; CZE, chalazal endosperm. Scale bars, 100 μ m. **c, d**, Localization of *TFL1*-GFP in an inflorescence meristem (**c**) and a root (**d**) of *tfl1-20 gTFL1-GFP*. Merge, merge of GFP and bright field images. Scale bars, 20 μ m. **e**, Localization of *TFL1*-GFP in developing *tfl1-20 gTFL1-GFP* seeds at 4 DAP by cryosectioning under enhanced fluorescence intensity compared to those shown in Fig. 3d. *TFL1*-GFP signal is much weaker in the chalazal endosperm than in the syncytial peripheral endosperm (upper panel). The lower panel is a close-up view of the chalazal endosperm and its nearby syncytial peripheral endosperm shown above. BF, bright field image. CZE, chalazal endosperm; EM, embryo; Merge, merge of GFP and BF images; SPE, syncytial peripheral endosperm. Scale bars, 50 μ m. **f**, Detection of background green fluorescence signal in a wild-type seed at 4 DAP. This image, acquired under the same condition for Fig. 3d, serves as a negative control for Fig. 3d. The auto-fluorescence (green fluorescence) signal was slightly detectable. Scale bar, 50 μ m. **g**, Western blot analysis of *TFL1*-3HA protein (arrowhead) in selected *tfl1-20 gTFL1-3HA* transgenic lines that possibly contain one T-DNA insertion site based on their segregation ratios. Total protein was extracted from young siliques, and subjected to Western blot analysis using anti-HA antibody. The experiments in **b–g** were repeated three times independently with similar results.



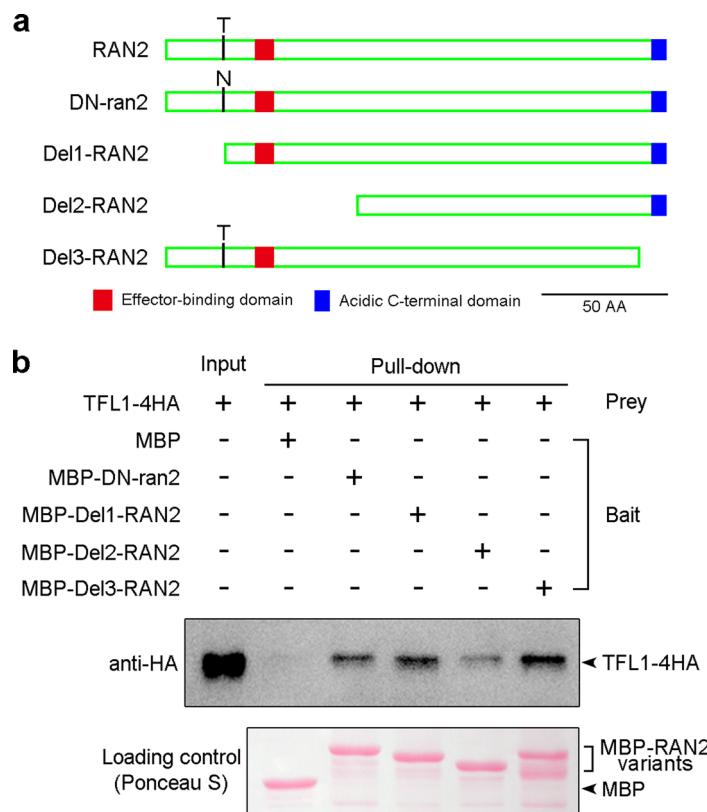
Extended Data Fig. 4 | Cytoplasmic localization of TFL1 in *Arabidopsis*. **a,b**, Localization of TFL1-GFP in developing *tfl1-20 gTFL1-GFP* seeds stained with DAPI at 4 (**a**) and 5 (**b**) days after pollination (DAP) by cryosectioning. TFL1-GFP signal in the syncytial (**a**) and cellularized (**b**) peripheral endosperm does not co-localize with DAPI staining. The inset in each panel in (**a**) or (**b**) shows one enlarged nucleus of the syncytial or cellularized peripheral endosperm, respectively. DAPI, fluorescence of 4,6-diamino-2-phenylindole; BF, bright-field image; Merge, merge of GFP, DAPI and BF images. Scale bars, 50 μ m. **c**, Analysis of TFL1-3HA cytoplasmic localization by immunogold electron microscopy using anti-HA antibody in the syncytial peripheral endosperm of *tfl1-20 gTFL1-3HA* seeds at 4 DAP. The lower panel shows a higher magnification of the area within the box indicated in the upper panel. Arrows indicate gold particles. There are no detectable gold particles in the nucleus. CV, central vacuole; iCW, cell wall of ii1; ii1, the innermost cell layer of the inner integument (endothelium cell); N, nucleus; Nu, nucleolus; PSV, protein storage vacuoles; SPE, syncytial peripheral endosperm. Scale bars, 2 μ m. **d**, Subcellular localization of TFL1-GFP in the cells of a *tfl1-20 gTFL1-GFP* inflorescence meristem. Merge, merge of GFP and bright field images. Scale bar, 10 μ m. The experiments in **a-d** were repeated three times independently with similar results.



