LETTER

https://doi.org/10.1038/s41586-019-1165-8

Arabidopsis FLL2 promotes liquid–liquid phase separation of polyadenylation complexes

Xiaofeng Fang[1,7], Liang Wang[2,7], Ryo Ishikawa[1,5,7], Yaoxi Li[1], Marc Fiedler[3], Fuquan Liu[1,6], Grant Calder[1], Beth Rowan[4], Detlef Weigel[4], Pilong Li[2]* & Caroline Dean[1]*

An important component of cellular biochemistry is the concentration of proteins and nucleic acids in non-membranous compartments[1,2]. These biomolecular condensates are formed from processes that include liquid–liquid phase separation. The multivalent interactions necessary for liquid–liquid phase separation have been extensively studied in vitro[1,3]. However, the regulation of this process in vivo is poorly understood. Here we identify an in vivo regulator of liquid–liquid phase separation through a genetic screen targeting factors required for *Arabidopsis* RNA-binding protein FCA function. FCA contains prion-like domains that phase-separate in vitro, and exhibits behaviour in vivo that is consistent with phase separation. The mutant screen identified a functional requirement for FLL2, a coiled-coil protein, in the formation of FCA nuclear bodies. FCA reduces transcriptional read-through by promoting proximal polyadenylation at many sites in the *Arabidopsis* genome[3,4]. FLL2 was required to promote this proximal polyadenylation, but not the binding of FCA to target RNA. Ectopic expression of FLL2 increased the size and number of FCA nuclear bodies. Crosslinking with formaldehyde captured in vivo interactions between FLL2, FCA and the polymerase and nuclease modules of the RNA 3′-end processing machinery. These 3′ RNA-processing components colocalized with FCA in the nuclear bodies in vivo, which indicates that FCA nuclear bodies compartmentalize 3′-end processing factors to enhance polyadenylation at specific sites. Our findings show that coiled-coil proteins can promote liquid–liquid phase separation, which expands our understanding of the principles that govern the in vivo dynamics of liquid-like bodies.

The regulation of the *Arabidopsis* floral repressor *FLC* involves conserved co-transcriptional mechanisms[5]. Autonomous pathway components lead to alternative 3′-end processing of the *FLC* antisense transcript (*COOLAIR*), chromatin modification and the generation of an *FLC* chromatin environment that reduces *FLC* transcriptional initiation and elongation[6]. The alternative 3′-end processing requires FCA[7,8], an RNA-binding protein that interacts with RNA 3′-end processing components[9,10]. To gain a mechanistic understanding of FCA function, we determined the subcellular localization of FCA. *fca-1* mutant plants that express a transgene (*FCA-eGFP*) for FCA fused to green fluorescence protein (eGFP) were generated (Extended Data Fig. 1a). The transgene fully complemented the *fca* mutation, and accelerated flowering and repressed *FLC* expression (Extended Data Fig. 1b, c). The mRNA and protein levels of the transgene were comparable to—although a little lower than—those of endogenous FCA (Extended Data Fig. 1d, e), which indicates that FCA–eGFP is not overexpressed and that it can functionally replace endogenous FCA. We acquired high-resolution images of FCA–eGFP in seven-day-old root-tip epidermal nuclei, and found that FCA–eGFP localized to multiple nuclear bodies (Fig. 1a).

FCA contains two N-terminal RNA-binding domains and a WW protein-interaction domain at its C terminus[11]. Outside of these domains, FCA is highly disordered—based on prediction[12]—and contains two prion-like domains (PrLDs)[13] (Fig. 1b). The PrLD has recently been identified as a driver for phase separation of RNA-binding proteins[14–17]. This prompted us to test whether FCA bodies have liquid-like characteristics. We first assessed the dynamicity of FCA bodies, using fluorescence recovery after photobleaching (FRAP). The spatio-temporal analysis of bleaching events showed that FCA redistributed rapidly from the unbleached area to the bleached area (Fig. 1c, d, Supplementary Video 1). Using time-lapse microscopy, we found that FCA bodies fused and relaxed into one body as soon as they intersected (Fig. 1e, Supplementary Video 2). We thus conclude that FCA localizes to nuclear bodies with liquid-like properties, suggesting that FCA can undergo liquid–liquid phase separation in vivo.

To determine whether FCA itself is a driver for the phase separation that we observed in vivo, we produced in vitro-recombinant FCA fused with a solubility tag (maltose-binding protein (MBP)), followed by a tobacco etch virus protease (TEV) cleavage site and a GFP tag (Fig. 1f). After TEV addition to cleave off MBP (Extended Data Fig. 2a), substantial green puncta formed; these were enhanced by addition of the crowding agent PEG8000 (PEG) (Fig. 1g) and low concentration of *Arabidopsis* total RNA (Extended Data Fig. 2b). To dissect which domain (or domains) of GFP–FCA is responsible for puncta formation, we separately expressed the N-terminal RNA recognition motif (RRM) domain or C-terminal PrLD of FCA (Fig. 1f). TEV cleavage of the C-terminal peptide yielded GFP–FCA–PrLD (Extended Data Fig. 2a) that formed extensive puncta, even without the addition of PEG (Fig. 1g). The GFP–FCA–RRM peptide did not form puncta, even with the addition of PEG (Fig. 1g, Extended Data Fig. 2a). These data indicate that the PrLD of FCA has the capacity to undergo phase separation.

We further investigated the detailed properties of in vitro phase separation of these proteins. GFP–PrLD was found to undergo phase separation at 1 μM, and the distinct droplets were spherical (Extended Data Fig. 2c). FRAP analysis indicated that GFP–PrLD molecules diffused rapidly within droplets, and exchanged between droplets and surrounding solution (Fig. 1h, i, Supplementary Video 3). Droplets were found to fuse upon contact (Fig. 1j, Supplementary Video 4). These data collectively indicate that GFP–PrLD undergoes liquid–liquid phase separation. However, the puncta formed by GFP–FCA in vitro did not show similar liquidity (Extended Data Fig. 2d–k), which suggests that there are additional regulators of FCA nuclear-body formation in cells.

To define the functional importance of and requirements for FCA nuclear-body formation, we undertook a mutagenesis screen for FCA function. We used an *Arabidopsis* progenitor line (C2) that contains three transgenes—*35S::FCAγ* (isoform that encodes the functional FCA protein) for overexpressing FCA, *FRIGIDA*, which encodes a strong activator of *FLC*, and *FLC* fused to the luciferase open reading

[1]John Innes Centre, Norwich, UK. [2]Beijing Advanced Innovation Center for Structural Biology, Tsinghua University-Peking University Joint Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing, China. [3]MRC Laboratory of Molecular Biology, Cambridge, UK. [4]Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany. [5]Present address: Graduate School of Agricultural Science, Kobe University, Kobe, Japan. [6]Present address: Institute of Global Food Security, School of Biological Sciences, Queen's University Belfast, Belfast, UK. [7]These authors contributed equally: Xiaofeng Fang, Liang Wang, Ryo Ishikawa. *e-mail: pilongli@mail.tsinghua.edu.cn; caroline.dean@jic.ac.uk

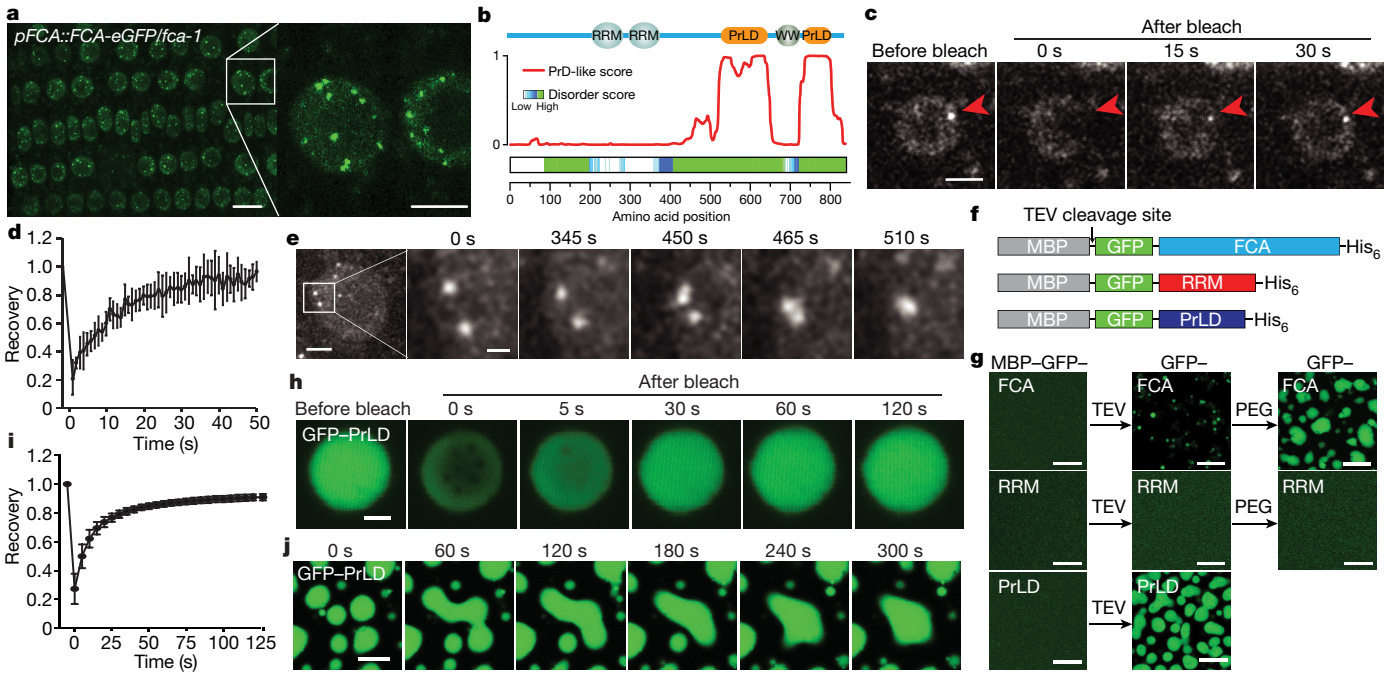


Fig. 1 | FCA phase separates in vitro, and exhibits behaviour in vivo that is consistent with phase separation. a, Fluorescence microscopy of *Arabidopsis* root-tip nuclei that express FCA–eGFP. Maximum projections of *z*-stacks that span the entire width of a nucleus were applied. Scale bars, 10 μm (left), 5 μm (right). Data are representative of five independent experiments. **b**, Top, protein domain structure of FCA. Bottom, predictions of PrLDs and disordered regions by 'Prion-like Amino Acid Composition' (PLAAC; http://plaac.wi.mit.edu/) and D^2P^2 (ref. [12]) algorithms, respectively. **c**, FRAP of FCA nuclear bodies. Time 0 indicates the time of the photobleaching pulse. Scale bar, 5 μm. Data are representative of seven independent experiments. **d**, Plot showing the time course of the recovery after photobleaching FCA nuclear bodies. Data are presented as mean ± s.d. ($n = 7$). **e**, Fluorescence time-lapse microscopy of *Arabidopsis* root-tip nuclei that express FCA–eGFP. Two fusing bodies are zoomed-in. Scale bars, 2 μm (left), 0.5 μm (right). Data are representative of three independent experiments. **f**, Schematic of protein fusions used for in vitro phase separation assay. **g**, In vitro phase separation assay of 10 μM GFP–FCA full-length and truncated proteins. Scale bars, 10 μm. Data are representative of three independent experiments. **h**, FRAP of GFP–FCA–PrLD droplets. Time 0 indicates the time of the photobleaching pulse. Scale bar, 2 μm. Data are representative of 12 independent experiments. **i**, Plot showing the time course of the recovery after photobleaching GFP–FCA–PrLD droplets. Data are presented as mean ± s.d. ($n = 12$). **j**, Fusion of GFP–FCA–PrLD droplets. Data are representative of three independent experiments.

frame (*FLC::LUC*) for monitoring the expression of *FLC*—to screen for the components that are required for FCA-mediated *FLC* repression[9,18,19]. This identified a new mutant that we term *sof78*, which has increased *FLC* expression and which flowered late compared to C2 wild type (Fig. 2a–c). The only other observed developmental defect was that flowers occasionally had five petals (Extended Data Fig. 3a). *sof78* carries a mutation in *FLL2* (also known as *AT1G67170*) that leads to a Glu-to-Lys amino acid substitution (E201K) (Fig. 2d). Transgenic expression of *FLL2* fully rescued the phenotype of *sof78* (Extended Data Fig. 3b–d). The effect of *sof78* mutation was not dependent on the *35S::FCAγ*, *FRIGIDA* and *FLC::LUC* transgenes, as shown by the fact that the mutant continued to flower late even after all three transgenes were removed (Extended Data Fig. 3e, f). We found that *sof78* mutation is semi-dominant (haploinsufficient) on the basis of three criteria. First, the heterozygous mutant *FLL2 sof78* has a flowering time that is intermediate between that of the wild type and the homozygous mutant (Extended Data Fig. 3g). Second, *fll2-2*—a loss-of-function allele that contains an *Agrobacterium* transfer (T-)DNA insertion (Extended Data Fig. 3h, i)—did not change the expression of *FLC* (Extended Data Fig. 3j) or affect flowering time (Extended Data Fig. 3k). Loss of the FLL2 protein potentially leads to use of the homologues FLL1 and FLL3 (Extended Data Fig. 3l), which associate with FLL2 in vivo (Supplementary Table 3). Third, the hybrid that carries *sof78* and *fll2-2* had higher levels of *FLC::LUC* expression than either wild-type or heterozygous *FLL2 fll2-2* plants (Extended Data Fig. 3m). We then performed an epistasis analysis and found that the effect of *sof78* mutation was not additive with *fca-1* with regard to flowering time (Fig. 2e). The *sof78* mutation did not affect the level of FCA protein (Fig. 2f). Together, these data suggest that FLL2 is required for FCA function.

