RBV, an evolutionarily conserved WD40 repeat protein, promotes microRNA biogenesis and loading into ARGONAUTE1 in *Arabidopsis*

Chao Liang
Shenzhen University
Qiang Cai
Shenzhen University https://orcid.org/0000-0001-5871-2519
Shaofang Li
State Key Laboratory of Plant Physiology and Biochemistry, College of Biological Sciences, China Agricultural University, Beijing 100193, China https://orcid.org/0000-0002-1903-023X
Chenjiang You
University of California, Riverside https://orcid.org/0000-0002-9935-1892
Chi Xu
Shenzhen University
Lei Gao
Shenzhen University https://orcid.org/0000-0002-6488-0292
Fei Wang
Shenzhen University
Dechang Cao
Shenzhen University https://orcid.org/0000-0002-0397-4316
Beixin Mo
Shenzhen University https://orcid.org/0000-0001-7279-9788
Xuemei Chen (xuemei.chen@ucr.edu)
University of California, Riverside https://orcid.org/0000-0002-5209-1157

Article

Keywords: microRNAs (miRNAs), biogenesis, splicing

Posted Date: June 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-536724/v1

Version of Record: A version of this preprint was published at Nature Communications on March 8th, 2022. See the published version at https://doi.org/10.1038/s41467-022-28872-x.

Abstract

MicroRNAs (miRNAs) play crucial roles in gene expression regulation through RNA cleavage or translation repression. Although miRNA biogenesis has been extensively studied, new factors involved in miRNA biogenesis are still being found from various genetic screens, suggesting that knowledge of miRNA biogenesis or its regulation is still incomplete. Here, we report the identification of a previously uncharacterized and evolutionarily conserved WD40 domain protein as a player in miRNA biogenesis in *Arabidopsis thaliana*. A mutation in the R*EDUCTION IN* B*LEACHED* V*EIN AREA* (*RBV*) gene encoding a WD40 domain protein led to the suppression of leaf bleaching caused by an artificial miRNA; the mutation also led to a global reduction in the accumulation of endogenous miRNAs. RBV may act at multiple steps in miRNA biogenesis. RBV promotes the transcription of *MIR* genes into pri-miRNAs by enhancing the occupancy of RNA polymerase II (Pol II) at *MIR* gene promoters. The nuclear protein RBV may also participate in pri-miRNA processing, as the mutation of *RBV* leads to the reduction of dicing body number. Moreover, mutation of *RBV* leads to a defect of miRNA loading into AGO1.RBV function is required in messenger RNA splicing, as RNA-seq uncovered a global intron retention defect in the mutant. Thus, this previously uncharacterized, evolutionarily conserved, nuclear WD40 domain protein acts in miRNA biogenesis and RNA splicing.

Introduction

MicroRNAs (miRNAs), 20 to 24 nucleotides (nt) in length, are one class of endogenous noncoding small RNAs in eukaryotes. miRNAs are players in gene regulatory networks involved in many biological processes such as development, metabolism and immunity in plants and animals[1]. Target gene expression is regulated by miRNAs post-transcriptionally through RNA cleavage or translation repression[2] .

Biogenesis of plant miRNAs entails a series of steps. Genes encoding miRNAs (*MIR*) are transcribed into pri-miRNAs that form imperfect stem-loop structures by DNA-dependent RNA polymerase II (Pol II)[3, 4]. *MIR* transcription is facilitated by Mediator[5], NEGATIVE ON TATA LESS2 (NOT2)[6], CELL DIVISION CYCLE5 (CDC5)[7], Elongator[8], and the TREX-2 component THP1[9]. Pri-miRNAs are processed into pre-miRNAs by DICER–LIKE1 (DCL1)[1, 10, 11, 12]. In humans, pri-miRNAs are thought to be co-transcriptionally processed, as they are associated with chromatin during transcription and processed at sites of transcription before splicing[13, 14]. It is proposed that the retention of pri-miRNAs at transcription sites may enhance processing[14]. In *Arabidopsis*, NOT2, CDC5 and Elongator interact with both Pol II and DCL1[6, 7, 8], suggesting that pri-miRNA transcription and processing may be also coordinated in plants.

