Linking key steps of microRNA biogenesis by TREX-2 and the nuclear pore complex in Arabidopsis

Bailong Zhang1, Chenjiang You1,2, Yong Zhang1, Liping Zeng1, Jun Hu3, Minglei Zhao1,4 and Xuemei Chen1✉

Unlike in metazoans, the stepwise biogenesis of microRNAs (miRNAs) in plants occurs within the nucleus. Whether or how the major steps in miRNA biogenesis are coordinated is largely unknown. Here we show that the plant TREX-2 complex promotes multiple steps in miRNA biogenesis, including transcription, processing and nuclear export. THP1 and SAC3A—the core subunits of TREX-2—interact and colocalize with RNA polymerase II to promote the transcription of MIR genes in the nucleoplasm. TREX-2 interacts with the microprocessor component SERRATE and promotes the formation of dicing bodies in the nucleoplasm. THP1 also interacts and colocalizes with the nucleoporin protein NUP1 at the nuclear envelope. NUP1 and THP1 promote the nuclear export of miRNAs and ARGONAUTE1. These results suggest that TREX-2 coordinates the transcription, processing and export steps in miRNA biogenesis to ensure efficient miRNA production.

Multicellular organisms use gene silencing as a means of regulating gene expression and genome stability. Small RNAs associate with ARGONAUTE (AGO) proteins and serve as determinants of sequence specificity in gene silencing1–2. In plants, the two main classes of small RNAs are miRNAs and small interfering RNAs (siRNAs). miRNAs are largely 21–22 nucleotides (nt) long, associate with AGO1, and regulate target-gene expression through mRNA cleavage and translational repression3,4. The crucial roles of miRNAs in diverse biological processes are reflected in the pleiotropic defects exhibited by mutants in miRNA biogenesis genes5.

In plants, miRNA biogenesis entails distinct steps: MIR genes are transcribed by DNA-dependent RNA polymerase II (Pol II) into primary miRNAs (pri-miRNAs). The pri-miRNAs undergo processing by a protein complex containing DICER-LIKE 1 (DCL1), HYponastic LEAVES1 (HYL1) and SERRATE (SE). The resulting miRNA–miRNA* duplexes are 2′–5′-methylated at the 3′ ends by the small-RNA methyltransferase HUA ENHANCER1 (HEN1). The miRNA strand is incorporated into AGO1 and the miRNA–AGO1 complexes, also known as miRNA-induced silencing complexes (miRISCs), are exported to the cytoplasm3,4. Several factors that facilitate and/or regulate these steps have been identified. The transcription of MIR genes requires Mediator1,3, Elongator3, NEGATIVE ON TATA LESS2 (NOT2), and MOS4-ASSOCIATED COMPLEX (MAC) subunits PRL1, PRL2, MAC3, MAC7 (refs.5–10) and CELL DIVISION CYCLE 5 (CDC5)11. Besides association with Pol II to promote transcription, NOT2, CDC5 and Elongator also associate with the dicing complex12,13. In fact, DCL1 is present at MIR genes (as determined by chromatin immunoprecipitation) in an Elongator-dependent manner1, suggesting that nascent pri-miRNAs undergo processing. Additionally, DCL1 and HYL1 are present both in the nucleoplasm and in nuclear dicing bodies (D-bodies)12. As pri-miR173 is found in D-bodies, D-bodies are thought to be sites of pri-miRNA processing11,12. After pri-miRNA processing, the loading of miRNAs into AGO1 is aided by HEAT SHOCK PROTEIN 90 (HSP90), which forms a complex with AGO1 (ref.13). Two importin-β family proteins, ENHANCED MiRNA ACTIVITY1 (EMA1) and TRANSPORTIN1 (TRN1), interact with AGO1 to inhibit or promote, respectively, the loading of miRNAs into AGO1 (refs.14,15). Although AGO1 is a cytoplasmic protein at the steady-state level, it undergoes nucleocytoplastic shuttling and probably loads miRNAs in the nucleus; AGO1–miRNA complexes are exported to the cytoplasm via a nuclear export signal in AGO1 (ref.16).

The nuclear envelope serves as a barrier to restrict certain molecular processes to either the nucleus or the cytoplasm16. miRNAs and small RNAs produced in the nucleus are transported through nuclear pore complexes (NPCs) to the cytoplasm where they function17,18. The evolutionarily conserved TREX-2 complex links transcription and mRNA export19. Yeast TREX-2 consists of Sac3, Thp1, Sem1 (Dss1), Sus1 and Cdc31 (Cen1)20 and is thought to be anchored to the inner side of the NPCs via the nucleoporin Nup1, which is localized to the nuclear pore basket21. Yeast TREX-2 directly associates with Mediator to coordinate transcription and mRNA export22,23.

The role of TREX-2 in transcription and mRNA export is best understood in yeast, an organism that does not have miRNAs. It is unknown whether TREX-2 has a role in miRNA biogenesis or export in plants or animals. Arabidopsis TREX-2 has been characterized in terms of its composition through yeast two-hybrid (Y2H) analyses24. It contains homologues of most of the yeast subunits except for Sus1, and these homologues are encoded by paralogous genes such as SAC3A, SAC3B, SAC3C, DSS1-(I), DSS1-(V), CEN1 and CEN2 (ref.24). However, the molecular and biological functions...
of plant TREX-2 are largely unknown. Here we show that mutants in the TREX-2 core component THP1 have defects in miRNA biogenesis and miRNA nuclear export. THP1 and another core TREX-2 component, SAC3A, physically interact with the miRNA transcriptional machinery Pol II, CDC5, and NOT2B. They also interact with SE and C-TERMINAL DOMAIN PHOSPHATASE-LIKE 1 (CPL1) (also known as FRIEY2 (FRY2)) to promote HYL1 protein dephosphorylation and D-body formation. THP1 colocalizes with Pol II and SE in the nucleoplasm and with NUP1 (also known as NUP136), a homologue of yeast Nup1, at the nuclear envelope. Reduction of THP1 or NUP1 function reduces the nuclear export of both miRNAs and AGO1, suggesting that TREX-2 promotes the export of miRISCs. In summary, TREX-2 coordinates miRNA biogenesis steps from transcription and processing to nuclear export.

Results

A mutation in THP1 causes a global reduction in miRNA accumulation. To identify players in miRNA biogenesis and/or activity, we performed an ethyl methanesulfonate (EMS) mutagenesis screen using the pSUC2: amiR-SUL (ams) line\(^2\). ams expresses the amiR-SUL artificial miRNA from the SUC2 promoter, which is specific for phloem companion cells. Silencing of SULFUR (SUL, also known as CHLORINA42, a gene required for chlorophyll synthesis\(^3\)) by amiR-SUL causes bleaching along the leaf veins. We isolated one mutant with reduced leaf bleaching, a phenotype indicative of compromised amiR-SUL activity (Fig. 1a). We later found that this mutation was in the gene encoding amiR-SUL (Extended Data Fig. 1c), in this mutant. To determine whether endogenous miRNAs were also affected in this mutant, we performed RNA gel blot analyses to examine seven miRNAs in three biological replicates. All examined miRNAs showed reduced accumulation in the thp1-5 ams mutant (Fig. 1b). Consistent with this, the expression of miRNA-target genes was increased (Fig. 1c). We also performed RNA-sequencing (RNA-seq) analysis of small RNAs with PTH1:THP1-eYFP and Col (wild type, used here as a negative control) using green fluorescent protein (GFP)-Trap beads. Three SAC3 proteins, SAC3A, SAC3B and SAC3C, and the nucleoporin NUP1 were recovered from the immunoprecipitation from pTHP1:THP1-eYFP with the most abundant peptide hits (Fig. 2a and Supplementary Table 2). We also examined direct protein–protein interactions of Arabidopsis TREX-2 components by Y2H analyses. THP1 directly interacted with SAC3A, SAC3B, DSS1-(1) and DSS1-(V), but not with CEN1 or CEN2, and SAC3B interacted with CEN1 (Extended Data Fig. 4a). Thus, the IP–MS and Y2H analyses, together with a previous study\(^4\), showed that THP1, SAC3, DSS1 and CEN proteins are subunits of Arabidopsis TREX-2, and that multiple forms of TREX-2 with distinct paralogues of SAC3, DSS or CEN proteins may exist (Fig. 2b). The absence of DSS1-(1) and DSS1-(V) in the THP1–eYFP IP–MS experiment might be because the proteins are short (70–100 amino acids) or are not an integral component of TREX-2. SUS1 in yeast is considered to be a linker protein that associates with both the SAGA complex and TREX-2 in co-regulation of gene transcription and miRNA turnover\(^5\). We did not identify Arabidopsis SUS1 (also known as ENY2) in our IP–MS experiment. A recent study showed that SUS1 associates with Arabidopsis SAGA but not with TREX-2 (ref. \(^6\)).