FLL2 has not been assigned any functional domain. We performed a deeper search using HHpred inside the MPI Bioinformatics Toolkit[20] and found FLL2 contains a coiled-coil segment with high probability (Extended Data Fig. 4a, b). The E201K amino acid change is predicted to influence a salt bridge that connects two coiled-coils[21] (Extended Data Fig. 4b, c). This Glu201 is highly conserved in all green plants (Extended Data Fig. 4d). As with FCA, FLL2 is highly disordered and contains a PrLD (Fig. 2d). We then examined the subcellular localization of FLL2. Stable transgenic expression of FLL2 tagged with yellow fluorescent protein (*FLL2-eYFP*) fully complemented the late-flowering phenotype of *sof78* (Extended Data Fig. 3h, n). FLL2–eYFP was detected in nuclear bodies (Fig. 2g), colocalizing with FCA after co-transformation of *FLL2-eYFP* and *FCA-CFP* (a construct expressing FCA fused to cyan fluorescent protein) into *Arabidopsis* cultured cells (Fig. 3a). FLL1 also colocalized with FCA in nuclear bodies, whereas FLL3 was evenly distributed in the nucleus (Extended Data Fig. 5a). Consistent with this, FLL1—but not FLL3—contains a PrLD (Extended Data Fig. 5b, c). However, using yeast two-hybrid assays, we found that FLL1 and FLL3 interacted with FCA, whereas FLL2 did not (Extended Data Fig. 5d)—future work is therefore required to fully elaborate the roles of the FLL proteins. The lack of a yeast two-hybrid interaction with phase-separated partners has also been found for FUS[22]. We reasoned that, given the dynamic property of FCA bodies, the association between FLL2 and FCA might be transient and dynamic; indeed, after formaldehyde crosslinking their interaction could be detected (Fig. 3b). We further interrogated the transient interactions between FLL2 and FCA in vitro. To this end, we first tested the in vitro phase separation propensity of FLL2. GFP–FLL2 resulted in few puncta and the addition of 10% PEG greatly enhanced puncta formation (Fig. 3c,

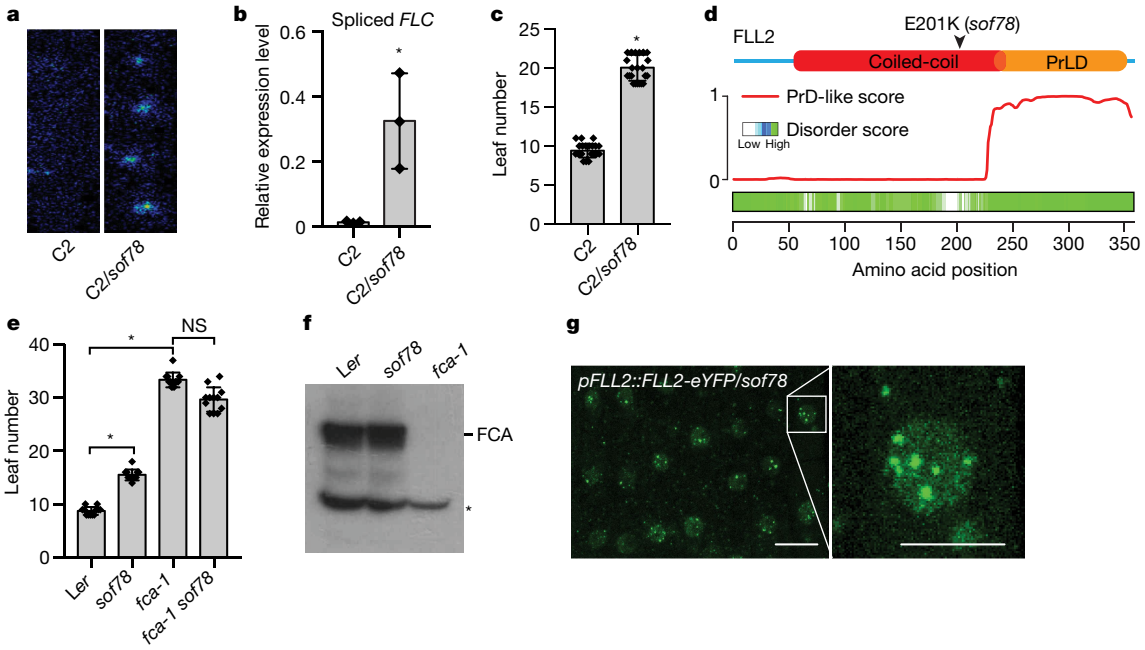


Fig. 2 | The coiled-coil domain protein FLL2 is required for the function of FCA. a, FLC–LUC bioluminescence signal of indicated plants taken by charge-coupled device camera. Data are representative of three independent experiments. **b**, Expression of spliced *FLC* relative to *UBC* in the indicated plants. Data are presented as mean ± s.d. ($n = 3$). Asterisk indicates a significant difference ($P = 0.0214$, two-tailed *t*-test). **c**, Flowering time of indicated plants (assayed as total leaf number produced by the apical meristem before it switched to producing flowers) grown in a long-day photoperiod. Data are presented as mean ± s.d. ($n = 20$). Asterisk indicates a significant difference ($P < 0.0001$, two-tailed *t*-test). **d**, Top, protein domain structure of FLL2. Bottom, predictions of PrLDs and disordered regions by PLAAC and D^2P^2 algorithms, respectively. **e**, Flowering time of indicated plants grown in a long-day photoperiod. Data are presented as mean ± s.d. ($n = 12$). Asterisks denote significant differences between the indicated plants ($P < 0.0001$, two-tailed *t*-test). NS, not significant. L*er*, Lansberg *erecta* wild type. **f**, The protein level of FCA in the indicated plants, as determined by western blot. Asterisk indicates non-specific signal. Data are representative of two independent experiments. For gel source data, see Supplementary Fig. 1. **g**, Fluorescence microscopy of *Arabidopsis* root-tip nuclei that express FLL2–eYFP. Maximum projections of *z*-stacks that span the entire width of the nucleus were applied. Scale bars, 10 μm (left), 5 μm (right). Data are representative of three independent experiments.

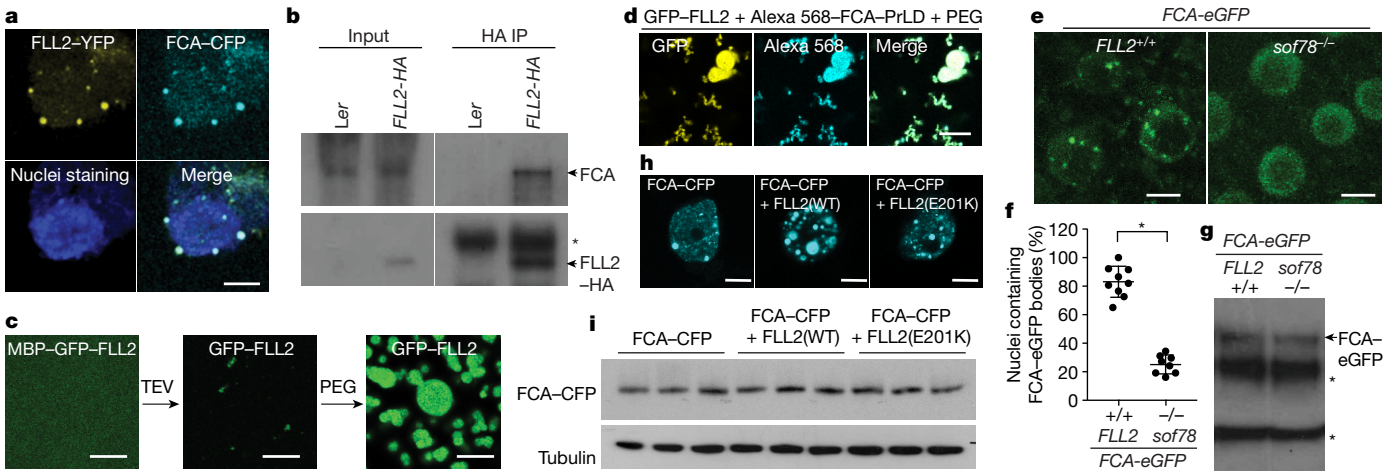


Fig. 3 | FLL2 promotes the phase separation of FCA to form nuclear bodies. a, Colocalization of FLL2–YFP with FCA–CFP in *Arabidopsis* cultured-cell nuclei. Scale bar, 5 μm. Images are representative of three independent experiments. **b**, Co-immunoprecipitation (IP) in stable transgenic plants after formaldehyde crosslinking to detect the association of haemagglutinin-tagged FLL2 (FLL2–HA) with FCA. Asterisk indicates non-specific signal. Data are representative of two independent experiments. For gel source data, see Supplementary Fig. 1. **c**, In vitro phase separation assay of 10 μM GFP–FLL2. Scale bars, 10 μm. Data are representative of three independent experiments. **d**, Co-separation of FCA–PrLD with GFP–FLL2. Alexa-Fluor-568-labelled 0.5 μM FCA–PrLD preferentially partitioned into GFP–FLL2 droplets. Scale bar, 10 μm. Data are representative of three independent experiments. **e**, Representative fluorescence microscopy images of FCA–GFP nuclear bodies in wild-type and *sof78*-mutant backgrounds. For each image, maximum projections of *z*-stacks that span the entire width of the nucleus were applied. Scale bars, 5 μm. Data are representative of eight independent experiments. **f**, Percentage of nuclei that contain FCA–GFP nuclear bodies in wild-type and *sof78*-mutant backgrounds. Data are presented as mean ± s.d. ($n = 8$). Asterisk indicates a significant difference ($P < 0.0001$, two-tailed *t*-test). **g**, The protein level of FCA–eGFP in wild-type and *sof78*-mutant backgrounds, as determined by western blot. Asterisks indicate non-specific signals. Data are representative of three independent experiments. For gel source data, see Supplementary Fig. 1. **h**, The effect of FLL2 overexpression on the pattern of FCA–CFP nuclear bodies assayed in tobacco leaf nuclei. Images are representative of three independent experiments. Scale bars, 5 μm. **i**, The protein level of FCA–CFP in indicated samples, as determined by western blot. Data are representative of three independent experiments. For gel source data, see Supplementary Fig. 1.

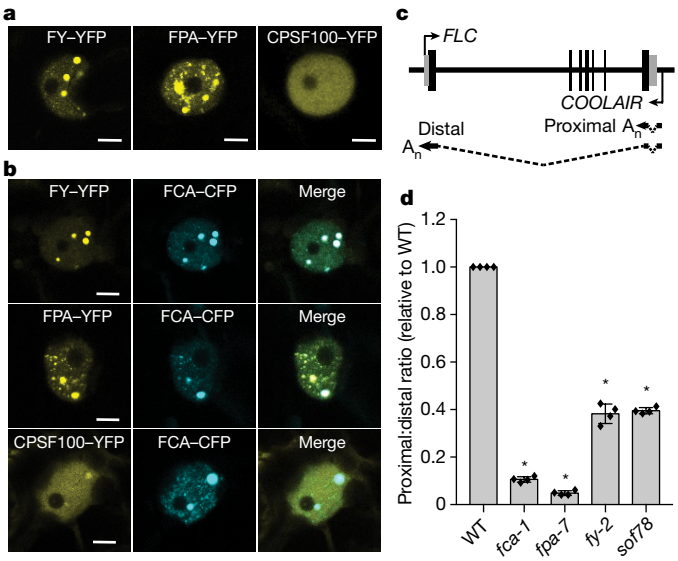


Fig. 4 | FCA and FLL2 associate with 3′-end processing factors and are important for polyadenylation at specific sites. a, Fluorescence microscopy of tobacco leaf nuclei that express the indicated proteins. Images are representative of three independent experiments. Scale bars, 5 μm. **b**, Colocalization of FY–YFP, FPA–YFP and CPSF100–YFP with FCA–CFP in tobacco leaf nuclei. Images are representative of three independent experiments. Scale bars, 5 μm. **c**, Schematic of *COOLAIR* transcripts from the *FLC* locus. A_n, polyA tail. Black rectangles denote exons and dashed lines denote introns. **d**, The ratio of proximal-to-distal isoforms of *COOLAIR* transcripts in the indicated plants relative to wild type (WT). Data are presented as mean ± s.d. ($n = 4$). Asterisks indicate significant differences ($P < 0.0001$, two-tailed *t*-test).