The efficient processing of pri-miRNAs requires the double-stranded RNA binding protein HYPONASTIC LEAVES1 (HYL1)[15, 16] and the zinc finger protein SERRATE (SE)[17, 18], which form the microprocessor complex with DCL1[19, 20]. The three proteins as well as pri-miRNAs are also found in nuclear foci called Dicing-bodies (D-bodies)[21]. Pre-miRNAs are processed to miRNA/miRNA* duplexes, which undergo 3’

methylation by the methyltransferase HEN1 to maintain miRNA stability[22]. Finally, the mature miRNA strands are loaded into ARGONAUTE1 (AGO1) to form RNA Induced Silencing Complexes (RISCs), which are active forms of miRNAs[23, 24, 25]. In Arabidopsis, size exclusion chromatography showed that the molecular weight of miRISCs is similar to that of AGO1, suggesting that miRISCs are bi-molecular AGO1-miRNA complexes, although larger complexes containing AGO1 can also be detected[23, 26, 27].

Here, we isolated a mutant with global defects in miRNA biogenesis in *Arabidopsis*. The mutation is in a previously uncharacterized gene encoding a protein containing seven WD40 repeats, which we designate as *RBV*. *RBV* promotes the transcription of *MIR* genes, as loss of function of *RBV* reduces *MIR* promoter activity and the occupancy of Pol II at *MIR* promoters. In addition, *RBV* promotes the localization of HYL1 in D-bodies. On the basis of these results, we propose that RBV may act to coordinate *MIR* transcription and pri-miRNA processing in plant miRNA biogenesis. Moreover, the association of miRNAs with AGO1 was drastically decreased and AGO1 resided in complexes larger than miRISCs in the *rbv-1* mutant, suggesting that RBV1 promotes the loading of miRNAs into AGO1. *RBV1* also has a global role in pre-mRNA splicing, affecting a set of short introns.

Results

Isolation of a mutant with defects in miRNA biogenesis

We performed an ethylmethane sulfonate mutagenesis screen for *Arabidopsis* mutants in miRNA biogenesis, utilizing the vein-centered leaf beaching phenotype caused by the phloem-specific expression of an artificial miRNA (amiR-SUL) targeting *SULFUR* (*SUL*) as a visible marker for miRNA activity[28]. A series of mutants with reduced leaf bleaching was isolated and found to be in genes with known roles in miRNA biogenesis/activity, such as *hyl1-11*, *dcl1-30*, *hen1-11* and *ago1-25*, suggesting that the genetic screen was effective (Supplementary Fig. 1a and b). The *ago1-25* allele isolated in our study was thus named because it harbored the same mutation as the one previously reported [29]. In addition, we isolated a new suppressor mutant with reduced leaf bleaching (Fig. 1a and b); the mutation was designated *rbv-1* as it was later shown to be in a previously uncharacterized gene that we named *REDUCTION IN BLEACHED VEIN AREA* (*RBV*). This mutant exhibited pleiotropic developmental phenotypes, such as reduced root length, smaller plant size, narrow leaves, short stature, increased branching and reduced fertility (Fig. 1c and Supplementary Fig. 2). Northern blot analyses showed that both amiR-SUL and endogenous miRNAs (miR156, miR159, miR164, miR165, miR167, miR319 and miR390) were moderately reduced in abundance in 14-day-old *amiR-SUL rbv-1* seedlings as compared to *amiR-SUL* seedings (Fig. 1d). We also preformed small RNA sequencing with 14-day-old *amiR-SUL* and *amiRSUL rbv-1* seedlings. Clustering analysis showed that the three biological replicates for each genotype were highly reproducible (Supplementary Fig. 3a). Reads corresponding to miRNAs were normalized against total mapped reads and miRNAs with RPM (reads per million mapped reads)>10 in either genotype (average of three replicates) were included in our analyses (Supplementary Data 1). An overall reduction in miRNA abundance in the mutant relative to the *amiR-SUL* line was observed (Fig. 1e and Supplementary Fig. 3b),

while there were some miRNAs such as miR845a and miR845b showing increased accumulation in the *amiRSUL rbv-1* mutant. We also performed RT-qPCR to examine the expression of some miRNAs target genes. The transcript levels of *SPL3* and *SPL10* (targets of miR156), *PHB*, *REV* and *PHV* (targets of miR165/6), *MYB33* and *MYB65* (targets of miR159), *ARF8* (a target of miR167), and *CUC2* (a target of miR164) were de-repressed in the *amiR-SUL rbv-1* mutant (Fig. 1f).