We next examined whether the SAC3 genes are required for miRNA biogenesis. The single mutants sac3a-2, sac3b-1, sac3b-2 and sac3c-1 were reported to not show any morphological defects\(^6\). We obtained two mutant alleles for each SAC3 gene (Extended Data Fig. 4b) and interrogated levels of SAC3 transcripts in the mutants by RT-PCR using different pairs of primers along the genes (Extended Data Fig. 4c). Judging from the levels of transcripts from the mutant alleles, sac3a-2 appeared to be a stronger allele than sac3a-1; sac3b-2 might be a stronger allele than sac3b-1; and sac3c-1 might be stronger than sac3c-2. We generated double mutants of SAC3B and SAC3C genes. The sac3b-1 sac3c-1 double mutant showed morphological defects (Fig. 2c). We failed to obtain the sac3b-2 sac3c-1 double mutant but were able to obtain the sac3b-1/sac3b-2 sac3c-1/sac3c-1 mutant, suggesting that the sac3b-2 sac3c-1 double mutant with two strong alleles is lethal. The developmental defects of sac3b-1/sac3b-2 sac3c-1/sac3c-1 were similar to those of the thp1-5 mutant (Figs. 1a and 2c). We then determined miRNA levels in sac3 mutants by northern blotting. The levels of eight miRNAs were slightly reduced in sac3b-1 sac3c-1 but were more severely reduced in sac3b-1/sac3b-2 sac3c-1/sac3c-1 (Fig. 2d). Indeed, the sac3b-1/sac3b-2 sac3c-1/sac3c-1 mutant was very similar to the thp1-5 mutant in

phenotype of ams, with ago1-27 showing the least level of suppression (perhaps owing to its being a weak allele) (Extended Data Fig. 3a). Next, we crossed the thp1-5 ams mutant with these mutants. Interestingly, we were only able to obtain viable double mutant plants for ago1-27 thp1-5 ams. The ago1-27 thp1-5 ams mutant showed more severe developmental defects as well as weaker leaf bleaching than ago1-27 ams or thp1-5 ams mutants (Extended Data Fig. 3a). We examined the genetic segregation of plants homozygous for hy1-2 ams, se-1 ams, hst-6 ams or hen1-8 ams and heterozygous for thp1-5. The progeny of these plants segregated 2:1 for thp1-5/+ and +/+, suggesting that double mutants between thp1-5 and these mutations were embryonic lethal (Extended Data Fig. 3b). All null mutants in DCL1 are embryonic lethal\(^7\); the embryonic lethality of the double mutants may be due to stronger defects in miRNA biogenesis than the respective single mutants.

SAC3 genes are required for miRNA biogenesis. THP1 is a subunit of TREX-2 and is conserved in yeast, animals and plants\(^8\). We investigated whether the role of THP1 in miRNA biogenesis reflected that of TREX-2. We first characterized Arabidopsis TREX-2 using immunoprecipitation followed by mass spectrometry (IP–MS). Immunoprecipitation was performed with pTHP1:THP1-eYFP and Col (wild type, used here as a negative control) using green fluorescent protein (GFP)-Trap beads. Three SAC3 proteins, SAC3A, SAC3B and SAC3C, and the nucleoporin NUP1 were recovered from the immunoprecipitation from pTHP1:THP1-eYFP with the most abundant peptide hits (Fig. 2a and Supplementary Table 2). We also examined direct protein–protein interactions of Arabidopsis TREX-2 components by Y2H analyses. THP1 directly interacted with SAC3A, SAC3B, DSS1-(1) and DSS1-(V), but not with CEN1 or CEN2, and SAC3B interacted with CEN1 (Extended Data Fig. 4a). Thus, the IP–MS and Y2H analyses, together with a previous study\(^4\), showed that THP1, SAC3, DSS1 and CEN proteins are subunits of Arabidopsis TREX-2, and that multiple forms of TREX-2 with distinct paralogues of SAC3, DSS or CEN proteins may exist (Fig. 2b). The absence of DSS1-(1) and DSS1-(V) in the THP1–eYFP IP–MS experiment might be because the proteins are short (70–100 amino acids) or are not an integral component of TREX-2. SUS1 in yeast is considered to be a linker protein that associates with both the SAGA complex and TREX-2 in co-regulation of gene transcription and miRNA turnover\(^5\). We did not identify Arabidopsis SUS1 (also known as ENY2) in our IP–MS experiment. A recent study showed that SUS1 associates with Arabidopsis SAGA but not with TREX-2 (ref. \(^6\)).

We next examined whether the SAC3 genes are required for miRNA biogenesis. The single mutants sac3a-2, sac3b-1, sac3b-2 and sac3c-1 were reported to not show any morphological defects\(^6\). We obtained two mutant alleles for each SAC3 gene (Extended Data Fig. 4b) and interrogated levels of SAC3 transcripts in the mutants by RT-PCR using different pairs of primers along the genes (Extended Data Fig. 4c). Judging from the levels of transcripts from the mutant alleles, sac3a-2 appeared to be a stronger allele than sac3a-1; sac3b-2 might be a stronger allele than sac3b-1; and sac3c-1 might be stronger than sac3c-2. We generated double mutants of SAC3B and SAC3C genes. The sac3b-1 sac3c-1 double mutant showed morphological defects (Fig. 2c). We failed to obtain the sac3b-2 sac3c-1 double mutant but were able to obtain the sac3b-1/sac3b-2 sac3c-1/sac3c-1 mutant, suggesting that the sac3b-2 sac3c-1 double mutant with two strong alleles is lethal. The developmental defects of sac3b-1/sac3b-2 sac3c-1/sac3c-1 were similar to those of the thp1-5 mutant (Figs. 1a and 2c). We then determined miRNA levels in sac3 mutants by northern blotting. The levels of eight miRNAs were slightly reduced in sac3b-1 sac3c-1 but were more severely reduced in sac3b-1/sac3b-2 sac3c-1/sac3c-1 (Fig. 2d). Indeed, the sac3b-1/sac3b-2 sac3c-1/sac3c-1 mutant was very similar to the thp1-5 mutant in
Fig. 1 | The *thp1-5* mutant exhibits pleiotropic developmental defects and reduced miRNA levels. **a**, Twenty-one-day-old *pSUC2:amiR-SUL* (*amS*) and *thp1-5 amS* plants with differences in the vein-bleaching phenotype. Scale bars, 10 mm. All plants of the same genotypes exhibit similar phenotypes and one representative is shown. **b**, RNA gel blot analysis of miRNAs from *amS* and *thp1-5 amS* plants. Replicate (Rep) 1 and Rep 2 were from 12-d-old seedlings grown on 0.5x Murashige–Skoog medium; Rep 3 was from 21-d-old seedlings grown in soil. The miRNA signals were quantified and normalized to those of U6; values are shown relative to *amS*. **c**, Determination of miRNA-target mRNA levels in 12-d-old seedlings of *amS* and *thp1-5 amS* plants by quantitative PCR with reverse transcription (RT-qPCR). The housekeeping gene IPP2 was included as a control. Error bars indicate s.d. from three biological replicates. **d**, Abundance of miRNAs in *amS* and *thp1-5 amS* as determined by small-RNA sequencing. Each dot represents a miRNA. Small-RNA libraries from three independent replicates for each background were generated from 12-d-old seedlings grown on Murashige–Skoog plates. small RNAs were normalized against 45S ribosomal RNA reads and abundance was expressed as RPMR (reads per million 45S rRNA reads). **e**, Schematic of the *THP1* gene. Two published *thp1* mutants and *thp1-5* are shown. The triangle represents a T-DNA insertion; the asterisk indicates the stop codon. **f**, Morphological phenotypes of 25-d-old *amS*, *thp1-5 amS* and *thp1-5 amS pTHP1:THP1-eYFP* plants showing that the *pTHP1:THP1-eYFP* transgene rescued the *thp1-5 amS* mutant defects. Scale bars, 10 mm. Note that the transgene was introduced into wild-type plants (Col) and a line with a single transgene insertion and a moderate level of the THP1-eYFP protein was chosen for crosses with *thp1-5 amS*. F2 plants that were homozygous for *thp1-5 amS* and the transgene were identified by genotyping and examined for phenotype. At least three different F2 homozygous plants showed similar phenotypes. **g**, Small-RNA gel blot analysis showing that the *thp1-5 amS* molecular defects were rescued by the *pTHP1:THP1-eYFP* transgene. Values are displayed as in **b**. Two independent repeats gave similar results.
TREX-2 associates with the transcription machinery and promotes MIR gene transcription. To determine which steps of miRNA biogenesis were affected in the thp1-5 amS mutant, we first examined MIR gene transcription. We measured pri-miRNA levels and found that all the tested pri-miRNAs showed reduced levels in thp1-5 amS compared with amS (Fig. 3a). The reduction in pri-miRNA levels may be due to reduced transcription or RNA stability. We crossed the thp1-5 amS mutant with pMIR167a:GUS, a β-glucoronidase (GUS) reporter of MIR167a promoter activity. Histochemical staining for GUS activity and quantitative PCR with reverse transcription (RT-qPCR) for measurement of GUS transcript levels revealed that GUS expression in thp1-5 was reduced compared with WT, suggesting that THP1 promotes MIR gene transcription (Fig. 3b,c). However, TREX-2 is known to prevent the epigenetic silencing of transgenes and some endogenous loci; thus, the reduction in pri-miR-SUL levels or pMIR167a:GUS expression in the thp1-5 amS mutant might be due to transcriptional silencing of the transgenes. This was unlikely, as the expression of the Bar gene in the same transfer DNA (T-DNA) as pSUC2:amiR-SUL was unaltered (Extended Data Fig. 5d).