Extended Data Fig. 2a). FLL2 co-phase-separated with the PrLD of FCA (Fig. 3d).

The association between FLL2 and FCA led us to investigate whether FLL2 regulates the formation of the FCA–GFP bodies. To test this, we introduced the *sof78* mutation into *FCA-eGFP* transgenic plants, and compared FCA bodies between *sof78*-mutant and corresponding wild-type backgrounds. The *sof78* mutation reduced the percentage of nuclei that contained FCA bodies (Fig. 3e, f) and their size (Extended Data Fig. 6a), without affecting the protein level of FCA–eGFP (Fig. 3g). Ectopically expressing wild-type FLL2 increased the size and number of FCA–CFP bodies in tobacco leaves (Fig. 3h). These bodies were spherical and highly dynamic, and mixed very rapidly after half-bleaching (Extended Data Fig. 6b–d, Supplementary Video 5), which supports the idea that FCA–GFP bodies can behave as liquid droplets. This increase was lost through overexpression of either FLL2(E201K) or FLL2(K202E), which are mutated in amino acids that are predicted to form a salt bridge (Fig. 3h, Extended Data Fig. 6e). The protein levels of FCA–CFP were not changed by FLL2 overexpression (Fig. 3i, Extended Data Fig. 6f). These data indicate that the coiled-coil protein FLL2 promotes the phase separation of FCA to form nuclear bodies.

FCA promotes proximal polyadenylation of transcripts from many loci in the genome that were previously classified as unannotated segments[4], as well as *COOLAIR* transcripts[9]. This process requires the RNA 3′-end processing and polyadenylation factors FY[10,23], CSTF64 and CSTF77[9], and the RRM-containing protein FPA[9,24]. The previous model was that FCA—by interacting with FY—targets 3′-end processing machinery to the proximal polyadenylation site[9]. However, efforts to prove the in vivo interaction of FCA with 3′-end processing and polyadenylation factors had so far been unsuccessful[23]. We speculated that FCA bodies are the sites at which FCA dynamically and transiently interacts with FPA and 3′-end processing factors. To test this hypothesis, we developed a technique that we term crosslinked nuclear immunoprecipitation and mass spectrometry (CLNIP–MS; see Methods) to determine the interactors of FCA upon formaldehyde crosslinking: this gave large heterogeneous FCA complexes (Extended Data Fig. 7a). We found that FPA, FY, cleavage and polyadenylation specificity factors[25], FIP1 (which links cleavage and polyadenylation factors to poly(A) polymerases)[26] and FLL2 all co-purified with FCA (Extended Data Table 1, Supplementary Table 1). These constitute the polymerase and nuclease modules of the 3′-end processing machinery[27]. Without crosslinking, these proteins were not detected (Extended Data Table 1, Supplementary Table 2). CLNIP–MS of FLL2 identified FCA, FPA and some of the 3′-end processing factors (Extended Data Table 1, Supplementary Table 3). CLNIP–MS of FLD—which functions downstream of FCA to induce chromatin changes[18]—did not capture any of these factors (Supplementary Table 4); this indicates that the 3′-end processing factors specifically associate with FCA and FLL2.

We next asked whether these 3′-end processing factors localized to FCA bodies. We overexpressed FY and FPA fused with YFP in tobacco leaves, and found that both of them formed nuclear bodies (Fig. 4a) that fully overlapped with FCA–CFP bodies (Fig. 4b). Consistent with this, FY and FPA are similar to FCA in that they are also highly disordered and contain PrLDs[13] (Extended Data Fig. 7b, c). Overexpression of cleavage and polyadenylation specificity factor 30 (CPSF30)–YFP and CPSF100–YFP alone did not give rise to nuclear bodies (Fig. 4a, Extended Data Fig. 7d), but these proteins were recruited to FCA–CFP bodies when they were co-expressed with FCA (Fig. 4b, Extended Data Fig. 8d). These data support the notion that the bodies are indeed sites at which FCA associates with RNA 3′-end processing factors. FCA has previously been found to associate with *FLC* chromatin[18]. We found that immunoprecipitation of FCA enriched nascent transcripts of *COOLAIR* and unannotated segments (Extended Data Fig. 8a–e)—which supports a co-transcriptional mechanism—and that this was independent of FLL2 (Extended Data Fig. 8a–e). Similar to *fca, fpa* and *fy* mutants, polyadenylation of proximal sites in *COOLAIR* and proximal polyadenylation in unannotated segments was defective in the *sof78* mutant (Fig. 4c, d, Extended Data Fig. 8f–i). Thus, we propose that the FCA nuclear bodies compartmentalize 3′-end processing factors and that this enhances polyadenylation at specific 3′-end processing sites (Extended Data Fig. 9). We envisage that there will be a wide size range of these highly dynamic FCA nuclear bodies; an important next step will be to simultaneously image target loci, FLL2 and FCA bodies. This will help to establish the size of the functionally important unit, and to determine whether the large FCA–GFP bodies that we observed using low-resolution microscopy are functional or whether they represent one extreme of the formation of biomolecular condensates.

Coiled-coils are α-helical super-secondary structures that mediate protein–protein interactions and oligomerization[28]. The yeast prion protein Sup35 contains a coiled-coil region within the middle domain[29]; this region is involved in inducing liquid-like phase separation upon pH sensing[30], which implies that coiled-coil domains may have a general role in phase separation. Our findings thus help to define the multivalent in vivo interactions that drive liquid–liquid phase separation.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41586-019-1165-8.

Received: 9 November 2018; Accepted: 3 April 2019;
Published online 1 May 2019.

1. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
2. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
3. Li, P. et al. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).
4. Sonmez, C. et al. RNA 3′ processing functions of *Arabidopsis* FCA and FPA limit intergenic transcription. *Proc. Natl Acad. Sci. USA* **108**, 8508–8513 (2011).
5. Whittaker, C. & Dean, C. The *FLC* locus: a platform for discoveries in epigenetics and adaptation. *Annu. Rev. Cell Dev. Biol.* **33**, 555–575 (2017).

6. Wu, Z. et al. Quantitative regulation of *FLC* via coordinated transcriptional initiation and elongation. *Proc. Natl Acad. Sci. USA* **113**, 218–223 (2016).
7. Koornneef, M., Hanhart, C. J. & van der Veen, J. H. A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol. Gen. Genet.* **229**, 57–66 (1991).
8. Sheldon, C. C. et al. The *FLF* MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. *Plant Cell* **11**, 445–458 (1999).
9. Liu, F., Marquardt, S., Lister, C., Swiezewski, S. & Dean, C. Targeted 3′ processing of antisense transcripts triggers *Arabidopsis FLC* chromatin silencing. *Science* **327**, 94–97 (2010).
10. Simpson, G. G., Dijkwel, P. P., Quesada, V., Henderson, I. & Dean, C. FY is an RNA 3′ end-processing factor that interacts with FCA to control the *Arabidopsis* floral transition. *Cell* **113**, 777–787 (2003).
11. Macknight, R. et al. *FCA*, a gene controlling flowering time in *Arabidopsis*, encodes a protein containing RNA-binding domains. *Cell* **89**, 737–745 (1997).
12. Oates, M. E. et al. D^2P^2: database of disordered protein predictions. *Nucleic Acids Res.* **41**, D508–D516 (2013).
13. Chakrabortee, S. et al. Luminidependens (LD) is an *Arabidopsis* protein with prion behavior. *Proc. Natl Acad. Sci. USA* **113**, 6065–6070 (2016).
14. Molliex, A. et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).
15. Patel, A. et al. A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066–1077 (2015).
16. Riback, J. A. et al. Stress-triggered phase separation is an adaptive, evolutionarily tuned response. *Cell* **168**, 1028–1040 (2017).
17. Kato, M. et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
18. Liu, F. et al. The *Arabidopsis* RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate *FLC*. *Mol. Cell* **28**, 398–407 (2007).
19. Marquardt, S. et al. Functional consequences of splicing of the antisense transcript *COOLAIR* on *FLC* transcription. *Mol. Cell* **54**, 156–165 (2014).
20. Zimmermann, L. et al. A completely reimplemented MPI Bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243 (2018).
21. Kayikci, M. et al. Visualization and analysis of non-covalent contacts using the Protein Contacts Atlas. *Nat. Struct. Mol. Biol.* **25**, 185–194 (2018).
22. Yang, L., Embree, L. J., Tsai, S. & Hickstein, D. D. Oncoprotein TLS interacts with serine–arginine proteins involved in RNA splicing. *J. Biol. Chem.* **273**, 27761–27764 (1998).
23. Manzano, D. et al. Altered interactions within FY/AtCPSF complexes required for *Arabidopsis* FCA-mediated chromatin silencing. *Proc. Natl Acad. Sci. USA* **106**, 8772–8777 (2009).
24. Hornyik, C., Terzi, L. C. & Simpson, G. G. The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. *Dev. Cell* **18**, 203–213 (2010).
25. Hunt, A. G. et al. *Arabidopsis* mRNA polyadenylation machinery: comprehensive analysis of protein-protein interactions and gene expression profiling. *BMC Genomics* **9**, 220 (2008).
26. Forbes, K. P., Addepalli, B. & Hunt, A. G. An *Arabidopsis* Fip1 homolog interacts with RNA and provides conceptual links with a number of other polyadenylation factor subunits. *J. Biol. Chem.* **281**, 176–186 (2006).
27. Casañal, A. et al. Architecture of eukaryotic mRNA 3′-end processing machinery. *Science* **358**, 1056–1059 (2017).
28. Lupas, A. N. & Bassler, J. Coiled coils – a model system for the 21st century. *Trends Biochem. Sci.* **42**, 130–140 (2017).
29. Fiumara, F., Fioriti, L., Kandel, E. R. & Hendrickson, W. A. Essential role of coiled coils for aggregation and activity of Q/N-rich prions and polyQ proteins. *Cell* **143**, 1121–1135 (2010).
30. Franzmann, T. M. et al. Phase separation of a yeast prion protein promotes cellular fitness. *Science* **359**, eaao5654 (2018).

Acknowledgements We thank M. M. Babu (MRC, Laboratory of Molecular Biology) and M. Howard (John Innes Centre) for discussions. This work was supported by the European Research Council grant 'MEXTIM', Wellcome Trust grant 210654/Z/18/Z, the BBSRC Institute Strategic Programme GEN (BB/P013511/1), the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant 800318, the JSPS overseas research fellowship to R.I., and the Medical Research Council grant U105192713.

Reviewer information *Nature* thanks Julia Chekanova, Jessica Hennacy, Martin Jonikas and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions X.F. and C.D. conceived the study. X.F., Y.L. and G.C. performed all in vivo imaging experiments and analyses. X.F. performed the mass spectrometry experiments. R.I., B.R., D.W., M. F. and F.L. did the genetic identification and analysis of the *sof78* mutant. L.W. and P.L. performed and analysed in vitro phase-separation assays. X.F., P.L. and C.D. wrote the manuscript, and all authors contributed ideas and reviewed the manuscript.

Competing interests The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41586-019-1165-8.
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1165-8.
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to P.L. or C.D.
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and investigators were not blinded to allocation during experiments and outcome assessment.

Plant materials and growth conditions. The parental C2 line[18] and flowering-time mutants *fca-1*[11], *fpa-7*[31] and *fy-2*[32] have previously been described. *fca-9* was provided by C.-H. Yang (National Chung Hsing University, Taiwan). *fll2-2* (GK-084H05) was obtained from the European *Arabidopsis* Stock Centre.

The seeds were surface-sterilized and sown on standard half-strength Murashige and Skoog (MS) medium (0.22% MS, 1% sucrose, 0.5% Phytagel, Sigma, P8169) plates and kept at 4 °C in the dark for 2 days before being transferred to long-photoperiod conditions (16 h light of 120 μmol m^{-2} s^{-1}/8 h dark). All RNA and protein experiments were done using 12-day-old seedlings unless specified.