Extended Data Fig. 5 | Isolation of *RAN1*, *RAN2*, and *RAN3* loss-of-function mutants. **a**, Schematic diagrams showing *RAN1*, *RAN2*, and *RAN3* genomic regions, and the CRISPR-Cas9 target sites (*ran1-3* and *ran1-4*) or T-DNA insertion sites (*ran2-1*, *ran3-1* and *ran3-2*). **b**, Quantitative analysis of *RAN2* or *RAN3* expression in their corresponding T-DNA insertion mutants. Results were normalized against the expression levels of the *U-BOX* gene as an internal control. Values are mean \pm SD of three biological replicates. **c,d**, CRISPR/Cas9-mediated target mutagenesis of *RAN1* in the wild-type (**c**) and *ran2-1* (**d**) background. The CRISPR/Cas9 target sites in *RAN1* are underlined in the upper panels. The newly created *ran1-3* single mutant (**c**) and *ran1-4 ran2-1* double mutant (**d**) contain a short deletion at the second exon and 1-bp thymine (T) insertion at the beginning of the fifth exon in *RAN1* (both highlighted by the black frame), respectively. The positions of the mutations in *ran1-3* and *ran1-4 ran2-1* in *RAN1* protein are indicated. **e**, Quantitative analysis of seed size parameters of *ran1-3*, *ran2-1* and *ran3-1*. Box plots show medians (lines), interquartile ranges (boxes), and whiskers (extending 1.5 times the interquartile ranges) of seed size parameters (area, perimeter, length, and width) of seeds of different genotypes (WT, $n=692$; *ran1-3*, $n=153$; *ran2-1*, $n=516$; *ran3-1*, $n=461$). Value represents the percentage change (%) in a seed parameter of a mutant relative to the mean value of WT plants set as 100%. Asterisks indicate significant differences between WT plants and other genotypes (two-tailed Mann-Whitney test, $P < 0.0001$). n.s., no statistical difference. P values for WT versus *ran1-3*, *ran2-1*, and *ran3-1* are 0.000193, $< 1 \times 10^{-15}$, and 0.461, respectively, for seed area; 0.0463, 4.378×10^{-7} , and 0.988, respectively, for seed perimeter; 0.0333, 0.800, and 2.43×10^{-6} , respectively, for seed length; and 0.000409, $< 1 \times 10^{-15}$, and 0.000164, respectively, for seed width.