Since TREX-2 associates with Mediator to regulate inducible gene transcription in yeast22, to probe how Arabidopsis TREX-2 regulates MIR gene transcription, we performed Y2H assays to test the protein–protein interactions between TREX-2 subunits and several Mediator subunits that have been reported to affect MIR gene transcription23. No protein–protein interactions were detected between TREX-2 subunits and these Mediator subunits (Extended Data Fig. 5a). We then tested the interactions between TREX-2 subunits and other factors known to promote MIR gene transcription, such as Pol II, NOT2 and subunits of Elongator and MAC (Fig. 3d and Extended Data Fig. 5b). In these Y2H assays, TREX-2 subunits SAC3A and THP1 interacted with the C-terminal domain of NRPB1, the largest subunit of Pol II, as well as with CDC5 and NOT2B (Fig. 3d). We validated the THP1–Pol II interaction by reciprocal coimmunoprecipitation (Fig. 3e) and the SAC3A–Pol II interaction by bimolecular fluorescence complementation (BiFC) (Extended Data Fig. 5c). Furthermore, we performed Pol II chromatin immunoprecipitation (ChiP) and found that Pol II occupancy at several MIR loci was reduced in the thp1-5 amS mutant (Extended Data Fig. 5e). Together, these results suggest that TREX-2 associates with Pol II and recruits it to MIR loci to promote MIR gene transcription.

To determine the global effects of TREX-2 on gene expression, we performed mRNA sequencing with thp1-5, sac3b-1/sac3b-2 sac3c-1/sac3c-1 and wild type. Three biological replicates of each genotype were highly correlated (Extended Data Fig. 6a). A significant overlap was found for both downregulated and upregulated DEGs in the two mutants (Extended Data Fig. 6b; P = 0.01), indicating that THP1 and SAC3 act in the TREX-2 complex to regulate the same group of genes. A total of 852 and 819 genes were commonly upregulated and downregulated, respectively (Extended Data Fig. 6a and Supplementary Table 3). The Gene Ontology terms enriched in the common DEGs (both upregulated and downregulated) included stress and stimulus responses. Ethylene signalling pathway genes were among the DEGs (Extended Data Fig. 6c,d), consistent with the enhanced ethylene response phenotype of thp1 mutants24,31.

TREX-2 associates with SE and CPL1 and regulates HYL1 phosphorylation status. Our Y2H assays also revealed that TREX-2 subunits THP1 and SAC3A interacted with SE, and SAC3A and SAC3B interacted with both the full-length and C-terminal portion (amino acids 484–967) of CPL1 (Fig. 4a). THP1 and SAC3A did not interact with the other components of the dicing complex, such as DCL1 or HYL1 (Extended Data Fig. 7a). The interactions

miRNA accumulation (Figs. 1b and 2d). Together, these results show that Arabidopsis TREX-2 components THP1 and SAC3 are required for miRNA production.

**Figure 2** | Arabidopsis TREX-2 is required for miRNA production. a, IP–MS using THP1–eYFP (expressed from pTHP1:THP1-eYFP) and NUP1–YFP (expressed from pNUP1:NUP1-YFP) as baits in Arabidopsis. Wild-type plants (Col) without the transgenes served as a negative control. The numbers of recovered peptides corresponding to each of the proteins are shown. b, Yeast and Arabidopsis TREX-2 components. c, Twenty-one-d-old seedlings of Col, sac3b-1/sac3b-2 sac3c-1/sac3c-1 (sac3b-1/b-2 c-1/c-1) and sac3b-1 sac3c-1. Scale bar, 10 mm. d, Small-RNA gel blot analysis to determine miRNA levels in Col, sac3b-1/sac3b-2 sac3c-1/sac3c-1 and sac3b-1 sac3c-1. Two independent repeats yielded similar results. U6 was used as an internal control. The numbers indicate the relative abundance of the miRNAs.
between THP1 and SE were further validated by coimmunoprecipitation (Fig. 4b), and those between SAC3A and SE were further validated by BiFC (Extended Data Fig. 7b). Since CPL1 and SE are responsible for HYL1 protein dephosphorylation and dephosphorylated HYL1 is the active form for miRNA processing\(^{22,33}\), we next examined the ratio of phosphorylated/non-phosphorylated (P+/P−) HYL1 using phos-tag, which decreases the mobility of phosphorylated proteins in polyacrylamide gels. We found that the total HYL1 protein level was slightly increased in the thp1-5 amS mutant (Fig. 4c, Extended Data Fig. 7c,d). However, that of \(\text{P}^+\) HYL1 was markedly increased and the P+/P− ratio was increased in the thp1-5 amS mutant (Fig. 4c), suggesting that inactive HYL1 (\(\text{P}^-\) HYL1) is the active form for miRNA processing, as DCL1 can be detected at D-bodies as compared to wild type (Extended Data Fig. 8b,c). MIR gene transcription is considered to be coupled with pri-miRNA processing, as DCL1 can be detected at D-bodies as compared to wild type (Extended Data Fig. 8b,c). MIR gene transcription is considered to be coupled with pri-miRNA processing, as DCL1 can be detected at D-bodies as compared to wild type (Extended Data Fig. 8b,c). MIR gene transcription is considered to be coupled with pri-miRNA processing, as DCL1 can be detected at D-bodies as compared to wild type (Extended Data Fig. 8b,c).

**Fig. 3** | TREX-2 promotes MIR gene transcription and associates with the transcription machinery. a, RT-qPCR to determine pri-miRNA levels in amS and thp1-5 amS. IPP2 was used as an internal control. Error bars indicate s.d. from three biological replicates. The experiment was repeated three times with similar results. b, GUS staining of pMIR167a::GUS and thp1-5 pMIR167a::GUS inflorescences. Three independent repeats gave similar results. c, Levels of GUS transcripts and pri-miR167a in pMIR167a::GUS and thp1-5 pMIR167a::GUS seedlings as determined by RT-qPCR. IPP2 was used as an internal control. Error bars indicate s.d. from three independent replicates. d, Y2H assays to test the interactions between TREX-2 components (THP1, SAC3A, SAC3B, DSS1-(I), DSS1-(V) and CEN2) and the C-terminal domain (CTD) of NRPB1, NOT2B and CDC5. The empty vectors GAL4-AD and GAL4-BD were included as controls. Diploid yeast cells from mating contained the bait and prey proteins interacted formed blue colonies on SD–Trp–/Leu double-dropout (DDO) medium, and diploid yeast cells in which the bait and prey proteins interacted formed blue colonies on SD–Trp–/–Leu/–His/–Ade/–X-a-gal/aureobasidin A (QDO/X/A) medium. The experiment was repeated two times with similar results. AD, activation domain; BD, DNA-binding domain. e, In vivo coimmunoprecipitation for validation of interactions between THP1 and NRPB1. Total protein extracts from inflorescences of pNRPB1::NRPB1-mRFP, pTHP1::THP1-mCerulean3 and pTHP1::THP1-mCerulean3 were immunoprecipitated (IP) using RFP-Trap beads. The coimmunoprecipitated proteins were detected by western blotting using antibodies for GFP and RFP, respectively. Two independent repeats gave similar results.

**Fig. 4** | TREX-2 promotes miRNA levels in thp1-5 mutants. The miRNA levels in thp1-5 amS and thp1-5 pMIR167a::GUS plants were determined by RT-qPCR. Error bars indicate s.d. from three biological replicates. The experiment was repeated three times with similar results.
in the thp1-5 amS mutant. Nuclear and cytoplasmic fractions were isolated from amS and thp1-5 amS seedlings; the distribution of U6 and the transfer RNA recognizing the triplet codon for methionine (tRNAMet) suggested that the fractionation was largely successful (Fig. 5a). Consistent with a global reduction in miRNA levels in the thp1-5 amS mutant, a reduction in miRNA levels was found in both nuclear and cytoplasmic fractions in the mutant compared with amS (Fig. 5a). Unlike hst mutants, the cytoplasmic/nuclear ratios of miRNAs were reduced in thp1-5 amS (Fig. 5a,b). This indicated that TREX-2 promotes the nuclear export of miRNAs.

We further performed mild formaldehyde fixation followed by nucleocytoplasmic fractionation to prevent the leakage of...
nucleoplasmic proteins or RNAs into the cytoplasm during fractionation. The conditions of formaldehyde fixation were optimized to avoid endomembrane contamination of the nuclear fraction (Extended Data Fig. 9b). HYL1 is a nucleoplasmic protein but tended to leak into the cytoplasm during fractionation in the absence of fixation; it was significantly enriched in the nuclear fraction using our improved method (Extended Data Fig. 9a). We used the improved fractionation method to examine the nucleocytoplasmic distribution of miRNAs and AGO1, and TREX-2 protein localization.