DNA constructs and generation of transgenic plants. To generate the *pFCA::FCA-eGFP* transgenic line, *FCA* genomic DNA including its promoter and 3′ untranslated region (UTR) was amplified and inserted into the pCambia1300 vector. The *FCA* ApaI fragment (*FCA* genomic DNA contains two ApaI sites) was swapped by *FCA* ApaI fragment fused with *GFP*, which was inserted before the stop codon using the HindIII restriction site. The resulting construct was transformed into the *fca-1* mutant.

For complementation of the *sof78* mutant, a 3.9-kb genomic fragment containing *AT1G67170* was cloned into the pCambia1300 vector and transformed into the *sof78* mutant. To generate *pFLL2::FLL2-eYFP* and *pFLL2::FLL2-HA* transgenic lines, *pFLL2::FLL2* with linker or haemagglutinin tag and 3′ UTR of *FLL2* were separately amplified from genomic DNA. The polymerase chain reaction (PCR) products were digested and sequentially cloned into a pENTR vector. *eYFP* sequence was then inserted between the above two fragments. The resulting *pENTR-FLL2* vectors were subsequently recombined into the binary destination vector pSLJ755I6[33]. These two constructs were transformed into the *sof78* mutant.

All the constructs described above were electroporated into *Agrobacterium tumefaciens* GV3101 for transformation of *Arabidopsis* by the floral-dip method.

The constructs for colocalization experiments were cloned by inserting coding sequences of *FCA*, *FLL1*, wild-type *FLL2*, *FLL3*, *FY*, *CPSF30* and *CPSF100* into the pCambia1300-35S-N1–YFP or pCambia1300-35S-N1–CFP vectors. *35S::FLL2*^{E201K}*-YFP* and *35S::FLL2*^{K202E}*-YFP* were cloned from *35S::FLL2-YFP* by PCR-directed mutagenesis. *35S::FPA-YFP* has previously been described[31].

The constructs for yeast two-hybrid assay were cloned by inserting coding sequences of *FCA*, *FLL1*, *FLL2* and *FLL3* into the pGADT7 and pGBKT7 vectors.

The constructs for in vitro protein expression were cloned by inserting coding sequences of *FCA*, *FLL2*, *FCA-PrLD* and *FCA-RRMs* into a modified pET11 expression vector (Novagen): a solubility tag, MBP, was followed by a TEV cleavage site and a GFP at the N terminus, and a non-cleavable 6×His tag at the C terminus. *FCA-PrLD* was also cloned into a similarly modified expression vector without the GFP tag.

Primers used for vector construction are listed in Supplementary Table 5.

Mutagenesis screening and cloning of *FLL2*. Chemical mutagenesis of the parental line C2 and screening of *sof78* mutants were carried out as previously described[18]. To map C2/*sof78*, the mutant was crossed to a Columbia line that contained the same transgenes, and the resulting F_2 plants were screened for FLC–LUC bioluminescence. Three hundred and ten plants with high FLC–LUC bioluminescence activity were pooled and the DNA was extracted and prepared for Illumina sequencing as previously described[34] using 2 × 150-bp paired-end reads to a depth of 20× coverage. Applying a previously described method[35], we identified a single genomic region enriched for L*er* alleles that did not correspond to the three transgenes, and identified candidate *sof78* causal mutations as polymorphisms that were not derived from L*er*. The identified mutations were confirmed by Sanger sequencing.

FLC–LUC detection. Seedlings around 12 days after germination on MS medium plates were sprayed with 1 mM of luciferin (Promega) substrate solution and incubated in the dark at room temperature for 20 min. LUC bioluminescence activity of the seedlings was assayed as described[9].

Flowering time analysis. Plants were grown in growth rooms with a photoperiod of 16 h light and 8 h dark. Temperature ranged between 23–25 °C during the day and 20–22 °C at night. The total leaf number produced by the main apical meristem before switching the developmental program to the initiation of flowering was counted to measure variation in flowering time.

Protein expression and purification. All proteins were expressed in *Escherichia coli* BL21 (DE3) cells (Tiangen) in the presence of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were induced overnight at 18 °C, collected and resuspended in lysis buffer (20 mM HEPES pH 7.4, 500 mM KCl, 1 mM PMSF). The cells were then lysed using a high-pressure homogenizer (ATS Engineering) and centrifuged. The supernatants were first purified with Ni-NTA and amylose resin (GenScript), followed by purification on a Superdex 200 increase 10/300 column (SD200) (GE healthcare). Proteins were stored in storage buffer (20 mM HEPES pH 7.4, 150 mM KCl, 1 mM DTT) at −80 °C after being flash-frozen in liquid nitrogen.

Protein labelling. MBP–FCA–PrLD–6×His protein solution was exchanged with reaction buffer (0.1 M sodium bicarbonate buffer pH 8.3) using a SD200 column. Alexa Fluor 568 carboxylic acid (Succinimidyl Ester) (Thermo Fisher, A20003) was added to MBP–FCA–PrLD–6×His proteins at a 1:1 molar ratio and incubated for 1 h at room temperature with continuous stirring. Free dye was removed using an SD200 column. The labelled proteins were stored in storage buffer at −80 °C.

In vitro phase separation assay. In vitro phase separation assay was performed in storage buffer. N-terminal MBP tags of MBP–GFP–FCA, MBP–GFP–FCA–PrLD, MBP–GFP–FCA–RRMs, MBP–GFP–FLL2 and MBP–FCA–PrLD were cleaved during droplet assembly with TEV protease overnight. Further droplet assembly for MBP–GFP–FCA, MBP–GFP–FCA–RRMs and MBP–GFP–FLL2 was mixed with 10% (w/v) final concentration of PEG8000 (Sigma). Phase separation of GFP–FCA in the presence of RNA was tested by adding *Arabidopsis* total RNA to 3.13 μM GFP–FCA to the final concentration of 1.3–0.09 μg/ml.

Droplets were assembled in 384 low-binding multi-well 0.17-mm microscopy plates (384-well microscopy plates) (In Vitro Scientific), sealed with optically clear adhesive film to prevent evaporation and observed under a Nikon A1 microscope equipped with 60× and 100× oil immersion objectives. Co-phase separation between FCA–PrLD and FLL2 was done by mixing Alexa-Fluor-568-labelled FCA–PrLD with MBP–GFP–FLL2 to final concentrations of 0.5 μM.

Microscopy. Plants were grown vertically on MS plates with 1% (w/v) sucrose and 0.5% (w/v) phytagel (Sigma-Aldrich, P8169). Analyses of subcellular localization were performed on a Zeiss LSM780 confocal microscope using a 40×/1.2 water objective and the GaAsP spectral detector of the LSM780. GFP was excited at 488 nm and detected at 491–535 nm, and YFP was excited at 514 nm and detected at 517–557 nm. All images are *z*-stack maximum projections using a step size of 0.45 μm, spanning the entire width of the nucleus.

Analyses of colocalization were also performed on a Zeiss LSM780 confocal microscope. YFP was excited at 514 nm and detected at 517–557 nm, CFP was excited at 458 nm and detected at 464–517 nm. YFP and CFP were acquired sequentially to avoid emission crosstalk. For microscopy of *Arabidopsis* cultured cells, cells were incubated with Vybrant DyeCycle Ruby Stain (Thermo Scientific, V10309) to stain the nuclei. Ruby was excited at 633 nm and detected at 638–686 nm.

For time-lapse microscopy of FCA nuclear bodies, a chamber was created on slides using SecureSeal adhesive sheets (Grace Bio-Labs, 620001) and filled with MS medium. Seven-day-old seedlings were transferred into the chamber and observed under an Andor Revolution XD Spinning disc confocal microscope using a 60× water immersion objective. Images were acquired every 15 s for 15 min. At each time point, maximum projections from *z*-stack of 14 steps with step size of 0.6 μm were applied. Image analysis was performed with FIJI/ImageJ.

FRAP. For the in vivo experiments, FRAP of FCA–eGFP bodies in *Arabidopsis* was performed as described[36] on an Andor Revolution XD spinning disc confocal microscope system with a Nikon ECLIPSE Ti microscope stand, Yokogawa CSU-X spinning disc and iXon 3 EMCCD camera. Using a 60× water immersion objective, a region of an FCA nuclear body was bleached using a laser intensity of 50% at 488 nm. Recovery was recorded for every second for a total of 50 s after bleaching. FRAP of FCA–YFP bodies formed in tobacco was performed on a Zeiss LSM880 Airy scan confocal microscope, using a 40×/1.1 water immersion objective lens. Bleaching was done using a 514-nm laser pulse (5 iterations, 50% intensity). To improve spatial resolution, the Airy scan detector was used in scan repetition mode and images were processed using Zen Black airy scan processing 2D method. Recovery was recorded every 233 ms. Analysis of the recovery curves was carried out with FIJI/ImageJ.

For the in vitro experiments, FRAP was carried out with samples in 384-well microscopy plates using a Nikon A1 microscope equipped with 60× and 100× oil immersion objectives, as above. Droplets were bleached with a 488-nm laser pulse (3 repeats, 70% intensity, dwell time 1 s). The recovery from photobleaching was recorded for the indicated time.

Yeast two-hybrid assay. Pairs of plasmids were co-transformed into the yeast strain AH109 following the manufacturer's handbook (Clontech). The co-transformed yeast clones were first grown on SD medium without Leu and Trp, and subsequently plated on SD medium without Ade, His, Leu and Trp.

RNA immunoprecipitation. About 2.0 g, 12-day-old wet seedlings were crosslinked in 1% formaldehyde. Crosslinked plants were ground into fine powder and lysed in 8 ml lysis buffer (20 mM Tris HCl pH 7.5, 20 mM KCl, 2 mM EDTA pH 8.0, 2.5 mM $MgCl_2$, 25% glycerol, 250 mM sucrose). The lysate was filtered through two layers of Miracloth (Merck, D00172956) and pelleted by centrifugation. The pellets were washed three times with 20 ml NRBT buffer (20 mM Tris HCl pH7.5, 2.5 mM $MgCl_2$, 25% glycerol, 0.2% Triton X-100), resuspended in 600 μl RIPA buffer (1× PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 20 U/ml RNaseOUT) and sonicated. Nuclear extract was incubated with FCA antibody and protein A magnetic beads (Pierce) for 2 h and washed sequentially with

low-salt, high-salt and TE buffers. The immunoprecipitates were then resuspended in 200 μl RNA Elution buffer (100 mM NaCl, 50 mM Tris HCl pH 7.0, 1 mM EDTA, 1% SDS) and boiled at 95 °C for 15 min. Ten microlitres proteinase K, 1 μl RNaseOUT and 2 μl 0.1 M DTT were added to each reaction and incubated at 65 °C for 60 min. The reaction was then stopped with 1 ml Trizol reagent (Ambion). RNA was then isolated and analysed.

RNA analysis. Total RNA or FCA-bound RNA was extracted, treated with TURBO DNase (Ambion) to remove DNA contamination and reverse-transcribed by SuperScript IV Reverse Transcriptase (Invitrogen) using gene-specific reverse primers. Quantitative PCR (qPCR) analysis was performed on a LightCycler480 II (ROCHE) and qPCR data were normalized to *UBC*. Semi-qPCR was performed with GoTaq G2 DNA polymerase (Promega) and PCR products were analysed by agarose gel electrophoresis.

For measuring the proximal:distal ratio of *COOLAIR*, the levels of proximal and distal *COOLAIR* were first normalized to total *COOLAIR*. The value of proximal *COOLAIR* was then divided by that of distal *COOLAIR*. All primers are described in Supplementary Table 5.

Western blot analysis. Protein samples were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis and transferred to PVDF membranes. Antibodies against FCA[37], haemagglutinin (Sigma, H3663) and GFP (Roche, 11814460001) were used as primary antibodies. After the primary antibody incubation, horseradish peroxidase-conjugated secondary antibodies (GE Healthcare) were used for protein detection by chemiluminescence (Thermo Scientific, 34095).