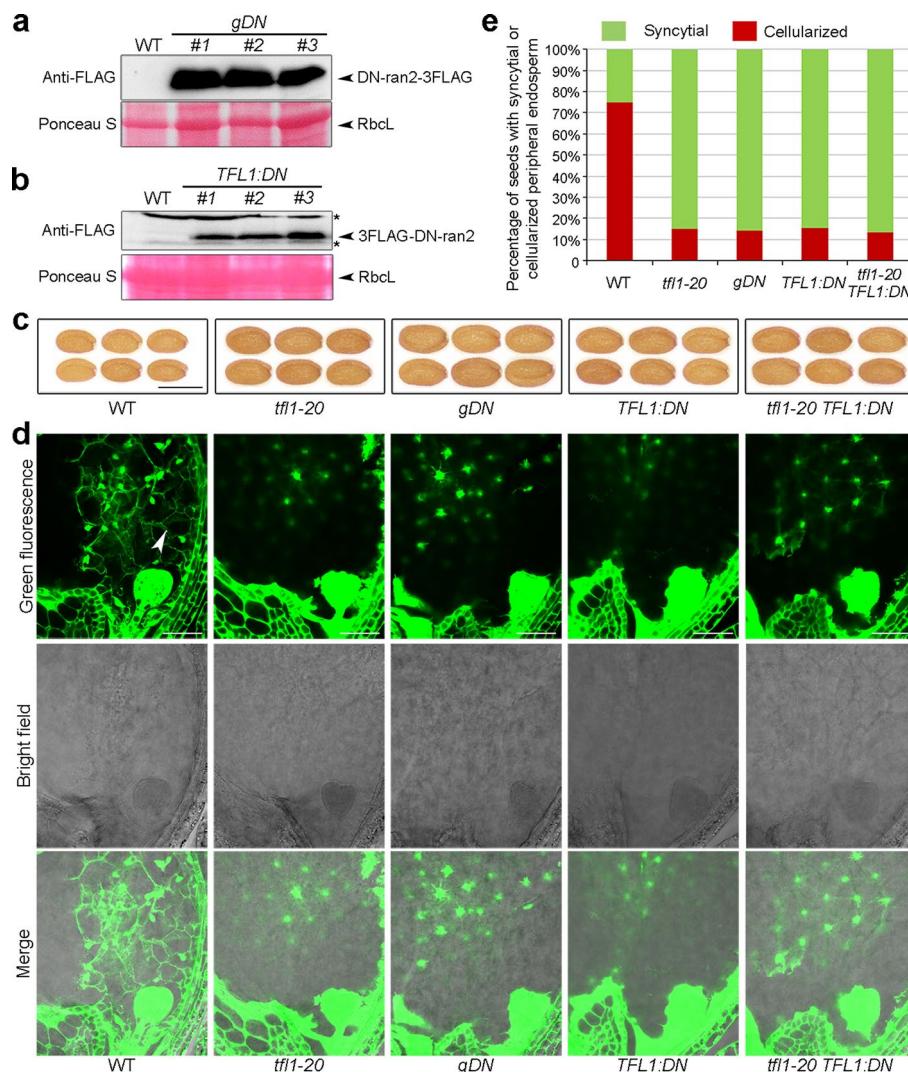


Extended Data Fig. 6 | Analysis of *RAN1*, *RAN2* and *RAN3* expression patterns. **a**, Quantitative analysis of *RAN1*, *RAN2* and *RAN3* expression in various tissues (upper panels) or developing siliques at 2 to 7 days after pollination (DAP, lower panels). OF, open flower; IS, inflorescence stem; RL, rosette leaf; Rt, root; Sil, silique; CL, cauline leaf; FB, flower bud. Results were normalized against the expression levels of the *U-BOX* gene as an internal control. Expression levels in lower panels are shown as relative values to the 2 DAP level set as 1. Values are mean \pm SD of three biological replicates. **b**, GUS staining of *GUS-gRAN2* seeds at 1 to 5 DAP. Scale bars, 50 μ m. The experiment was repeated three times independently with similar results.