Fig. 5 | Nucleocytoplasmic partitioning of miRNAs and AGO1, and TREX-2 protein localization. a, Small-RNA gel blot analysis to determine the nuclear and cytoplasmic accumulation of miRNAs in amS and thp1-5 amS. U6 and tRNA/m* served as nuclear and cytoplasmic RNA markers, respectively. They also served as the loading controls for the nuclear and cytoplasmic fractions for quantification of miRNA levels. The numbers represent relative levels of miRNAs. Two independent experiments gave similar results. The nucleocytoplasmic partitioning of AGO1 in Col, thp1-5 and hst-6. The fractionated samples were analysed by western blot using AGO1, cFBPase, BIP and H3 antibodies, respectively. H3 is a nuclear marker, cFBPase is a cytoplasmic marker and BIP is an endoplasmic reticulum marker used to indicate the extent of endomembrane contamination in the nuclear fraction. The nuclear (N) and cytoplasmic (C) AGO1 signals were quantified and normalized to H3 and cFBPase, respectively, and values are expressed relative to T (total extract). The experiment was repeated three times with similar results.

b, The cytoplasmic/nuclear (C/N) ratios of miRNAs as determined in a. c, The nucleocytoplasmic partitioning of AGO1 in Col, thp1-5 and hst-6. d–k, Protein colocalization analyses of NRPB1 (pNRPB1:NRPB1-mRFP) with THP1 (pTHP1:THP1-mEGFP) (d), NRPB1 with SAC3A (pSAC3A:SAC3A-mEGFP) (e), SE (pSE:SE-mRuby3) with THP1 (f), SE with SAC3A (g), NRPB1 with NUP1 (pNUP1:NUP1-mEGFP) (h), SE with NUP1 (i), NUP1 (pNUP1:NUP1-mRuby3) with THP1 (j) and NUP1 with SAC3A (k). All fluorescent-protein-fusion transgenic plants were generated in the Col background and crossed with each other. F1 plants were examined by confocal microscopy. Scale bars, 1 µm. At least three plants were examined and gave similar results.

THP1 localizes both in the nucleoplasm and at the nuclear envelope and interacts with NUP1. Yeast TREX-2 localizes at the nuclear envelope, but Arabidopsis TREX-2 associates with many nucleoplasmic proteins, including Pol II and SE. This prompted us to examine the localization of Arabidopsis TREX-2 proteins. We generated transgenic lines harbouring the transgenes pTHP1:THP1-mEGFP, pSAC3A:SAC3A-mEGFP, pNRPB1:NRPB1-mRFP, pSE:SE-mRuby3, pNUP1:NUP1-mEGFP and pNUP1:NUP1-mRuby3 individually and crossed them with one another to perform pairwise comparisons of expression patterns in the root tips of F1 seedlings. NRPB1 and SE colocalized with THP1 and SAC3A in the nucleoplasm (Fig. 5d–g), suggesting that THP1, but not HST, is required for the nuclear export of AGO1. Together, these data strongly support that THP1 promotes the export of miRISCs.
NUP1 promotes miRISC nuclear export. Given that THP1 and NUP1 colocalize and interact with each other, we investigated whether NUP1 has a role in miRNA biogenesis. The null allele nup1-2 is male-gametophytic lethal22, which precludes the analysis of nup1-2 homozygous mutants. During the generation of a series of pNUP1::NUP1-megFP transgenic lines, we observed plants with developmental defects and suspected that they may be nup1 co-suppression lines (nup1-cs) (Fig. 6a and Extended Data Fig. 10a). We examined the levels of NUP1 transcripts in six pNUP1::NUP1-megFP transgenic lines, two with normal morphology and four with developmental defects. RT–PCR was performed with primers that detected both endogenous NUP1 transcripts and those from NUP1-megFP. As expected, the two lines with normal morphology had similar levels of NUP1 RNA compared with Col, whereas the four lines with developmental defects had very low levels of NUP1 RNA (Extended Data Fig. 10b). Consistently, NUP1-megFP protein was only detectable in the two lines with normal morphology (Extended Data Fig. 10c). This suggested that the four lines with low NUP1 mRNA and no NUP1-megFP protein were nup1-cs lines. The strong nup1-cs lines showed abnormal flower development and greatly reduced fertility (Fig. 6a and Extended Data Fig. 10a), and we chose a weaker line that could produce some seeds for further experiments. We collected nup1-cs inflorescence tissue and performed small-RNA sequencing together with Col and thpl-5 for comparison. We found that most miRNAs were similar in abundance between Col and nup1-cs, whereas a global reduction of miRNA levels was found in thpl-5 (Fig. 6b). RNA gel blot assays showed that mir166 and miR172 were similar in abundance between Col and the nup1-cs line (Extended Data Fig. 10d), suggesting that miRNA levels are not altered in nup1-cs.

We then examined the transcript levels of several miRNA-target genes in nup1-cs by RT–qPCR. Most miRNA targets were derepressed in nup1-cs as compared with Col (Fig. 6c), implying that miRNA activities were compromised in nup1-cs. On the basis of the evidence that NUP1 interacted and colocalized with THP1 at the nuclear envelope, we suspected that the nuclear export of miRNAs might be affected in nup1-cs. We examined the nucleocytoplasmic distribution of miRNAs and AGO1 in nup1-cs, finding that both examined miRNAs and the AGO1 protein showed increased levels in the nuclear fraction in nup1-cs (Fig. 6d–f and Extended Data Fig. 10e,f). This indicated that NUP1 also promotes the nuclear export of miRNAs and AGO1.

To probe the possible mechanism by which THP1 and NUP1 promote the nuclear export of miRISCs, we first isolated nuclei and then performed AGO1 IP–MS. We found that NUP1 was among the proteins identified in the nuclear AGO1 IP–MS (Fig. 6g and Supplementary Table 4). Besides NUP1, many other nucleoporins, exportins and importins were associated with nuclear AGO1. These proteins were also present in the lists of protein hits from IP–MS using NUP1-YFP or THP1-eYFP as the bait (Fig. 6g and Supplementary Table 2). These results suggested that miRISCs are exported through the nuclear pore and that THP1, importins and exportins are involved in the export.

Discussion

miRNA biogenesis is a multistep process in the nucleus that culminates in the export of miRISCs to the cytoplasm. Although pri-miRNA processing is probably cotranscriptional, and several factors aid both the transcription and processing steps in miRNA biogenesis, it is not known whether subsequent steps in miRNA biogenesis, such as miRISC formation and miRISC export, are coordinated with the transcription and processing steps. In this study, we show that TREX-2 coordinates the transcription, processing and miRNA-export steps in miRNA biogenesis. We found that TREX-2: (1) interacts with the transcriptional machinery to promote MIR gene transcription; (2) interacts with the dicing complex component SE and the phosphatase CPL1 to dephosphorylate the dicing complex component HYL1 and promote D-body formation, and recruits HYL1 and DCL1 to MIR loci; and (3) promotes miRISC export, possibly through its interactions with the NPC and exportins and importins, including XPO1 (also known as CRM1) (Supplementary Table 4). Coordinating the transcription, processing and export steps in miRNA biogenesis may increase the efficiency of miRNA biogenesis.

TREX-2 is required for the coordination of MIR gene transcription and pri-miRNA processing. In human cells, both intronic and independently transcribed pri-miRNAs are co-transcriptionally processed by Drosha36,37. Plant MIR genes are similar to protein-coding genes in that their promoters contain the TATA-box motif and other transcription factor binding motifs38,39. The transcription of MIR genes requires general and specific transcription regulators39. To date, many MIR transcription regulators, such as NOT2, DDL, Elongator and MAC subunits PRL1, PRL2 and CDC5 have been shown to interact with both Pol II and the dicing complex10,11. Moreover, mutants with defects in D-body formation, such as those in MAC7, PP4R3, PRL1 and CDC5 genes, also exhibit reduced pri-miRNA levels that reflect defects in transcription26–11,40. These studies point to cotranscriptional processing of pri-miRNAs. Indeed, DCL1 has been found to associate with MIR gene promoters in an RNA-dependent manner40. In this study, we found that TREX-2 subunits THP1 and SAC3A interact with Pol II, NOT2B and CDC5 and promote MIR gene transcription. TREX-2 also interacts with the dicing complex component SE and promotes HYL1 dephosphorylation through CPL1. Furthermore, we showed that HYL1 is present at MIR loci. Therefore, our study further supports the conclusion that pri-miRNAs are co-transcriptionally processed.