CLNIP–MS. Twelve-day-old seedlings were crosslinked in 1% formaldehyde. Vacuum was applied gently and released slowly to avoid disruption of subcellular structures. Crosslinked plants were ground into fine powder and lysed in lysis buffer (20 mM Tris HCl pH 7.5, 20 mM KCl, 2 mM EDTA pH 8.0, 2.5 mM $MgCl_2$, 25% glycerol, 250 mM sucrose). The lysate was filtered through two layers of Miracloth (Merck, D00172956) and pelleted by centrifugation. The pellets were washed three times with NRBT buffer (20 mM TrisHCl pH 7.5, 2.5 mM $MgCl_2$, 25% glycerol, 0.2% Triton X-100), resuspended in RIPA buffer (1× PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) and sonicated. Nuclear extract was incubated with FCA antibody, anti-HA magnetic beads (Thermo Scientific, 88836) or anti-Flag M2 magnetic beads (Sigma, M8823) and washed sequentially with low-salt, high-salt and TE buffers. The immunoprecipitates were boiled at 95 °C for 15 min to reverse crosslinking. The protein samples were gel-purified and subjected to mass spectrometry analysis by nano-liquid chromatography with tandem mass spectrometry on an Orbitrap Fusion Tribrid mass spectrometer coupled to an UltiMate 3000 RSLCnano LC system (Thermo Scientific). Data were searched using Mascot server (Matrix Science), and analysed using MaxQuant software[38].

Transient transformation of *Arabidopsis thaliana* cell-suspension cultures. The *Arabidopsis* Col-0 cells were subcultured in MS medium with B5 vitamins (1× MS basal salt mix, 1× Gamborg's B5 vitamins, 3% sucrose, 0.59 g/l MES buffer pH 5.7, 1 mg/l 2,4-dichlorophenoxyacetic acid) weekly and maintained at 100–110 r.p.m. at 24 °C in the dark. *Agrobacterium* was inoculated into 10 ml LB medium and incubated overnight. Cells were collected, resuspended in the same volume of sterile M buffer (10 mM $MgCl_2$, 10 mM MES pH 5.6, 50 μg/ml acetosyringone) and incubated at room temperature (20–22 °C) for 2–4 h. Cells were collected, resuspended in 4 ml sterile M buffer. Two hundred microlitres of this culture was added to 50 ml freshly subcultured *Arabidopsis* cells and incubated for 48 h before microscopy analysis.

Agro-infiltration of tobacco leaves. Agro-infiltration experiments were performed essentially as described[39], except that *Nicotiana benthamiana* plants were used.

Genotyping. To genotype C2 plants, *FRI* was genotyped by mixing two pairs of primers (FRI-JU223_F6+R6 and FRI-del_F1+R1) together. The *35S::FCA* transgene was genotyped with primers 35S::FCA_F6+R6. The *FLC-LUC* transgene was genotyped with primers LUC3_F+R.

To genotype the *sof78* and *fca-1* mutations, genomic DNA was amplified with primers sof78-dCAPs_F+R and fca-1-dCAPS_F+R, respectively. PCR products were digested with MseI, followed by 3% agarose gel electrophoresis.

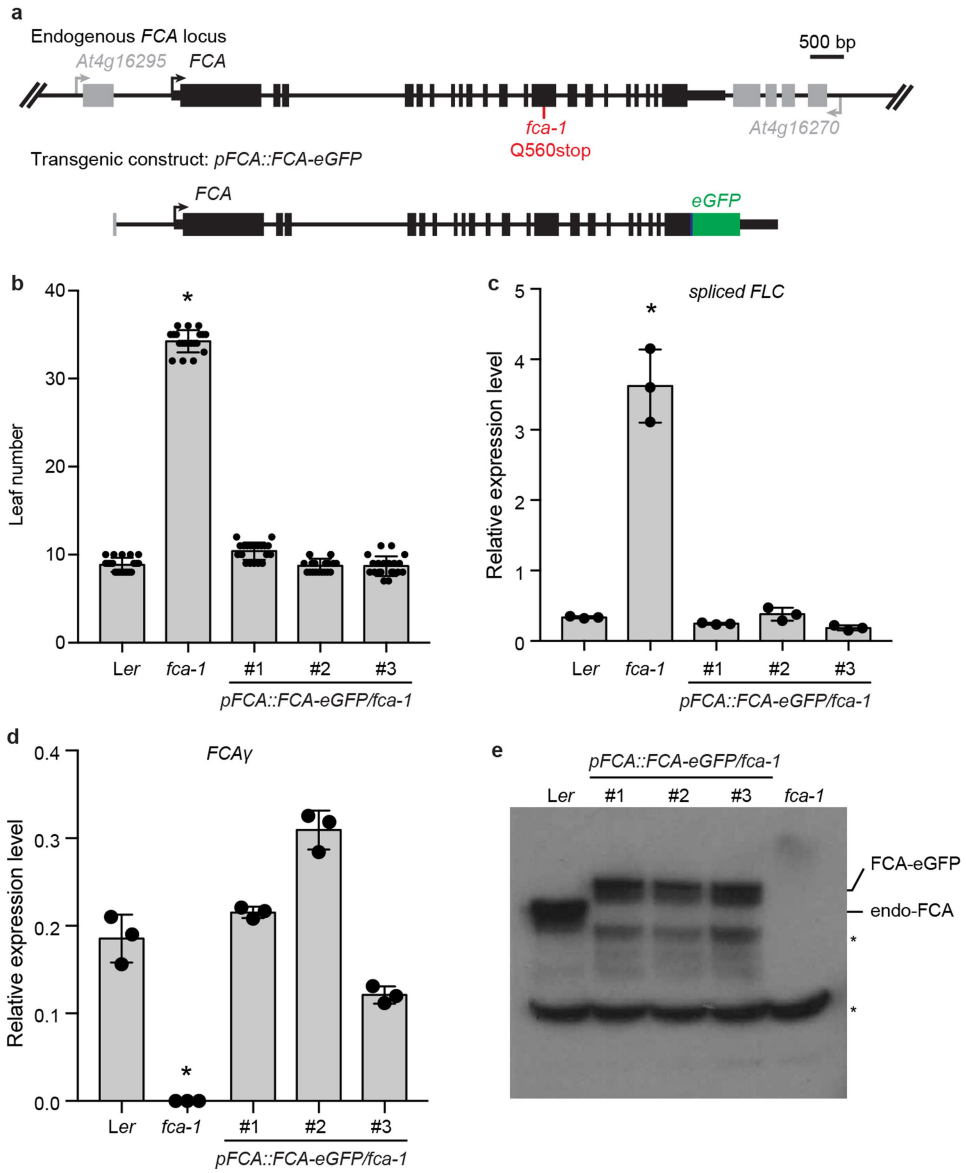
fll2-2 was genotyped with the transfer DNA left-border-specific primer GABI_F and At1g67170_R3 to detect the presence of the insertion. PCR using At1g67170_F3+R3 was performed to test whether the transfer DNA insertion was homozygous.

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

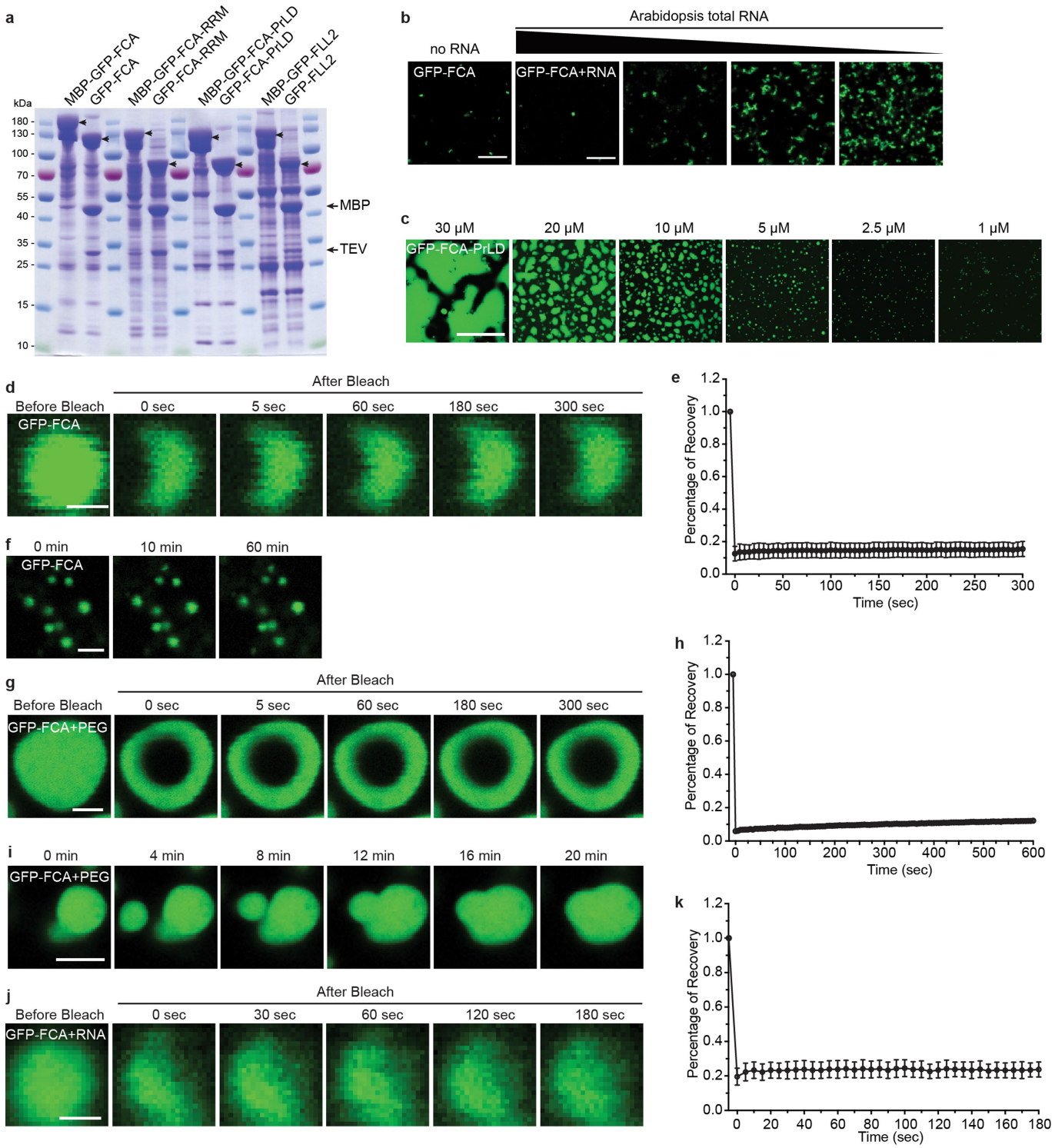
Data availability

Full lists of mass spectrometry are provided as Supplementary Tables 1–4. All the other raw data that support the findings of this study are available from the corresponding authors upon reasonable request.

31. Bäurle, I., Smith, L., Baulcombe, D. C. & Dean, C. Widespread role for the flowering-time regulators FCA and FPA in RNA-mediated chromatin silencing. *Science* **318**, 109–112 (2007).
32. Henderson, I. R., Liu, F., Drea, S., Simpson, G. G. & Dean, C. An allelic series reveals essential roles for FY in plant development in addition to flowering-time control. *Development* **132**, 3597–3607 (2005).
33. Jones, J. D. et al. Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. *Transgenic Res.* **1**, 285–297 (1992).
34. Rowan, B. A., Patel, V., Weigel, D. & Schneeberger, K. Rapid and inexpensive whole-genome genotyping-by-sequencing for crossover localization and fine-scale genetic mapping. *G3 (Bethesda)* **5**, 385–398 (2015).
35. Schneeberger, K. et al. SHOREmap: simultaneous mapping and mutation identification by deep sequencing. *Nat. Methods* **6**, 550–551 (2009).
36. Rosa, S. Measuring dynamics of histone proteins by photobleaching in *Arabidopsis* roots. *Methods Mol. Biol.* **1675**, 455–465 (2018).
37. Macknight, R. et al. Functional significance of the alternative transcript processing of the *Arabidopsis* floral promoter FCA. *Plant Cell* **14**, 877–888 (2002).
38. Tyanova, S. et al. Visualization of LC-MS/MS proteomics data in MaxQuant. *Proteomics* **15**, 1453–1456 (2015).
39. Fang, Y. & Spector, D. L. Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living *Arabidopsis* plants. *Curr. Biol.* **17**, 818–823 (2007).