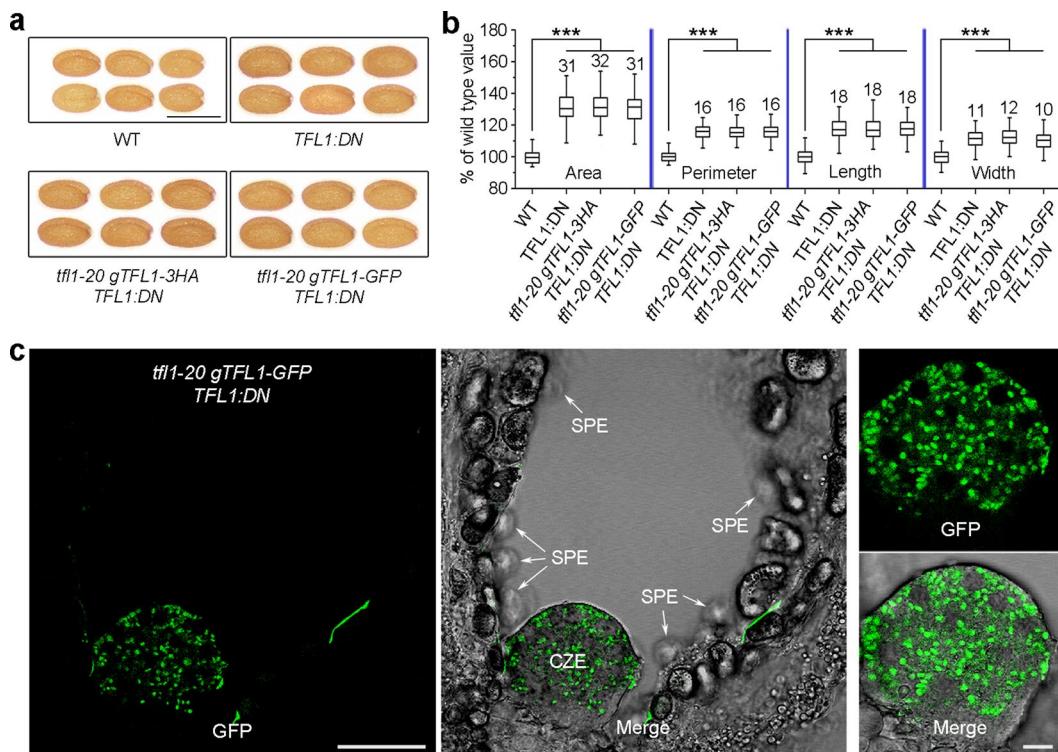


Extended Data Fig. 7 | Pull-down assay of the interaction between TFL1 and the dominant-negative (DN-ran2) or truncated versions of RAN2.

a, Schematic diagrams of DN-ran2 and RAN2 truncated (Del1-Del3) proteins that were fused to MBP. The full-length RAN2 protein contains the effector binding domain and acidic C-terminal domain implicated in protein-protein interaction. The DN-ran2 protein contains a mutation of threonine (T) at the residue 27 to asparagine (N), which is located near the effector-binding domain of RAN2. **b**, Pull-down assay result. MBP and various MBP fusion proteins were used as baits, and the corresponding loading control was stained with Ponceau S (lower panel). The input of the prey protein TFL1-4HA extracted from 35 S:TFL1-4HA siliques and its corresponding pull-downed signals were examined by immunoblot analysis using anti-HA antibody (upper panel). The experiment was repeated three times independently with similar results.