HYL1 is a major component of the microprocessor and forms subnuclear foci called D-bodies32,41. The activity and dynamics (such as levels and nucleocytoplasmic shuttling) of HYL1 are highly regulated via phosphorylation and dephosphorylation32,33,42. In this study, we show that TREX-2 regulates HYL1 activity, as
demonstrated by (1) interaction of TREX-2 with CPL1, a HYL phosphatase, and the increase in phosphorylated/non-phosphorylated HYL1 ratio in *thp1-5*; (2) compromised D-body formation in *thp1-5*; and (3) a reduction in HYL1 occupancy at *MIR* gene loci in *thp1-5*. HYL1 protein phosphorylation is increased under ethylene treatment\(^3\). *THP1* was also identified as *EER5* (*Enhanced Ethylene Response 5*), which regulates the expression of a subset of ethylene-responsive genes\(^3\). Thus, TREX-2 may also regulate HYL1 protein dynamics to influence miRNA biogenesis in response to environmental signals.

Yeast TREX-2 functions in mRNA export and promotes the transcription of highly inducible genes by targeting these genes to the nucleus.
We note that amiR-SUL was only reduced to 40% of wild-type levels in thp1-5 amS, whereas the leaf-bleaching phenotype was nearly completely suppressed. We suspect that the strong phenotype is contributed by both reduced levels of amiR-SUL and inefficient miRNA nuclear export. Alternatively, THP1 may also promote the cell-to-cell trafficking of amiR-SUL, for which no evidence is available yet.

In summary, our studies show that Arabidopsis TREX-2 associates with Pol II and the microprocessor in the nucleoplasm and couples the transcription and processing steps of miRNA biogenesis. In addition, THP1, together with NUP1, promotes the export of miRISCs to the cytoplasm through the nuclear pore (Fig. 7). The promotion of the transcription, processing and export steps of miRNA biogenesis by TREX-2 could be explained by its dual localization in both the nucleoplasm and at the nuclear pore. Although SAC3A is localized only in the nucleoplasm (Fig. 5k), SAC3B is localized at the nuclear envelope. Thus, it is possible that the nucleoplasmic TREX-2 consists of SAC3A and other subunits, whereas the nuclear pore-localized TREX-2 consists of SAC3B and other subunits. THP1 might be shuttling between the nucleoplasm and the nuclear pore to coordinate nucleoplasmic miRNA biogenesis with miRISC nuclear export.

Methods

Plant materials and growth conditions. All Arabidopsis mutants and transgenic lines are in the Columbia (Col-0) ecotype. All plants were grown in soil or on 0.5x Murashige–Skoog plates in growth rooms at 22 °C under full-spectrum white light under long-day (16h light:8h dark) or short-day (8h light:16h dark) conditions. The pSUC2:amiR-SUL line was from D. Weigel (Max Planck Institute for Developmental Biology, Tubingen, Germany). thp1-5 was isolated from our EMS mutagenesis screen with the pSUC2:amiR-SUL (amS) line; thp1-3 (esp1) and pNUP1:THP1-NUP1-YFP were from Y. Cui (Agriculture and Agri-Food Canada, London, Ontario, Canada); and thp1-1 (ser5-2, SAIL_82_A02) was from the Arabidopsis Biological Resource Center. pMIR167a:GUS and p35S:HYL1-YFP were described. sac3 single mutants were from the Arabidopsis Biological Resource Center and the sac3 double and triple mutants were generated in this study. pTHP1:THP1-eYFP, pTHP1:THP1-meGFP, pTHP1:THP1-mCardenal3, pNRBP1:NRPB1-mRFP, pSE:SE-mRuby3, pSAC3A:SAC3A-meGFP, pNUP1:NUP1-meGFP and pNUP1:NUP1-mRuby3 transgenic lines were generated in this study. To generate the plasmids of the above transgenes, full-length genomic sequences were amplified from the respective mutant DNA, ligated into the pTOK108 vector. The clones were verified by sequencing and inserts with different genes were transferred into the pGWB binary vectors or our newly created pGW binary vectors (Supplementary Table 5) using LR Clonase II enzyme mix (Invitrogen) to generate the T-DNA constructs. The binary vectors were used in Agrobacterium-mediated transformation to produce transgenic plants in the Col background.

EMS mutagenesis and mapping. EMS mutagenesis was performed as described. The identification of the thp1-5 mutation through whole-genome resequencing was performed as described. The thp1-5 amS mutant was back-crossed with the parental line amS line. Genomic DNA was extracted from 150 pooled F2 mutants and used in library construction. The library was paired-end sequenced on Illumina HiSeq 2000 at about 30x coverage and the reads were mapped to the TAIR10 genome. Three candidate mutations (the other two mutated genes are AT2G13540 and AT5G42090) with 100% mutation rate in the reads were EMS-typical C:G to T:A transitions that were predicted to cause nonsynonymous substitutions in the coding regions of genes. Linkage analysis was performed on 96 individual mutant plants in the F2 population of the backcross using three derived cleaved amplified polymorphic sequences (dCAPS) markers to assess linkage between the mutations and the vein-bleaching phenotype. Only the mutation in THP1 was found to be homozygous in all 96 plants, suggesting that this mutation was responsible for the phenotype.

Small-RNA gel blot analysis and sequencing. Small-RNA gel blot analysis was performed as described. Approximately 5–10μg total RNA was extracted from seedling tissue, resolved on 15% urea–PAGE gels. Small-RNA libraries were constructed following instructions from the NEBNext Multiplex Small-RNA Library Prep Set for Illumina (E7300). The libraries were sequenced on an Illumina HiSeq 2500. The resulting data were analysed using our pipeline pRNASeqTools v.0.4.
NATURE PLANTS

PCR-based transcript analysis. To perform RT–PCR, total RNAs from 10-d-old seedlings or inflorescences were treated with DNase I (Roche) and purified. Five micrograms of purified RNA was reverse-transcribed using ReverTraAiD Reverse Transcriptase (Thermo Fisher Scientific) with oligo-d(T) primers according to the manufacturer's instructions. RT–qPCR was carried out in triplicate using the iQ SYBRGreen Supermix (BioRad) on the BioRad CFX96 system. Primers used are listed in Supplementary Table 6.

Y2H assays. The Y2H assays were performed as described8 according to the Matchmaker Gold Y2H system user manual (Clontech). Constructs of the TRX-2 components were from the laboratory of Y. Cui (Agriculture and Agri-Food Canada, Laval, Ontario, Canada)9. The pCTD pGBK7–CDCK, pGBK7–NOT2B and pGBK7–SE were from the laboratory of Y. Qi (Tsinghua University, Beijing, China). Other full-length or truncated cDNAs were amplified and cloned into the pGBK7 or pGADT7 vectors (Clontech), and verified by sequencing. For the Y2H assays, the pGBK7 and pGADT7 plasmids with the genes to be tested were transformed into the Y2H Gold and Y187 yeast strains, respectively. The two yeast strains were mixed for mating, and the mating mixture was transferred to the DDO and QDO/X/A media for selection of diploids and those with reporter gene expression, respectively.

Western blot, coimmunoprecipitation and IP–MS analyses. Twelve-day-old seedlings or inflorescences were ground to a fine powder in liquid nitrogen and the powder was resuspended in IP buffer (50 mM Tris–HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 0.1% NP-40 and complete protease inhibitor (Roche)) and incubated for 30 min with gentle rotation at 4 °C. The protein suspensions were centrifuged at 16,000g for 10 min to remove debris. The supernatant was incubated with GTP-Trap or red fluorescent protein (RFP)-Trap beads (ChromoTek) for 2 h at 4 °C, and then the beads were washed with IP buffer five times and boiled in SDS sample buffer (50 mM Tris–HCl at pH 6.8, 2% SDS, 10% glycerol, 1% bromophenol blue and 1% 2-mercaptoethanol). The proteins were resolved by 8%–12% SDS–PAGE and detected with corresponding antibodies.

For IP–MS, the same immunoprecipitation procedure was performed and the proteins were resolved by 5% SDS–PAGE. After Coomassie staining, the entire lane was analysed by mass spectrometry as described10.

ChIP assays. Twelve-day-old Arabidopsis seedlings were cross-linked in buffer (1% formaldehyde) and crosslinking was stopped with 100 mM (final concentration) glycine, with both treatments carried out at room temperature under vacuum. The plant material was ground to a fine powder, which was resuspended in cold nuclei-isolation buffer (0.25 M sucrose, 15 mM Pipes pH 6.8, 6 mM MgCl2, 60 mM KCl, 15 mM NaCl, 1 mM CaCl2, 0.9% Triton X-100, 1 mM PMSF and 1% Protease inhibitor cocktail). The sample was centrifuged at 16,000g for 10 min at 4 °C. The supernatant was incubated with GTP-Trap or red fluorescent protein (RFP)-Trap beads (ChromoTek) for 2 h at 4 °C, and then the beads were washed with IP buffer five times and boiled in SDS sample buffer (50 mM Tris–HCl at pH 6.8, 2% SDS, 10% glycerol, 1% bromophenol blue and 1% 2-mercaptoethanol). The proteins were resolved by 8%–12% SDS–PAGE and detected with corresponding antibodies.

For IP–MS, the same immunoprecipitation procedure was performed and the proteins were resolved by 5% SDS–PAGE. After Coomassie staining, the entire lane was analysed by mass spectrometry as described10.