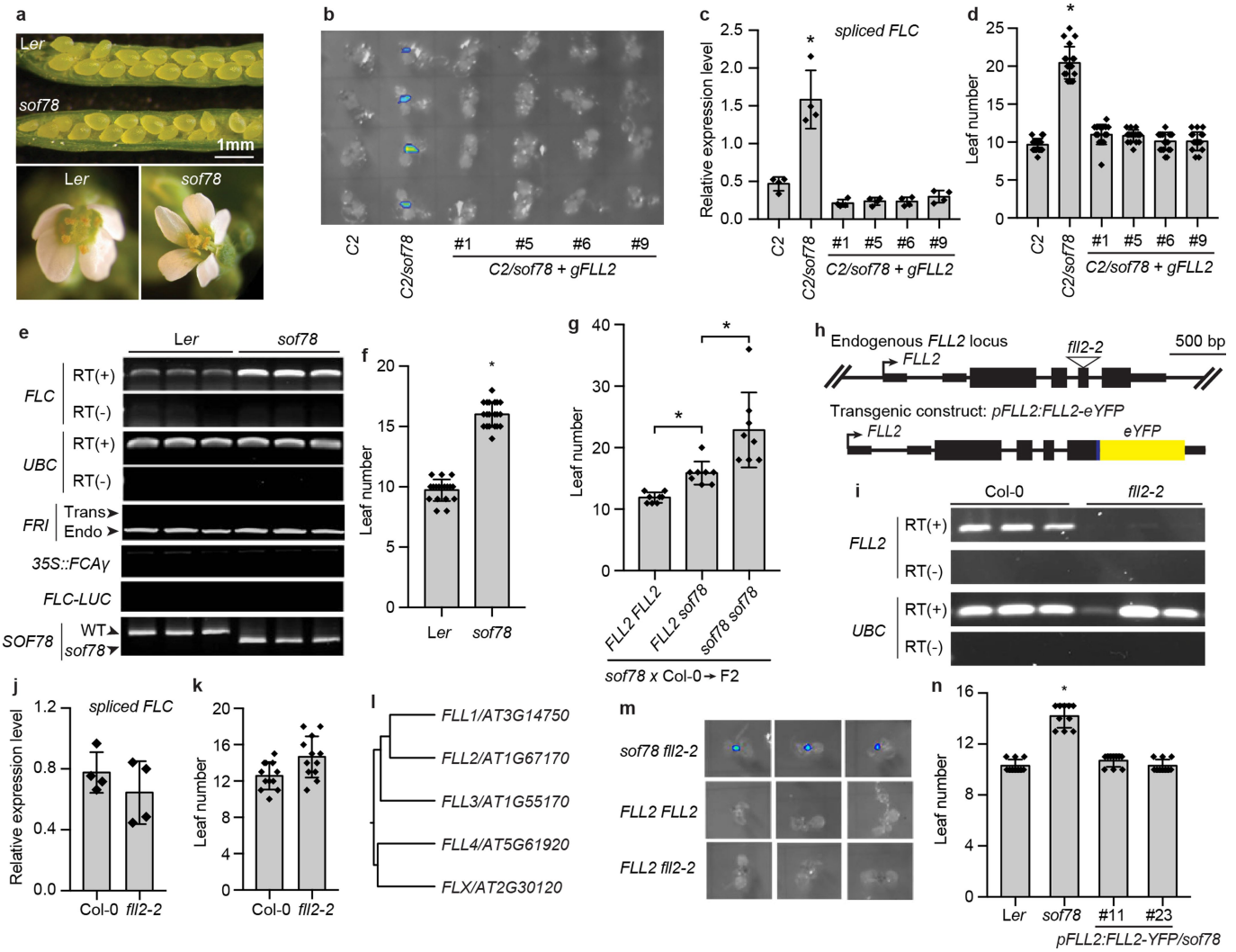


Extended Data Fig. 1 | Transgenic FCA–eGFP is functionally equivalent to endogenous FCA. a, Top, genomic *FCA* locus indicating upstream and downstream genes (grey), and the position of the *fca-1* mutation. Bottom, illustration of transgenic *FCA-eGFP* construct. Thick black boxes indicate exons, thin black boxes indicate UTRs and black lines indicate introns. **b**, Flowering time of indicated plants grown in a long-day photoperiod. Data are presented as mean ± s.d. ($n = 20$). Asterisk indicates a significant difference ($P = 0.0001$, two-tailed *t*-test). **c**, Expression of spliced *FLC* relative to *UBC* in the indicated plants. Data are presented as mean ± s.d. ($n = 3$). Asterisk indicates a significant difference ($P = 0.0004$, two-tailed *t*-test). **d**, Expression of spliced *FCAγ* relative to *UBC* in the indicated plants. Data are presented as mean ± s.d. ($n = 3$). Asterisk indicates a significant difference ($P = 0.0003$, two-tailed *t*-test). **e**, The protein levels of FCA and FCA–eGFP in the indicated plants as determined by western blot. Asterisks indicate non-specific signals. Data are representative of two independent experiments. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | FCA undergoes phase separation in vitro. **a**, Coomassie staining of indicated protein samples before and after TEV cleavage of the MBP tag. Arrowheads indicate the proteins labelled on top of the gel. Data are representative of three independent experiments. **b**, Phase separation of GFP–FCA in the presence of *Arabidopsis* total RNA was tested using 3.13 μM GFP–FCA and RNA ranging from 0.09 to 1.3 μg ml^{-1}. Scale bars, 10 μm. Data are representative of three independent experiments. **c**, In vitro phase separation assay of GFP–FCA–PrLD at various protein concentrations. Scale bar, 50 μm. Data are representative of three independent experiments. **d**, FRAP of GFP–FCA puncta. Time 0 indicates the time of the photobleaching pulse. Scale bar, 1 μm. Data are representative of eight independent experiments. **e**, Plot showing the time course of the recovery after photobleaching GFP–FCA puncta. Data are presented as mean ± s.d. ($n = 8$). **f**, GFP–FCA puncta do not grow in size or coalesce with each other. Time points are indicated above (in minutes). Scale bar, 10 μm. Data are representative of three independent experiments. **g**, FRAP of GFP–FCA puncta in the presence of 10% (w/v) PEG. Time 0 indicates the time of the photobleaching pulse. Scale bar, 2 μm. Data are representative of nine independent experiments. **h**, Plot showing the time course of the recovery after photobleaching GFP–FCA puncta in the presence of 10% (w/v) PEG. Data are presented as mean ± s.d. ($n = 9$). **i**, Fusion of GFP–FCA puncta in the presence of 10% (w/v) PEG. Time points are indicated above (in minutes). Scale bar, 2 μm. Data are representative of three independent experiments. **j**, FRAP of GFP–FCA puncta in the presence of *Arabidopsis* total RNA. Time 0 indicates the time of the photobleaching pulse. Scale bar, 1 μm. Data are representative of eight independent experiments. **k**, Plot showing the time course of the recovery after photobleaching GFP–FCA puncta in the presence of *Arabidopsis* total RNA. Data are presented as mean ± s.d. ($n = 8$).

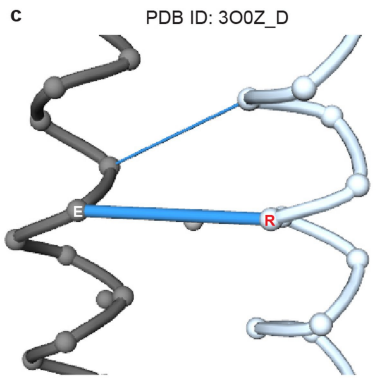


Extended Data Fig. 3 | Characterization of the *sof78* mutation. a, The seed development (top) and the petal number (bottom) of *sof78* mutant and L*er* wild type. Photos are representative of at least five independent experiments. **b**, FLC–LUC bioluminescence signal of indicated plants taken by charge-coupled device camera. Data are representative of three independent experiments. **c**, Expression of spliced *FLC* relative to *UBC* in the indicated genotypes. Data are presented as mean ± s.d. ($n = 4$). Asterisk indicates a significant difference ($P = 0.0014$, two-tailed *t*-test). **d**, Flowering time of indicated plants grown in long-day photoperiod. Data are presented as mean ± s.d. ($n = 20$). Asterisk indicates a significant difference ($P < 0.0001$, two-tailed *t*-test). **e**, PCR with reverse transcription (RT–PCR) detection of *FLC* and *UBC* transcripts, or PCR amplification of indicated fragments from genomic DNA. Data are representative of three independent experiments. **f**, Flowering time of indicated plants grown in long-day photoperiod. Data are presented as mean ± s.d. ($n = 20$). Asterisk indicates a significant difference ($P < 0.0001$, two-tailed *t*-test). **g**, Flowering time of indicated plants grown in long-day photoperiod. Data are presented as mean ± s.d. ($n = 8$). Asterisks indicate significant differences ($P \leq 0.0077$, two-tailed *t*-test). **h**, Top, genomic *FLL2* locus, indicating upstream and downstream genes and positions of *sof78* mutation and *fll2-2* transfer DNA insertion. Bottom, illustration of transgenic *FLL2-eYFP* construct. Thick black boxes indicate exons, thin black boxes indicate UTRs and black lines indicate introns. **i**, RT–PCR detection of *FLL2* and *UBC* transcripts in Col-0 and *fll2-2* plants. Data are representative of three independent experiments. **j**, Expression of spliced *FLC* relative to *UBC* in the indicated genotypes. Data are presented as mean ± s.d. ($n = 4$). **k**, Flowering time of indicated plants grown in long-day photoperiod. Data are presented as mean ± s.d. ($n = 12$). **l**, Phylogenetic tree of proteins of the FLX family. The tree was drawn using the PHYLIP program. Bootstrap values from 1,000 trials are shown. **m**, FLC–LUC bioluminescence signal of indicated plants taken using a charge-coupled device camera. Data are representative of three independent experiments. **n**, Flowering time of indicated plants grown in long-day photoperiod. Data are presented as mean ± s.d. ($n = 10$). Asterisk indicates a significant difference ($P < 0.0001$, two-tailed *t*-test).

a

PDB ID	Protein Name	Propability	E-value
6F1T_X	Actin Related Protein 1 (ARP1)	97.27	0.012
5NNV_C	coiled-coil segment of Smc	96.61	0.041
1C1G_A	Tropomyosin	96.60	0.031
5XG2_A	coiled-coil segment of Smc	96.52	0.038
3O0Z_D	coiled-coil domain of ROCK1	96.40	0.065
1I84_S	Muscle Myosin heavy chain	96.24	0.33
2EFR_A	C-terminal of Tropomyosin	96.15	0.094
5Y06_A	Coiled coil of MSMEG	96.13	0.15
4ILO_A	Coiled coil of CT398	96.03	0.18
4CGK_B	Coiled coil of PcsB	95.88	0.24