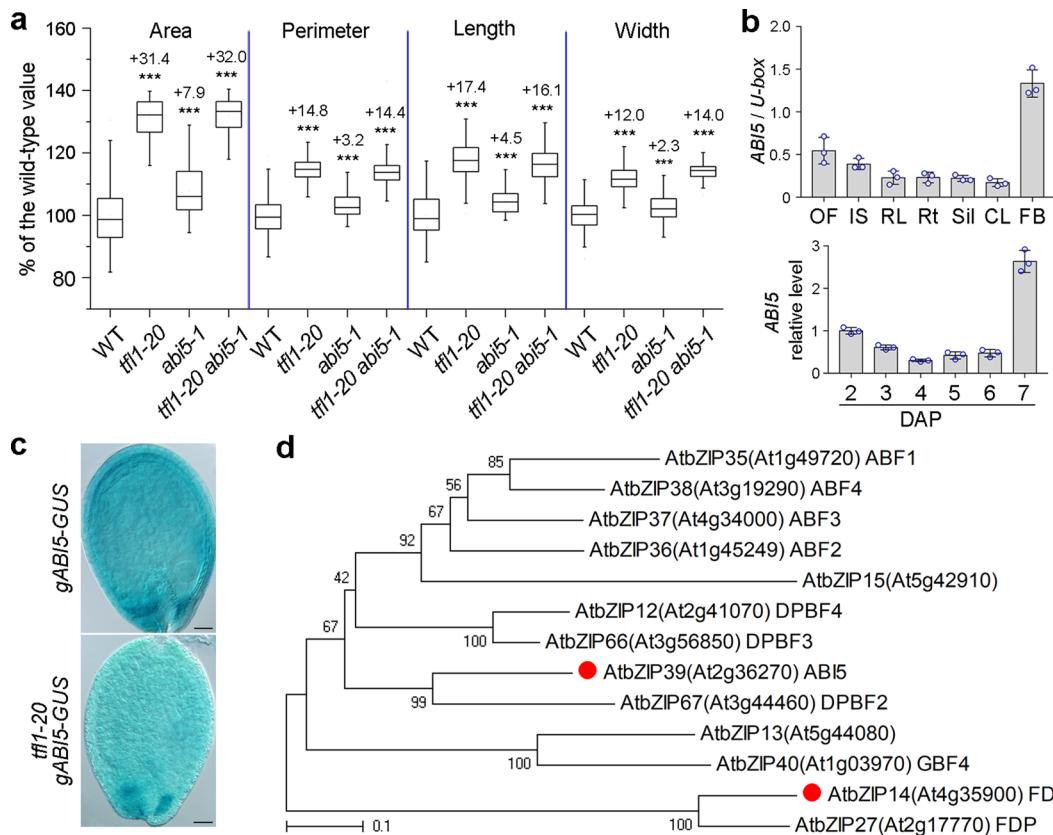


Extended Data Fig. 8 | Seed size phenotypes of *gDN*-*ran2*-3FLAG (*gDN*) and *TFL1*:3FLAG-DN-*ran2* (*TFL1*:DN) transgenic plants. **a,b**, Western blot analysis of DN-*ran2* protein expression in *gDN* (**a**) or *TFL1*:DN (**b**) transgenic lines. Total protein extracted from siliques was analysed using anti-FLAG antibody. Asterisks indicate non-specific bands. **c**, Comparison of mature dry seeds of wild-type (WT), *tfl1-20*, *gDN*, *TFL1*:DN and *tfl1-20 TFL1*:DN plants. Scale bar, 500 µm. **d**, Comparison of the peripheral endosperm development of cleared whole-mount seeds of WT, *tfl1-20*, *gDN*, *TFL1*:DN and *tfl1-20 TFL1*:DN at 4 DAP by confocal microscopy. The auto-fluorescence (green fluorescence) signal was generated by glutaraldehyde treatment. Arrowhead indicates new cell wall during endosperm cellularization. Merge, merge of green fluorescence and bright field images. Scale bars, 50 µm. The experiments in **a-d** were repeated three times independently with similar results. **e**, Percentage of seeds with syncytial or cellularized peripheral endosperm in WT, *tfl1-20*, *gDN*, *TFL1*:DN and *tfl1-20 TFL1*:DN at 4 DAP examined by confocal microscopy. Randomly selected seeds from more than 10 siliques for each genotype were examined.



Extended Data Fig. 9 | *tfi1-20 gTFL1-3HA TFL1:3FLAG-DN-ran2* and *tfi1-20 gTFL1-GFP TFL1:3FLAG-DN-ran2* produce large seeds like *tfi1-20*.

a, Comparison of mature dry seeds of WT, *TFL1:3FLAG-DN-ran2* (*TFL1:DN*), *tfi1-20 gTFL1-3HA TFL1:DN* and *tfi1-20 gTFL1-GFP TFL1:DN* plants. Scale bar, 500 μ m. **b**, Quantitative analysis of seed size parameters of different genotypes. Box plots show medians (lines), interquartile ranges (boxes), and whiskers (extending 1.5 times the interquartile ranges) of seed size parameters (area, perimeter, length, and width) of seeds of different genotypes (WT, n=129; *TFL1:DN*, n=110; *tfi1-20 gTFL1-3HA TFL1:DN*, n=118; *tfi1-20 gTFL1-GFP TFL1:DN*, n=105). Percentage change (%) in a seed parameter of a specific genotype is shown relative to the mean value of WT plants set as 100%. The increase in the percentage in a specific genotype over WT is indicated above each box. Asterisks indicate significant differences between WT plants and other genotypes (two-tailed Mann-Whitney test, $P < 0.0001$). P values are all less than 1×10^{-15} . **c**, Localization of *TFL1-GFP* in another *tfi1-20 gTFL1-GFP TFL1:DN* seed at 4 DAP by cryosectioning. *TFL1-GFP* signal is observable in the chalazal endosperm (CZE; left and middle panels) and its close-up view (right two panels). Note that *TFL1-GFP* is absent in all the syncytial peripheral endosperm (SPE). Merge, merge of GFP and bright field images. Scale bars, 40 μ m (left) and 10 μ m (right). The experiments in **a,c** were repeated three times independently with similar results.