Nuclear–cytoplasmic fractionation. Twelve-day-old seedlings were ground to a fine powder in liquid nitrogen and the powder was resuspended in 2 ml/g•1 solution buffer (20 mM Tris–HCl pH 7.5, 10 mM MgCl2, 2.5 mM MgCl2, 2.5% glycerol, 250 mM sucrose, 5 mM DTT and protease inhibitor cocktail (Roche)). The suspension was filtered through two layers of Miracloth. The flow through was centrifuged at 1,500g for 10 min at 4 °C. The supernatant was centrifuged at 10,000g for 15 min at room temperature, and the supernatant was collected and crosslinked with 0.5% formaldehyde/1X PBS buffer under vacuum for 15 min. The inflorescence tissues were then transferred into the DDO and QDO/X/A media for selection of diploids and those with reporter gene expression, respectively.

ChIP assays. Twelve-day-old Arabidopsis seedlings were cross-linked in buffer (1% formaldehyde) and crosslinking was stopped with 100 mM (final concentration) glycine, with both treatments carried out at room temperature under vacuum. The plant material was ground to a fine powder, which was resuspended in cold nuclei-isolation buffer (0.25 M sucrose, 15 mM Pipes pH 6.8, 6 mM MgCl2, 60 mM KCl, 15 mM NaCl, 1 mM CaCl2, 0.9% Triton X-100, 1 mM PMSF and 1% Protease inhibitor cocktail). The sample was centrifuged at 16,000g for 45 min at 4 °C. The final nuclear pellet was used for RNA isolation using the Trizol method.

For the improved nucleocytoplasmic fractionation approach, 12-d-old seedlings were cross-linked in 0.5% formaldehyde/1X PBS buffer under vacuum for 15 min. The inflorescence tissues were then centrifuged at 10,000g for 10 min at 4 °C to a final concentration of 100 mM followed by incubation for 5 min under vacuum at room temperature. The plant material was washed in 1X PBS buffer and frozen in liquid nitrogen. The frozen tissue then underwent the same nucleocytoplasmic fractionation procedure as described above.

GUS staining. Inflorescences from Col-0 and thp1-5 harbouring a homozygous pMIR167a:GUS transgene were vacuum infiltrated with the GUS staining solution (1 mM EDTA, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 100 mM sodium phosphate, 1% Triton X-100, 1 mg/ml X-Gluc) for 10 min and then incubated at 37 °C for 2 h. Tissue clearing was performed with 70% ethanol for 2 d before imaging.

Immunofluorescence and confocal microscopy. For DCL1 immunofluorescence, inflorescence tissue was collected and fixed with 4% paraformaldehyde in 1X PBS for 20 min under vacuum. The tissue was quenched by 50 mM NH4Cl (in 1X PBS) and washed three times with 1X PBS. The tissue was chopped on a glass slide in nuclei-isolation buffer (500 mM sucrose, 10 mM KCl, 10 mM Tris–HCl, pH 9.0, 10 mM EDTA, 4 mM spermidine and 1 mM spermine) and nuclei were isolated by filtering the tissue suspension through a 100 μm nylon mesh and centrifuging the filtrate at 500g for 3 min. The nuclei were resuspended in 5% acrylamide-bisacrylamide (29:1) in nucleus-isolation buffer. A 2.5 μl aliquot of the suspension was spotted on a slide, covered with coverslip, and allowed to air dry at room temperature for 2 h. The coverslip was removed and a hybridization chamber was formed around the spot containing the nuclei. The entire immunofluorescence procedure was performed in the hybridization chamber: the samples were treated with 0.5% Triton X-100 (10 min), 3% CAS630 in 1X PBS (1 h), 1X PBS (5 min, 3 times), 1% BSA (30 min at 37°C), 1:100 primary antibody (1 h at 37 °C), 1X PBS (10 min, 6 times), 1:3,000 DAPI (20 min at room temperature) and 1X PBS (5 min, 2 times). The hybridization chamber was removed and the sample was mounted with Vectashield and sealed with nail polish. Images were captured with a Zeiss 880 fluorescence confocal microscope.

To generate DCL1 antibody, a recombinant protein of DCL1 (amino acids 1–250)–glutathione S-transferase (GST) was expressed, purified and used as the antigen to raise polyclonal antibodies in rabbits. The antisera was affinity-purified using a DCL1–GST–His-conjugated column. The specificity of the antibody was validated by immunoblotting with wild-type and dcl1-9 plants. DNA transfection experiments, root tips from 5-d-old seedlings were directly examined for fluorescent signals under a Zeiss 880 fluorescence confocal microscope.

RNA-seq library construction and data analysis. Three batches of 12-d-old Col-0, thp1-5 and suc3-1/suc3-1/suc3-1/suc3-1/suc3-1/suc3-5 seedlings were used for total RNA extraction. Polyadenylated RNA was isolated from total RNA using the Magnetic mRNA Isolation Kit (New England Biolabs). RNA-seq libraries were prepared using NEBNext mRNA Library Prep Reagent Set for Illumina (New England Biolabs) and sequenced on an Illumina HiSeq 2500 platform to generate high-quality single-end reads of 101 bp in length. Data analysis was performed with the pRNASeqTools pipeline (https://github.com/grubbybio/pRNASeqTools). The raw reads were aligned to the TAIR10 genome using HISAT2 (ref. 12). Transcript levels were measured in reads per million total read counts. Differentially expressed genes were identified using DESeq2 with fold change of 2 and P < 0.01 as the parameters13. The P-value of overlapped genes between two sets was calculated by SuperExactTest14. Gene Ontology analysis was performed with agriGO15.
References

1. Achkar, N. P., Cambiagi, D. A. & Manarella, P. A. miRNA biogenesis: a dynamic pathway. Trends Plant Sci. 21, 1034–1044 (2016).

2. Chen, X. Small RNAs and their roles in plant development. Annu. Rev. Cell Dev. Biol. 25, 21–44 (2009).

3. Bolognani, N. G. et al. Nucleo-cytoplasmic shuttling of ARGONAUTE1 prompts a revised model of the plant microRNA pathway. Mol. Cell 69, 709–719 (2018).

4. Yu, Y., Jia, T. & Chen, X. The 'how' and 'where' of plant microRNAs. N. Phytol. 216, 1002–1017 (2017).

5. Kim, Y. J. et al. The role of Mediator in small and long noncoding RNA production in Arabidopsis thaliana. EMBO J. 30, 814–822 (2011).

6. Fang, X., Cui, Y., Li, Y. & Qi, Y. Transcription and processing of primary microRNA transcripts are coupled by Elongator complex in Arabidopsis. Nat. Plants 1, 15075 (2015).

7. Wang, L. et al. NOT2 proteins promote Polymerase II-dependent transcription and interact with multiple microRNA biogenesis factors in Arabidopsis. Plant Cell 25, 715–727 (2013).

8. Li, S. et al. MAC3A and MAC3B, two core subunits of the MOS4-associated complex, positively influence miRNA biogenesis. Plant Cell 30, 481–494 (2018).

9. Jia, T. et al. The Arabidopsis MOS4-associated complex promotes microRNA biogenesis and precursor mRNA splicing. Plant Cell 29, 2626–2643 (2017).

10. Zhang, S., Liu, Y. & Yu, B. PRL1, an RNA-binding protein, positively regulates the accumulation of miRNAs and siRNAs in Arabidopsis. PLoS Genet. 10, e1004841 (2014).

11. Zhang, S., Xie, M., Ren, G. & Yu, B. CDC5, a DNA binding protein, positively regulates posttranscriptional processing and/or transcription of primary microRNA transcripts. Proc. Natl Acad. Sci. USA 110, 17588–17593 (2013).

12. 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).

13. Iki, T. et al. In vitro assembly of plant RNA-induced silencing complexes facilitated by the nuclear chaperone HSP90. Mol. Cell 39, 282–291 (2010).

14. Wang, W. et al. An importin beta protein negatively regulates microRNA with ARGONAUTE1 in Arabidopsis. Plant Cell 18, 1041–1053 (2006).

15. Wang, S. K. et al. The PROTEIN PHOSPHATASE4 complex promotes transcription and interact with multiple microRNA biogenesis factors in Arabidopsis. Mol. Cell 31, 486–501 (2019).

16. Song, L., Han, M.-H., Lesicka, J. & Fedoroff, N. Arabidopsis primary microRNA processing proteins HY1 and DCL1 define a nuclear body distinct from the Cajal body. Proc. Natl Acad. Sci. USA 104, 5437–5442 (2007).

17. Cho, S. K., Ben Chaabane, S., Shah, P., Poulson, C. P. & Yang, S. W. COP1 E3 ligase protects HY1 to retain microRNA biogenesis. Nat. Commun. 5, 5867 (2014).

18. Des Lauriers, S. D., Alvarez, A. A., Lacey, R. F., Binder, B. M. & Larsen, P. B. Dominant gain-of-function mutations in transmembrane domain III of ERS1 and ETR1 suggest a novel role for this domain in regulating the magnitude of ethylene response in Arabidopsis. N. Phytol. 208, 442–455 (2015).