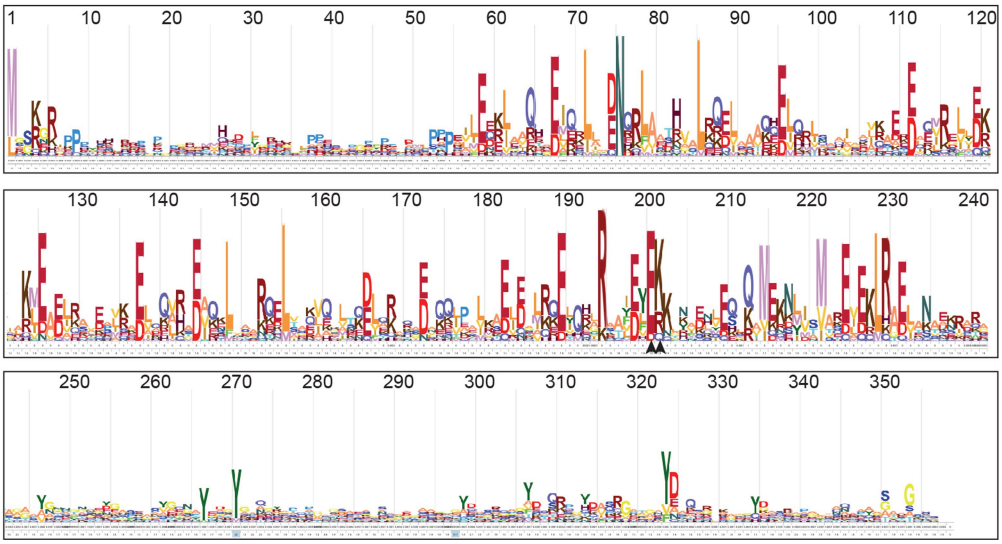
c



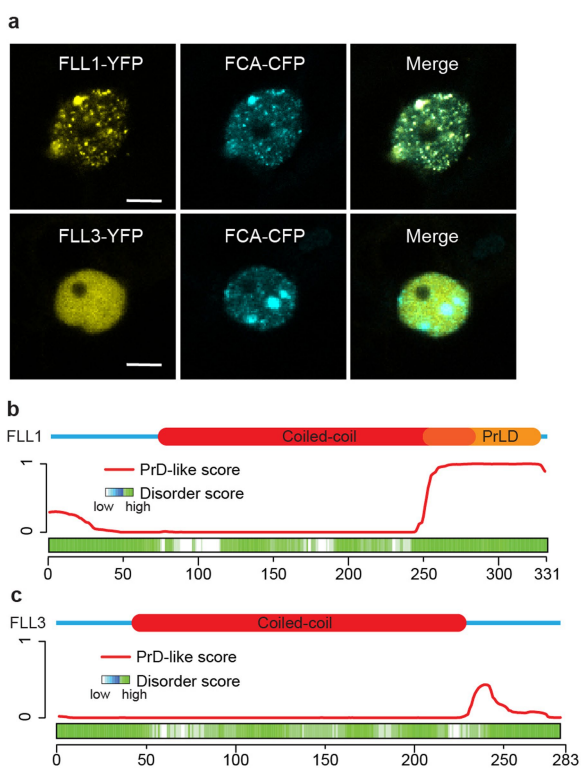
b

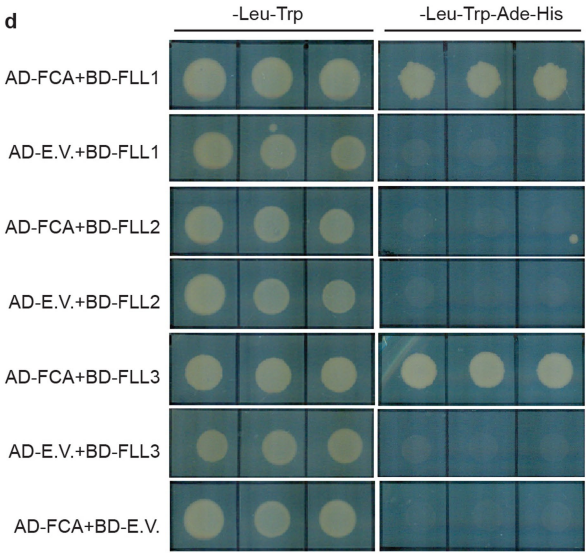
```
  FLL2  65 EKVAKMETELQKSEAVKLEMQQARAEARSLVVAREELMSKVHQLTQELQKSRSDVQQIPALMSELENLRQ
           |++..||-.+    ..+...|.........+|.....||.......+.+.+...+.....    |...++.+.+
3O0Z_D   6 EKLSQLQKQL---EEANDLLRTESDTAVRLRKSHTEMSKSISQLESLNRELQERNRI---LENSKSQTDK

  FLL2 135 EYQQCRATYDYEKKFYNDHLESLQAMEKNYMTMAREVEKLQAQLM
           ++..+.+.++.|+..+....+.....+..+.++-.+++.|+..+.
3O0Z_D  70 DYYQLQAILEAERRDRGHDSEMIGDLQARITSLQEEVKHLKHNLE
```

d

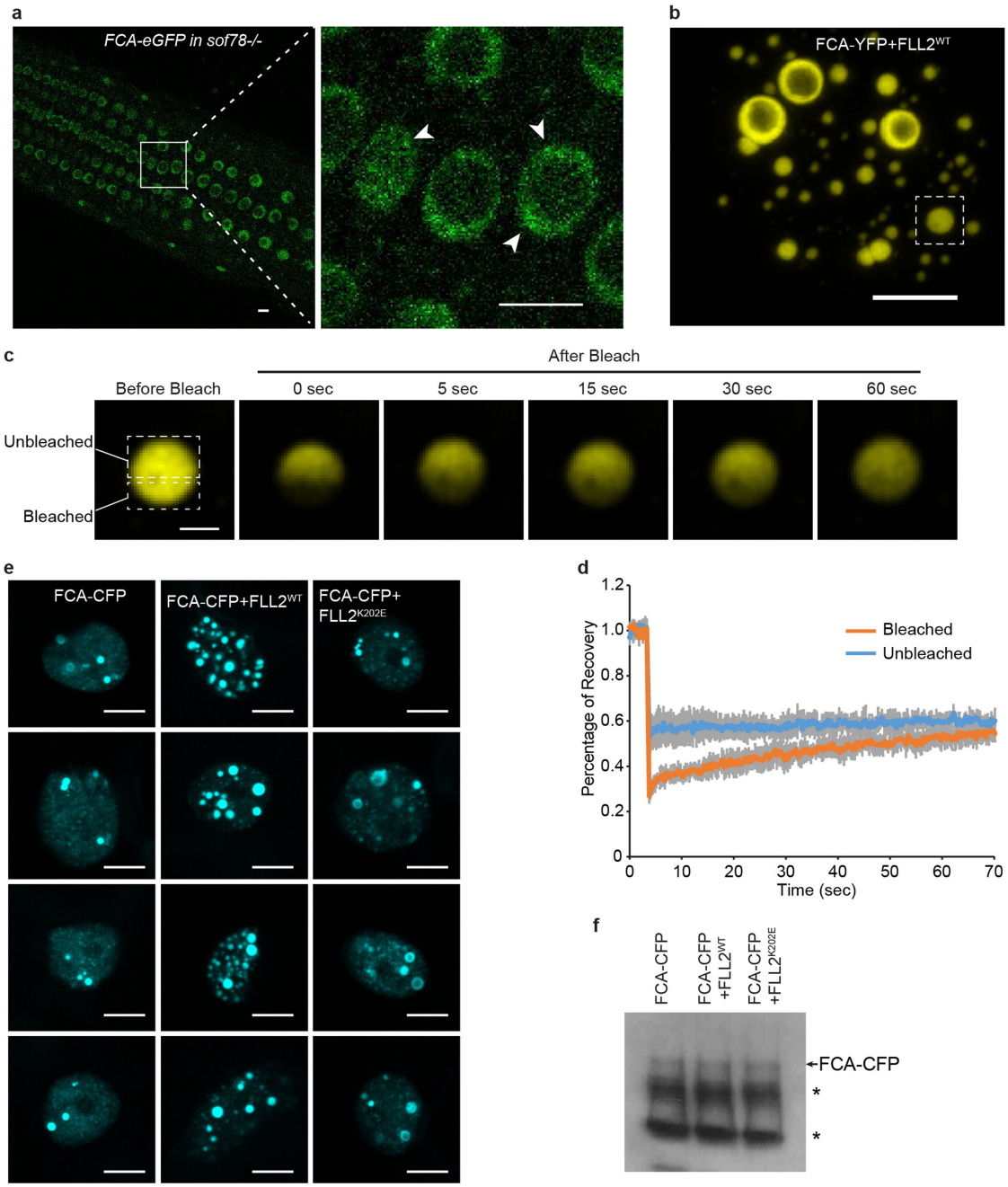


Extended Data Fig. 4 | *FLL2* encodes a coiled-coil domain protein. **a**, A fragment (55–243 amino acids) of FLL2 protein was blasted against the 'PDB_mmCIF70_5_Oct' database using HHpred of the MPI Bioinformatics Toolkit (https://toolkit.tuebingen.mpg.de/#/). The top 10 hits are shown. When the probability is larger than 95%, the homology is nearly certain. **b**, The alignment between coiled-coil domains of FLL2 and human ROCK1. Black arrowhead indicates the Glu that is mutated in *sof78*. **c**, A salt bridge was formed between Glu and Arg (indicated by red arrowheads in **b**) on two molecules of ROCK1. Data were obtained from the Protein Contacts Atlas (http://www.mrc-lmb.cam.ac.uk/pca/). **d**, Plot showing the sequence conservation of FLL2. Analysis was done using the HmmerWeb server (https://www.ebi.ac.uk/Tools/hmmer/) by searching with *A. thaliana* FLL2 within the taxonomy of plants 'Ensembl genomes plants', which yielded 520 homologues within the Streptophyta. The HMM logo shows the conservation for each amino acid for the 520 homologues. Black arrowheads indicate the two amino acids that are predicted to form a salt bridge.