Extended Data Fig. 10 | Examination of seed size phenotypes and gene expression pertaining to *ABI5*. **a**, Quantitative analysis of seed size parameters of *tf1-20*, *abi5-1*, and *tf1-20 abi5-1*. Box plots display medians (lines), interquartile ranges (boxes), and whiskers (extending 1.5 times the interquartile ranges) of seed size parameters (area, perimeter, length and width) of seeds of different genotypes (WT, $n=198$; *tf1-20*, $n=133$; *abi5*, $n=110$; *tf1-20 abi5*, $n=138$). Percentage change (%) in a seed parameter of a mutant is shown relative to the mean value of wild-type (WT) plants set as 100%. The increase in the percentage in a mutant over WT is indicated above each box. Asterisks indicate significant differences between WT plants and other mutants (two-tailed Mann-Whitney test, $P < 0.001$). P values for seed area, perimeter, length, and width between WT and *abi5* are 1.48×10^{-12} , 1.74×10^{-8} , 1.55×10^{-11} , and 1.78×10^{-4} , respectively, while P values between WT and *tf1-20* or *tf1-20 abi5* are all less than 1×10^{-15} . **b**, Quantitative analysis of *ABF5* expression in various tissues (upper panel) or developing siliques at 2 to 7 days after pollination (DAP, lower panels). OF, open flower; IS, inflorescence stem; RL, rosette leaf; Rt, root; Sil, silique; CL, cauline leaf; FB, flower bud. Results were normalized against the expression levels of the *U-BOX* gene as an internal control. Expression levels in the lower panel are shown as relative values to the 2 DAP level set as 1. Values are mean \pm SD of three biological replicates. **c**, Comparison of GUS staining signals between *gABI5-GUS* and *tf1-20 gABI5-GUS* seeds at 4 DAP. The *gABI5-GUS* construct was generated through translational fusion of a *GUS* gene with a 4.1-kb *ABI5* genomic fragment before the stop codon TAA. Scale bar, 50 μ m. The experiment was repeated three times independently with similar results. **d**, Phylogenetic analysis of the A group of basic leucine zipper (bZIP) transcription factors. The phylogenetic tree was generated by MEGA4 using the Neighbor-joining algorithm. Numbers on the branches indicate bootstrap values in 1,000 replicates. *ABI5* and *FD* are marked with red dots.

Corresponding author(s): Hao Yu

Last updated by author(s): Jul 10, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Seed size parameters (area, perimeter, length and width) were measured using the ImageJ 1.49g software (<https://imagej.nih.gov/ij/>). Quantification of Western blot signals was conducted by tracing out the signal (Western blot signal or loading control stained by Ponceau S solution) in individual bands using the ImageJ 1.49g software.

Data analysis

Two-tailed Student's t test and two-tailed Mann-Whitney test were performed using GraphPad Prism 7.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The authors declare that all data supporting the findings of this study are available within the article and its supplementary information files or from the corresponding author upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Experimental sample sizes were estimated based on our past experience performing similar experiments or published papers. More than 100 seeds of each genotype were analyzed for measuring different seed size parameters (area, perimeter, length, width, and length/width ratio) as previously reported (Garcia. et al., Arabidopsis haiku mutants reveal new controls of seed size by endosperm. <i>Plant Physiology</i> , 131, 1661-1670, 2003). Five sample batches with each containing 500 seeds were used for measuring seed mass as previously reported (Jofuku et al., Control of seed mass and seed yield by the floral homeotic gene APETALA2. <i>PNAS</i> , 102, 3117-3122, 2005).
Data exclusions	No data from experiments other than the preliminary analyses/experimental optimization were excluded. When measuring seed size by the ImageJ 1.49g software, we only examined clearly separate, non-aggregated seeds to avoid generation of incorrect seed size parameters from aggregated seeds.
Replication	Experiments were independently repeated by multiple times and similar results were obtained. The information on replication has been included in all relevant figure legends.
Randomization	Samples were grown on the same condition and randomly allocated in the growth chamber. Experimental plant materials were collected randomly without any bias.
Blinding	No blinding. The experimental materials are plants, thus the blinding design is not applicable to this system.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	<i>Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).</i>
Research sample	<i>State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.</i>
Sampling strategy	<i>Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.</i>
Data collection	<i>Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.</i>
Timing	<i>Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.</i>
Data exclusions	<i>If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Non-participation	<i>State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.</i>
Randomization	<i>If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.</i>