19. Cabal, G. G. et al. SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. Nature 441, 770–773 (2006).

20. Voinnet, O. Diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. Annu. Rev. Plant Biol. 65, 473–503 (2014).

21. Ghioldi, M. & Zamore, P. D. Plant siRNAs: An expanding universe. Nat. Rev. Genet. 10, 94–108 (2009).

22. Du, T. T. & Zamore, P. D. microPreimer: The biogenesis and function of microRNA. Development 132, 4465–4465 (2005).

23. Nakagawa, T. et al. Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. J. Biosci. Bioeng. 104, 34–41 (2007).

24. Li, S. et al. MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. Cell 153, 562–574 (2013).

25. Li, H., Ruan, J. & Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. 18, 1851–1858 (2008).

26. Li, S. et al. Biogenesis of phased siRNAs on membrane-bound polycyomes in Arabidopsis. eLife 5, e22750 (2016).

27. Johnson, N. R., Yeoh, J. M., Coruh, C. & Axtell, M. J. Improved placement of dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).

28. Zhang, B., Wang, L., Zeng, L., Zhang, C. & Ma, H. Arabidopsis TOE proteins convey a photoperiodic signal to antagonize CONSTANS and regulate flowering time. Genes Dev. 29, 975–987 (2015).

29. Zeng, L. et al. Evolution and protein interactions of AP2 proteins in Brassicaceae: Evidence linking development and environmental responses. J. Integr. Plant Biol. 58, 549–563 (2016).
56. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. Nat. Methods 12, 357–360 (2015).

57. Wang, M., Zhao, Y. & Zhang, B. Efficient test and visualization of multi-set intersections. Sci. Rep. 5, 16923 (2015).

58. Du, Z., Zhou, X., Ling, Y., Zhang, Z. & Su, Z. agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res. 38, W64–W70 (2010).

Acknowledgements
We thank D. Weigel (Max Planck Institute for Developmental Biology, Tubingen, Germany) for sharing the pSUC2:amiR-SUL line; Y. Cui (Agriculture and Agri-Food Canada, London, Ontario, Canada) for sharing the rpl1-3 and sac3 mutants and TREX-2 Y2H plasmids; M. Chen (University of California, Riverside, CA, USA) for sharing the fluorescence protein-containing vectors and stimulating discussions; and B. Le, H. Hu and N. Flynn for comments on the manuscript. This work was funded by NIH GM129373 to X.C.

Author contributions
X.C. and B.Z. designed the research. B.Z. performed the majority of experiments. C.Y. and L.Z. conducted data analyses. Y.Z. performed nuclear AGO1 IP–MS. M.Z., J.H. and B.Z. performed genotyping for double mutant identification. B.Z. and X.C. 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/s41477-020-0726-z. Supplementary information is available for this paper at https://doi.org/10.1038/s41477-020-0726-z.

Correspondence and requests for materials should be addressed to X.C.

Peer review information Nature Plants thanks Hiro-oki Iwakawa, Pablo Manavella, Rodrigo Reis and the other, anonymous, reviewer(s) 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 | Small-RNA sequencing of thp1-5 amS. a, Heatmaps to show reproducibility among amS and thp1-5 amS biological replicates. The whole genome was divided into 100-bp bins, and small-RNA reads whose 5’ ends located in a bin were assigned to this bin. Color density indicates distance calculated by log-transformed normalized read counts assigned to each bin. b, Distribution of fold changes for all detected miRNAs (n = 232) and 21 nt (n=59742 windows of 100bp) and 24 nt (n=201374 windows of 100bp) small RNAs in thp1-5 amS relative to amS. The lower extreme, the lower hinge, the white dot, the upper hinge, and the upper extreme of the black box represent the minimum, the first quartile, the median, the third quartile, and the maximum of the data. The violin shape corresponds to the density of data. ***, P-value < 2.2e-16. P-values were determined by a paired two-sided Wilcoxon test.
Extended Data Fig. 2 | Defects of thp1 mutants. a–b, Flowers (a) and siliques (b) of amS and thp1-5 amS plants. Note that the shoot terminates in thp1-5 amS but not in wild type (amS). Scale bar in (a) = 1 mm. Scale bar in (b) = 10 mm. c, 21-day-old Col, thp1-1 and thp1-3 plants with differences in leaf shape. Scale bar = 10 mm. d, RNA gel blot analysis of 6 miRNAs in Col, thp1-1 and thp1-3. The miRNA signals were quantified using U6 as a loading control, and values were relative to Col. These experiments were repeated three times with similar results.
Extended Data Fig. 3 | Genetic interactions between mutants in THP1 and miRNA biogenesis genes. a, Morphological phenotypes of 25-day-old amS, thp1-5 amS, se-1 amS, hyl1-2 amS, hst-6 amS, hen1-8 amS, ago1-27 amS, and thp1-5 ago1-27 amS plants. Scale bar = 10 mm. All plants with the same genotype exhibit the same phenotype and one representative for each genotype is shown. b, Genetic segregation analyses. F2 plants of the indicated genotypes were selfed and ~48 F3 plants were genotyped for the thp1-5 mutation. The ratios of thp1-5/+ : +/+ are around 2:1, indicating that thp1-5 homozygosity leads to embryo lethality.
Extended Data Fig. 4| See next page for caption.
Extended Data Fig. 4 | Arabidopsis TREX-2 components and soc3 mutant alleles. a, Yeast two-hybrid assays to test interactions among TREX-2 components (THP1, SAC3A, SAC3B, DSSI-(I), DSSI-(V), CEN1, and CEN2). The empty vectors AD and BD were included as negative controls. After mating, diploid yeast cells containing both the bait and prey plasmids grew on SD-Trp/-Leu (DDO) medium, and diploid yeast cells in which the bait and prey proteins interacted grew into blue colonies on SD-Trp/-Leu/-His/-Ade+X-gal Aureobasidin A (QDO/X/A) medium. The experiment was repeated two times with similar results. b, Diagrams of SAC3A, B, and C genes showing the different mutant alleles. Triangles indicate positions of T-DNA insertions. Black arrows indicate primers used for RT-PCR. Lines below the gene models indicate conserved protein domains in homologous proteins. c, RT-PCR analyses of SAC3 gene expression in different soc3 mutants. The primers used are shown as black arrows in (b). The experiment was repeated two times with similar results.
Extended Data Fig. 5 | TREX-2 associates with the transcription machinery. a, Yeast two-hybrid assays to test interactions between TREX-2 components (THP1, SAC3A, SAC3B, DSS1-(I), DSS1-(V), and CEN1) and subunits of Mediator (MED20A, MED20B, MED20C and MED18). The empty vectors AD and BD were included as negative controls. After mating, diploid yeast cells containing both the bait and prey plasmids grew on SD-Trp/-Leu (DDO) medium, and diploid yeast cells in which the bait and prey proteins interacted grew into blue colonies on SD-Trp/-Leu/-His/-Ade + Xα-gal Aureobasidin A (QDO/X/A) medium. The experiment was repeated two times with similar results. b, Yeast two-hybrid assays to test interactions between TREX-2 components (THP1, SAC3A, SAC3B, DSS1-(I), DSS1-(V), and CEN2) and subunits of Elongator (ELP4 and ELP5) or subunits of MAC (PRL1 and MAC8). The empty vectors AD and BD were included as negative controls. After mating, diploid yeast cells containing both the bait and prey plasmids grew on SD-Trp/-Leu (DDO) medium, and diploid yeast cells in which the bait and prey proteins interacted grew into blue colonies on SD-Trp/-Leu/-His/-Ade + Xα-gal Aureobasidin A (QDO/X/A) medium. The experiment was repeated two times with similar results. c, BiFC analysis to test interactions between SAC3A and Pol II CTD (C-terminal domain). Scale bar = 10 µm. Three independent experiments yielded similar results. d, RT-qPCR to determine pri-miR-SUL and Bar transcript levels in amS and thp1-5 amS. IPP2 was used as an internal control. More than 30 plants from each genotype were used in the experiment. e, ChIP-qPCR analyses to determine the occupancy of four MIR loci by Pol II in 12-day-old seedlings of amS and thp1-5 amS. IgG amS was served as a negative control. More than 200 plants from each genotype were used in the experiment.
Extended Data Fig. 6 | See next page for caption.
Extended Data Fig. 6 | mRNA sequencing of wild type (Col), *thp1-5*, and *sac3b-1/sac3b-2 sac3c-1/sac3c-1*. a, Correlation analysis across WT, *thp1-5* and *sac3b-1/sac3b-2 sac3c-1/sac3c-1* samples. The three samples from each genotype are biological replicates. Color density indicates Euclidean distance calculated by log-transformed normalized read counts assigned to each gene. b, Venn diagrams showing the overlap in down-regulated (left) and up-regulated (right) genes between *thp1-5* and *sac3b-1/sac3b-2 sac3c-1/sac3c-1*. c, GO enrichment analysis of 852 commonly up-regulated genes in *thp1-5* and *sac3b-1/sac3b-2 sac3c-1/sac3c-1*. d, GO enrichment analysis of 819 commonly down-regulated genes in *thp1-5* and *sac3b-1/sac3b-2 sac3c-1/sac3c-1*. 
**Extended Data Fig. 7 | TREX-2 associates with the microprocessor.** **a**, Yeast two-hybrid assays to test interactions between TREX-2 components (THP1, SAC3A, SAC3B, DSS1-(I), DSS1-(V), and CEN2) and the microprocessor components DCL1 and HYL1. The empty vectors AD and BD were included as negative controls. After mating, diploid yeast cells containing both the bait and prey plasmids grew on SD-Trp/-Leu (DDO) medium, and diploid yeast cells in which the bait and prey proteins interacted were expected to grow into blue colonies on SD-Trp/-Leu/-His/-Ade + Xα-gal Aureobasidin A (QDO/X/A) medium. DCL1-N, amino acids 1 to 831 of DCL1; DCL1-C, amino acids of 832 to 1909 of DCL1. The experiment was repeated two times with similar results. **b**, BiFC assay to test interactions between SAC3A and SE. Scale bar = 10 µm. Two independent experiments yielded similar results. **c**, Determination of HYL1 phosphorylation status. The same blot as that of the upper panel in Fig. 4c with different exposure time. The different forms of HYL1 are as indicated (P⁺ = phosphorylated HYL1; P⁻ = non-phosphorylated HYL1). The experiment was repeated three times with similar results. **d**, Another replicate for Fig. 4c. **e**, A western blot to show the levels of HYL1-YFP and endogenous HYL1 in 3SS:HYL1-YFP and 3SS:HYL1-YFP thp1-5. RPN6 is a loading control. Two independent experiments yielded similar results.
Extended Data Fig. 8 | DCL1 D-body formation is compromised in thp1-5. a, A western blot to detect DCL1 with the anti-DCL1 antibody in WT and dcl1-9 (a truncation mutant). This shows that the antibody was able to specifically detect DCL1. The experiment was conducted one time. b, Detection of DCL1 D-body by immunofluorescence of isolated nuclei from inflorescence tissue of Col and thp1-5. DCL1 D-body signals are in green and DAPI stained nuclei are in blue. One representative image is shown for Col and thp1-5 each. Scale bar =1µm. The experiment was repeated independently two times with similar results. c, Quantification of DCL1 D-body numbers in Col (n=78) and thp1-5 (n=101) nuclei. The numbers 0–4 represent the numbers of D-bodies in each nucleus. d, ChIP-qPCR analyses to determine the occupancy of four MIR loci by DCL1 in 12-day-old seedlings of amS and thp1-5 amS. amS and thp1-5 amS were immunoprecipitated by DCL1 antibody. SN1 (located between At3g44000 and At3g44005 at nucleotides 15,805,617–15,805,773 of chromosome 3) was used as a negative control. IgG amS served as a negative control.
Extended Data Fig. 9 | The nucleo-cytoplasmic partitioning of miRNAs and AGO1 in thp1-5. a, Comparison of an improved nucleo-cytoplasmic fractionation method and the traditional method. The fractionated samples were subjected to protein gel blot analysis using anti-HYL1, anti-GAPDH and anti-H3 antibodies, respectively. T = total extract; C = cytoplasm; N = nucleus. HYL1 and H3 are nuclear proteins; GAPDH is a cytoplasmic protein. The experiment was repeated two times with similar results. b, Optimization of nucleo-cytoplasmic fractionation in terms of the duration of paraformaldehyde crosslinking (8 min, 15 min and 20 min). The fractionated samples were subjected to protein gel blot analysis using anti-AGO1, anti-cFBPase, anti-HYL1, anti-BIP and anti-H3 antibodies, respectively. T = total extract; C = cytoplasm; N = nucleus. H3 is a nuclear marker; cFBPase is a cytoplasmic marker. They were also used in the quantification of AGO1 levels (represented by the numbers) between T and N and between T and C, respectively. Three independent experiments gave similar results. c, Small RNA gel blot assays to determine the levels of various miRNAs in Col and thp1-5 following fractionation with the improved method. T = total extract; C = cytoplasm; N = nucleus. Signal intensity of T was arbitrarily set to 1.0; that of C and N was normalized to T against U6 and tRNA-Met, respectively, as nuclear and cytoplasmic RNA markers. The experiment was repeated two times with similar results. d, Cytoplasmic/nuclear ratios of various miRNAs as determined in (c).
Extended Data Fig. 10 | Characterization of nup1-cs lines. a, 45-day-old plants of wild type (Col) and three lines harboring the pNUP1:NUP1-mEGFP transgene (the three on the right). Two of the three are nup1-cs (nup1 co-suppression) lines. Scale bar = 10 mm. b, RT-PCR analysis of NUP1 transcripts in Col, pNUP1:NUP1-mEGFP and nup1-cs lines. Note that the signal represents transcripts from both the endogenous NUP1 gene and the pNUP1:NUP1-mEGFP transgene. The experiment was repeated two times with similar results. c, Protein gel blot analysis of the NUP1-mEGFP protein in Col, pNUP1:NUP1-mEGFP and nup1-cs lines. The experiment was conducted one time. d, RNA gel blot assays to determine the levels of miR166 and miR172 in Col and a nup1-cs-1 line. Inflorescences were used for RNA extraction. U6 served as an internal control. The experiment was repeated two times with similar results. e–f, Protein gel blot assays to determine the nucleo-cytoplasmic partitioning of AGO1 in Col and nup1-cs (e) and in Col, nup1-cs, and thp1-5 (f). Three independent experiments gave similar results. The fractions were subjected to protein gel blot analysis using anti-AGO1, anti-cFBPase, anti-H3 and anti-BIP antibodies, respectively. H3 is a nuclear marker; cFBPase is a cytoplasmic marker. They were also used in the quantification of AGO1 levels (represented by the numbers) between T and N and between T and C, respectively. BIP is an ER marker to indicate the level of ER contamination in the nuclear fractions.
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**