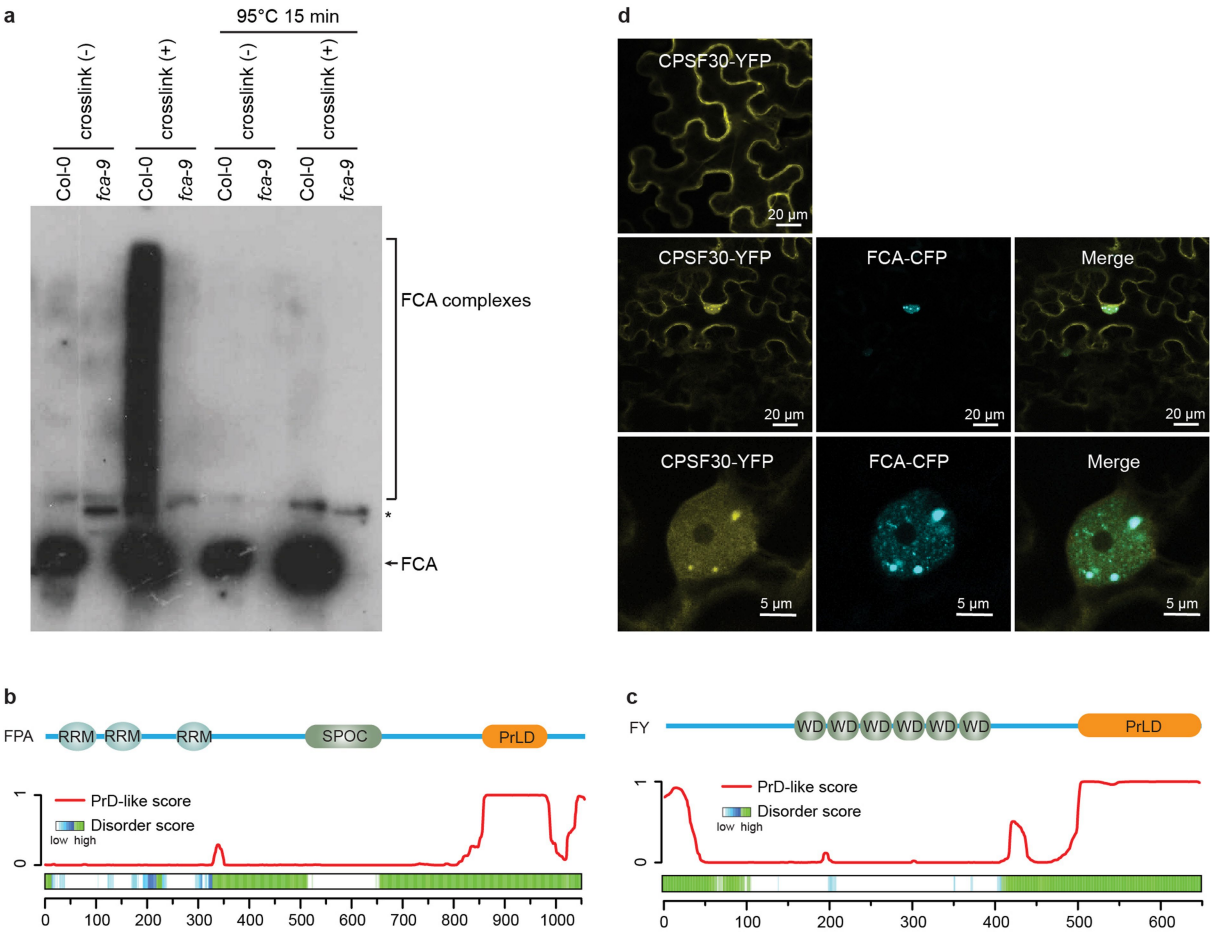




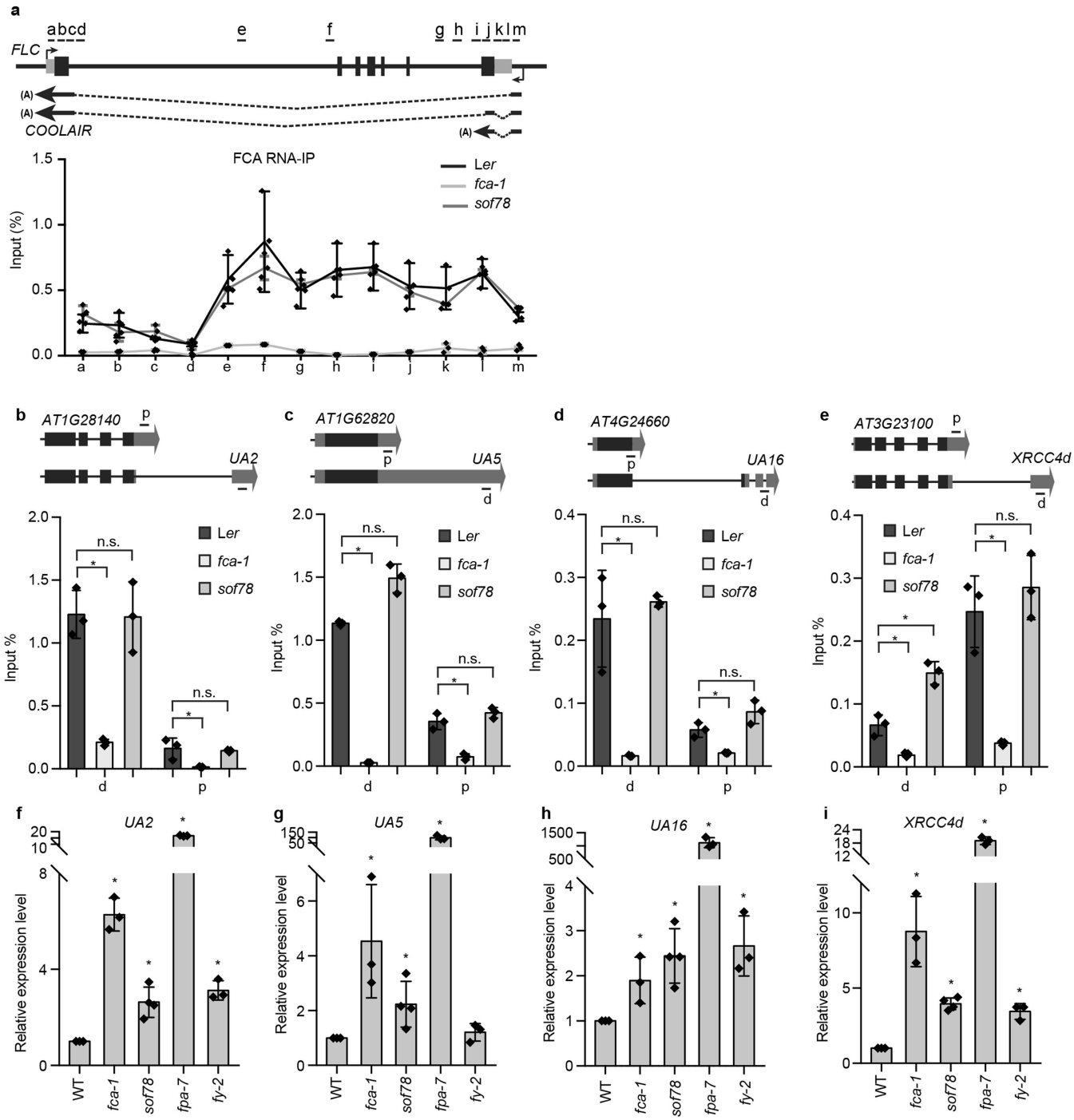
Extended Data Fig. 5 | Characterization of FLL1 and FLL3. **a**, Colocalization of FLL1–YFP and FLL3–YFP with FCA–CFP in tobacco leaf nuclei. Scale bars, 5 μm. Data are representative of three independent experiments. **b**, **c**, Top, protein domain structures of FLL1 and FLL3. Bottom, predictions of PrLDs and disordered regions by PLAAC and D^2P^2 algorithms, respectively. **d**, Interactions of FCA with FLL proteins in yeast two-hybrid assay. Combinations of constructs were transformed into yeast AH109 strain, and assayed on stringent medium. Three independent colonies were tested. AD, activating domain; BD, binding domain; E.V., empty vector. Data are representative of three independent experiments.



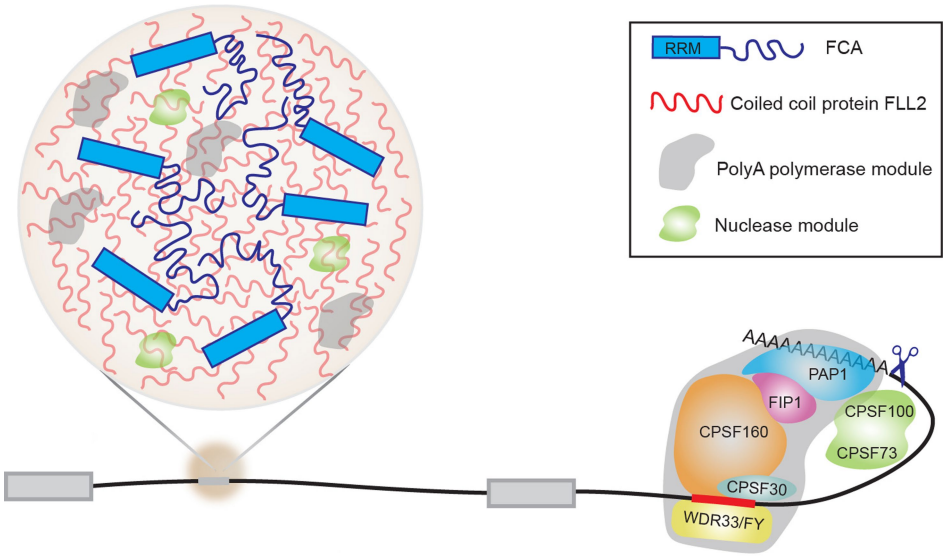
Extended Data Fig. 6 | FLL2 promotes the formation of FCA nuclear bodies. a, An example image showing the FCA–eGFP nuclear bodies in *sof78*-mutant background. A seven-day-old *Arabidopsis* root tip was observed under the confocal microscope. The region indicated in the left panel is shown magnified in the right panel. Scale bars, 5 μm. Image is representative of eight independent experiments. **b**, A tobacco nucleus overexpressing FCA–YFP and FLL2. Scale bar, 5 μm. Data are representative of six independent experiments. **c**, Half-bleach of the FCA–YFP body indicated in **b**. Time 0 indicates the time of the photobleaching pulse. Scale bar, 1 μm. Data are representative of six independent experiments. **d**, Plot showing the time course of the recovery after photobleaching the FCA body. Data are presented as mean ± s.d. ($n = 6$). **e**, The effect of FLL2 overexpression on the pattern of FCA–CFP nuclear bodies, assayed in tobacco leaf nuclei. Scale bars, 5 μm. Data are representative of three independent experiments. **f**, The protein level of FCA–CFP in indicated samples as determined by western blot. Asterisks indicate non-specific signal. Data are representative of three independent experiments. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 7 | Analysis of FCA, FPA, FY and CPSF30. a, In vivo formaldehyde crosslinking gives heterogenous FCA complexes that are much larger than those obtained without crosslinking. Nuclear extracts were prepared from crosslinked or non-crosslinked plants; half of the extracts were mixed with NuPAGE LDS sample buffer and boiled at 70 °C and, for the remaining half, extracts were reverse-crosslinked by heating at 95 °C. Samples were analysed by western blot using FCA antibody. Data are representative of two independent experiments. For gel source data, see Supplementary Fig. 1. **b**, **c**, Top, the annotated functional domains of FPA (**b**) and FY (**c**). Bottom, predictions of PrLDs and disordered regions by PLAAC and D^2P^2 algorithms, respectively. **d**, The colocalization of CPSF30–YFP with FCA–CFP. CPSF30–YFP alone (top) or together with FCA–CFP (middle and bottom) are expressed in tobacco leaves. Images are representative of three independent experiments. Scale bars, 20 μm (top and middle), 5 μm (bottom).



Extended Data Fig. 8 | Effect of *sof78* mutation on the binding of FCA to the nascent transcripts of *COOLAIR* and unannotated segments, and proximal polyadenylation of unannotated segments. a–e, RNA immunoprecipitation and qPCR analysis of FCA enrichment on the transcripts of *COOLAIR* (**a**), *AT1G28140* and unannotated segment 2 (UA2) (**b**), *AT1G62820* and unannotated segment 5 (UA5) (**c**), *AT4G24660* and unannotated segment 16 (UA16) (**d**), and *AT3G23100* and distal *XRCC4* (*XRCC4d*) (**e**). Gene structures are shown above the plots. Data are presented as mean ± s.d. ($n = 3$). Asterisks indicate significant differences ($P \leq 0.0381$, two-tailed *t*-test). Short black lines indicate positions of the primers that were used for qPCR amplification. **f–i**, The expression levels of distally polyadenylated isoforms of UA2 (**f**), UA5 (**g**), UA16 (**h**) and *XRCC4d* (**i**) in the indicated plants relative to wild type. Data are presented as mean ± s.d. ($n = 3$). Asterisks indicate significant differences ($P \leq 0.0099$, two-tailed *t*-test).



Extended Data Fig. 9 | A working model for the role of the coiled-coil protein FLL2 to promote nuclear bodies that are important for polyadenylation at specific sites. At efficient polyadenylation sites, the cleavage and polyadenylation specificity factor complex specifically recognizes the *cis*-acting motif upstream of the cleavage site, catalyses pre-mRNA cleavage and recruits polyadenylation polymerase to initiate polyadenylation. At other sites, phase-separated FCA bodies compartmentalize 3′-end processing factors to enhance polyadenylation.

Extended Data Table 1 | List of proteins identified by FCA and FLL2–HA affinity purification, with or without formaldehyde crosslinking

Identified Protein	Accession No.	Mol. Mass (kDa)	No. of Matched Unique Peptides							
			FCA IP Crosslinking			FCA IP No crosslinking		FLL2-HA IP Crosslinking		
			IP1	IP2	IP3	IP1	IP2	IP1	IP2	IP3
FCA	AT4G16280	92	42	43	40	33	60	4	2	2
FPA	AT2G43410	100	26	27	35	0	0	17	14	9
FY	AT5G13480	72	8	9	10	0	0	0	0	0
CPSF160	AT5G51660	158	6	16	9	0	0	3	3	2
CPSF100	AT5G23880	82	3	10	6	0	0	0	0	0
CPSF73	AT1G61010	77	8	9	8	0	0	0	2	0
CPSF30	AT1G30460	28/70	2	3	3	0	0	0	0	0
CFIS2	AT4G25550	23	11	14	12	0	0	7	3	5
AtFIP1	AT5G58040	133	9	7	6	0	0	0	3	0
FLL2	AT1G67170	40	4	3	4	0	0	21	22	25

nature research

Corresponding author(s): Caroline Dean, Pilong Li

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.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	All movies and fluorescence microscopic images were collected on Zeiss LSM780 confocal microscope, Zeiss LSM880 confocal microscope, Andor Revolution XD Spinning disc confocal microscope system or NIKON A1 microscope as indicated in the Methods.
Data analysis	ImageJ (Version 1.49), Microsoft Excel (2011), MaxQuant software package (version1.5.5.1), Zen Black, MPI Bioinformatics Toolkit, Protein Contacts Atlas, HmmerWeb

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 upon request. 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 raw data that support the findings of this study are available from the corresponding author upon reasonable request. Full lists of Mass spectrometry were provided as Supplementary Tables 1-4.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	No statistical approach was used to predetermine sample size. The determined sample size was adequate as the differences between experimental groups was significant and reproducible.
Data exclusions	No data was excluded from analysis.
Replication	All data except for the immunoblots are representative of at least three independent biological replicates. The immunoblot data are representative of two independent biological replicates. Similar results were obtained from independent biological replicates.
Randomization	Randomization of samples were performed. Seedlings from different plates were collected. Multiple plants were scored for flowering time.
Blinding	Blinding was not possible as the author who performed the experiment also analyzed the data.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	For Western blot: FCA (homemade; 1:8,000), HA (Sigma H3663, Clone HA-7; 1:2,000), GFP (Roche 11814460001, a mixture of clones 7.1 and 13.1; 1:3,000), secondary antibodies anti-Mouse IgG (GE NXA931; 1:40,000), anti-Rabbit IgG (GE NA934; 1:20,000) For immunoprecipitation: FCA (homemade; 1:1,000), anti-HA Magnetic Beads (Thermo Scientific 88836; 1:1,000), anti-FLAG M2 Magnetic Beads (Sigma M8823: 1:1,000)
Validation	FCA (homemade) - Arabidopsis total extract from fca mutant did not produce any signal (Extended Data Fig. 1e), HA (Sigma H3663, Clone HA-7) - validation stated by the manufacturer "a working concentration of 0.25-1 µg/ml is determined using cell extracts expressing N-terminal HA fusion proteins", GFP (Roche 11814460001, a mixture of clones 7.1 and 13.1) - validation stated by the manufacturer "Anti-GFP can be used to

verify the expression of Green Fluorescernt Protein (GFP) and GFP fusion proteins by western blot analysis. It also works well for the immunoprecipitation of GFP and GFP fusion proteins. Anti-GFP recognises both wild- type and mutant forms of GFP."
anti-HA Magnetic Beads (Thermo ScientificTM, 88836) - validation stated by the manufacturer "Pierce Anti-HA Magnetic Beads are compatible with immunoprecipitation and analyses by Western blot",
anti-FLAG M2 Magnetic Beads (Sigma, M8823) - validation stated by the manufacturer "Suitable for immunoprecipitation procedures"
anti-Mouse IgG (GE NXA931) and anti-Rabbit IgG (GE NA934) - validation stated by the manufacturer "Optimized for use with our range of ECL™ Detection Reagents."