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	<i>Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.</i>
Research sample	<i>Describe the research sample (e.g. a group of tagged <i>Passer domesticus</i>, all <i>Stenocereus thurberi</i> within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.</i>
Sampling strategy	<i>Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.</i>
Data collection	<i>Describe the data collection procedure, including who recorded the data and how.</i>
Timing and spatial scale	<i>Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken</i>
Data exclusions	<i>If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>
Reproducibility	<i>Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.</i>
Randomization	<i>Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.</i>
Blinding	<i>Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.</i>

Did the study involve field work? Yes No

Field work, collection and transport

Field conditions	<i>Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).</i>
Location	<i>State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).</i>
Access and import/export	<i>Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).</i>
Disturbance	<i>Describe any disturbance caused by the study and how it was minimized.</i>

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

Antibodies

Antibodies used	Antibodies used in this study, including anti-HA (Santa Cruz, SC-7392), anti-FLAG (Sigma, F3165), anti-mouse IgG-HRP (Santa Cruz, SC-516102), anti-HA-HRP (Santa Cruz, SC-7392 HRP), anti-H3 (Abcam, AB1791), mouse anti-rabbit IgG-HRP (Santa Cruz, SC-2357), anti-HIS (NewEast Biosciences, 28004), anti-MBP-HRP (NEB, E8038S), and 15 nm gold-conjugated goat anti-mouse antibody (EMs, 25132) are commercially available.
Validation	The validation statements for the species and application of all the commercial antibodies used in this study can be found in the manufacturers' websites through catalogue numbers.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	<i>State the source of each cell line used.</i>
Authentication	<i>Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.</i>
Mycoplasma contamination	<i>Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.</i>
Commonly misidentified lines (See ICLAC register)	<i>Name any commonly misidentified cell lines used in the study and provide a rationale for their use.</i>

Palaeontology

Specimen provenance	<i>Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).</i>
Specimen deposition	<i>Indicate where the specimens have been deposited to permit free access by other researchers.</i>
Dating methods	<i>If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.</i>

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	<i>For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.</i>
Wild animals	<i>Provide details on animals observed in or captured in the field; report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.</i>
Field-collected samples	<i>For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.</i>
Ethics oversight	<i>Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.</i>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	<i>Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."</i>
Recruitment	<i>Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.</i>
Ethics oversight	<i>Identify the organization(s) that approved the study protocol.</i>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	<i>Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.</i>
Study protocol	<i>Note where the full trial protocol can be accessed OR if not available, explain why.</i>
Data collection	<i>Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.</i>
Outcomes	<i>Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.</i>

ChIP-seq

Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links <i>May remain private before publication.</i>	<i>For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.</i>
Files in database submission	<i>Provide a list of all files available in the database submission.</i>
Genome browser session (e.g. UCSC)	<i>Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.</i>

Methodology

Replicates	<i>Describe the experimental replicates, specifying number, type and replicate agreement.</i>
Sequencing depth	<i>Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.</i>
Antibodies	<i>Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.</i>
Peak calling parameters	<i>Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.</i>
Data quality	<i>Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.</i>
Software	<i>Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.</i>

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	<i>Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.</i>
Instrument	<i>Identify the instrument used for data collection, specifying make and model number.</i>
Software	<i>Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.</i>

Cell population abundance Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type Indicate task or resting state; event-related or block design.

Design specifications Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition

Imaging type(s) Specify: functional, structural, diffusion, perfusion.

Field strength Specify in Tesla

Sequence & imaging parameters Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.

Area of acquisition State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.

Diffusion MRI Used Not used

Preprocessing

Preprocessing software Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).

Normalization If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.

Normalization template Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.

Noise and artifact removal Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).

Volume censoring Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

Statistical modeling & inference

Model type and settings Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: Whole brain ROI-based Both

Statistic type for inference Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.
(See [Eklund et al. 2016](#))

Correction Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a Involved in the study

Functional and/or effective connectivity

Graph analysis

Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.