RBV encodes an evolutionarily conserved WD40 domain protein

The *amiR-SUL rbv-1* mutant was backcrossed with the parental *amiR-SUL* line. In a total of 614 F2 plants, 141 (23%) exhibited the mutant phenotypes, which is consistent with the phenotype being caused by a single, nuclear, and recessive mutation (c^2 =1.257; P =0.244; (Supplementary Table 1). In order to identify the causal mutation in *amiR-SUL rbv-1*, pooled DNA from mutant plants in the F2 of the *amiR-SUL rbv-1* x *amiR-SUL* cross was used for whole-genome re-sequencing. The results revealed that the *amiR-SUL rbv-1* phenotype was linked to a single nucleotide change (G-to-A) in the first exon of AT5G64730, causing the change of the encoded amino acid from glycine to glutamic acid (see Methods; Fig. 2a).

In order to confirm that *RBV* is indeed AT5G64730, a construct of *RBV* (AT5G64730)-eYFP driven by its native promoter was generated and introduced into *amiR-SUL rbv-1*. Phenotypes of the transgenic plants showed that *pRBV:RBV-eYFP* fully rescued the morphological defects of *amiR-SUL rbv-1* (Fig. 2c). Furthermore, the transgene restored amiR-SUL accumulation as well as leaf bleaching in *amiR-SUL rbv-1* (Fig. 2d). While the expression of the amiR-SUL target gene, *SUL*, was increased in the mutant at the RNA and protein levels, the transgene also restored *SUL* RNA (Fig. 2e) and SUL protein (Fig. 2f) in *amiR-SUL rbv-1* to wild-type levels. Moreover, the transgene rescued the defects in miR159 and miR319 accumulation in the *amiR-SUL rbv-1* mutant (Fig. 2d). Therefore, the miRNA biogenesis and morphological defects of *amiR-SUL rbv-1* were attributable to the mutation in *RBV* (AT5G64730).

RBV encodes a previously uncharacterized protein with seven WD40 repeats (Fig. 2a). In plants, WD40 repeat proteins are numerous, interact with diverse proteins, and act in a variety of biological processes, such as plant development and immunity[30, 31]. To investigate whether RBV is an evolutionarily conserved protein, a phylogenetic tree including homologs of RBV in various plants was generated (Supplementary Data 2); Supplementary Fig. 4a). According to the phylogenetic analysis, RBV orthologs can be found in plants ranging from single-cell green algae to core eudicots and grasses, and there is no close paralog of RBV in the latest common ancestor of land plants. In most eudicots that have undergone gamma whole genome duplication (WGD) and Brassicaceae that has also undergone beta and alpha WGDs, *RBV* remained as a single copy. Only in species that have undergone recent, specific WGDs, such as apple, soybean, and maize, there are multiple copies of *RBV* (Fig. 2b and Supplementary Fig. 4). This result indicated that RBV was functionally conserved and extra copies might be deleterious.

We also obtained two lines with T-DNA insertions in or near *RBV* (Salk_126634 and Salk_075672) (Supplementary Fig. 5a and 5c). Interestingly, the phenotypes of the T-DNA mutants were the same as Col (Supplementary Fig. 5b). Next, we examined *RBV* transcripts in the two T-DNA insertion lines. RT-PCR

showed that transcripts corresponding to the full-length coding region of *RBV* were present in the two mutants (Supplementary Fig. 5d). In addition, real-time RT-PCR showed that the levels of *RBV* transcripts were only mildly reduced in the two mutants (Supplementary Fig. 5e**)**, consistent with the lack of morphological phenotypes.