All Illumina sRNA-seq data in this study were collected by the manufacturer's software following the manufacturer's pipeline.

**Data analysis**

All sequencing data were analyzed by an in-house pipeline pRNASeqTools v0.6 (https://github.com/grubbybio/pRNASeqTools).

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 temporary small RNA sequencing datasets link: http://gofile.me/460c2/DaEDvRvwL

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  No sample-size calculations were performed. Sample size was determined to be adequate based on the magnitude and consistency of measurements.

Data exclusions  No data were excluded in this study.

Replication  The reproducibility of biological replicates were shown in Figure1d, Figure3a and 3c.

Randomization  All experiments were in bulk, using RNA extracted from biological replicates which were consisted of over 20 seedlings of Arabidopsis thaliana.

Blinding  Investigators were not blinded to Arabidopsis genotypes during experiments. Data reported are not subjective.

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**

| n/a | Involved in the study |
|-----|-----------------------|
|     | Antibodies           |
|     | Eukaryotic cell lines|
|     | Palaeontology        |
|     | Animals and other organisms|
|     | Human research participants|
|     | Clinical data        |

**Methods**

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

**Antibodies**

Antibodies used  antiAGO1 (agrisera,Cat. AS09 527), anti-GFP (Roche, Cat. 11814460001), anti-RFP (Chromotek, Cat. 6G6-100), anti-HYL1 (agrisera, Cat. AS06 136), anti-SE (agrisera, Cat. AS09 532A), anti-H3 (abcam, Cat. ab1791), anti-GAPDH (santa Cruz, Cat. sc-365062), anti-cFBPase (agrisera, Cat. AS04 043), GFP-trap (Chromotek, Cat. gtma-20), RFP-trap (Chromotek, rtma-20), anti-DCL1 was home-made and validated in Supplementary Figure 8a

Validation  anti-DCL1 was home-made and validated in Supplementary Figure 8a. All other antibodies used are commercial and were validated.

**Eukaryotic cell lines**

Policy information about  cell lines

| Cell line source(s) | State the source of each cell line used. |
|---------------------|----------------------------------------|
| Authentication      | Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated. |
| Mycoplasma contamination | 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. |
| Commonly misidentified lines (See  ICOLC register) | Name any commonly misidentified cell lines used in the study and provide a rationale for their use. |

**Palaeontology**

Specimen provenance  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).

Specimen deposition  Indicate where the specimens have been deposited to permit free access by other researchers.

Dating methods  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
Animals and other organisms

Policy information about studies involving animals: ARRIVE guidelines recommended for reporting animal research

| Laboratory animals | For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals. |
|--------------------|------------------------------------------------------------------------------------------------------------------|
| Wild animals       | 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. |
| Field-collected samples | 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. |
| Ethics oversight   | 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. |

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 | 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." |
| Recruitment               | 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. |
| Ethics oversight          | Identify the organization(s) that approved the study protocol. |

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

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 | For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data. |
| Files in database submission | Provide a list of all files available in the database submission. |
| Genome browser session (e.g. UCSC) | 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. |

Methodology

| Replicates | Describe the experimental replicates, specifying number, type and replicate agreement. |
| Sequencing depth | Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of... |
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
Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

Instrument
Identify the instrument used for data collection, specifying make and model number.

Software
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.

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.