In order to remove the *amiR-SUL* transgene background, we crossed the *amiR-SUL rbv-1* mutant with wild-type (Col) plants. In the F2 population, we identified homozygous *rbv-1* plants without the *amiR-SUL* transgene through genotyping; these plants showed the same pleiotropic phenotypes as *amiR-SUL rbv-1* (Fig. 3a). The morphological phenotypes of *rbv-1* were completely rescued by the *pRBV:RBV-eYFP* transgene (Fig. 3a). We further confirmed that the *rbv-1* mutant without the *amiR-SUL* transgene had defects in miRNA accumulation. RNA gel blots showed that the levels of several endogenous miRNAs, such as miR156, miR159, miR164 and miR167, were mildly reduced in *rbv-1* as compared to wild-type Col plants. Other miRNAs, such as miR319 and miR398, were unaffected (Supplementary Fig. 6a). To rule out the possibility that RBV indirectly regulated miRNA accumulation through affecting the expression of the miRNA biogenesis machinery, we examined the expression of the known genes involved in miRNA biogenesis, and no significant changes were observed in *rbv-1* (Supplementary Fig. 6b and 6c).

RBV promotes the transcription of *MIR* genes

We next investigated how *RBV* promotes miRNA biogenesis. *MIR* gene transcription to produce pri-miRNAs is the first step in miRNA biogenesis. RT-qPCR showed that the levels of pri-miRNAs were reduced by ~30-50% in the *rbv-1* mutant (Fig. 3b). For the pri-miRNAs, 298 pri-miRNAs annotated in Araport11 were examined by RNA-seq. In total, only 41 pri-miRNAs gave reads in any sample (Supplementary Fig. 7). In the RNA-seq data, we only detected three (pri-miR159b, pri-miR166a and pri-miR319b) of the seven pri-miRNAs that we examined by qRT-PCR in Fig. 3b. All three pri-miRNAs showed reduced abundance in the *rbv-1* mutant, which is consistent with results from qRT-PCR. The reduction in pri-miRNA abundance could be attributed to impaired *MIR* genes transcription, reduced stability of pri-miRNAs or enhanced pri-miRNA processing. To determine whether transcription was affected in the *rbv-1* mutant, we crossed *rbv-1* with a GUS reporter line (*pMIR167a:GUS*) under the control of the *MIR167a* promoter and obtained *pMIR167a:GUS rbv-1* (with both the transgene and the mutation being homozygous). GUS activity was visibly lower in *pMIR167a:GUS rbv-1* than in *pMIR167a:GUS* as revealed by GUS staining (Fig. 3c). RT-qPCR analysis confirmed that the *rbv-1* mutant had lower GUS transcript levels (Fig. 3d). Thus, reduced transcription of *MIR* genes could be one of the problems in miRNA biogenesis in *rbv-1*.

To further confirm a positive role of *RBV* in *MIR* gene transcription, the occupancy of Pol II at *MIR* loci was determined by chromatin immunoprecipitation (ChIP) with an antibody against the Pol II C-terminal repeats. ChIP without an antibody served as the negative control. *MIR166a*, *MIR167a* and *MIR171a* promoter regions, as well as C1 (a region between the genes AT2G17460 and AT2G17470 known to not engage Pol II[7], were examined by RT-qPCR after ChIP. The *MIR166a* and *MIR167a* promoter regions were enriched in the immunoprecipitates in both *rbv-1* and Col relative to C1 (Fig. 3e). Pol II occupancy at *MIR166a* and *MIR167a* promoter regions was reduced in the *rbv-1* mutant relative to Col, while the signal

on *MIR171a* was too low to be calculated (Fig. 3e). Thus, *RBV* facilitates the recruitment of Pol II to *MIR* loci.

RBV is localized in the nucleoplasm and required for the proper localization of HYL1 in D-bodies

We studied the expression of *RBV* in various tissues and the subcellular localization of the protein. RT-PCR analyses of RNAs from seedlings, roots, cauline leaves, rosette leaves, stems and inflorescences showed ubiquitous *RBV* expression (Supplementary Fig.8). *RBV:RBV-eYFP* transgenic plants exhibited YFP fluorescence in the nucleoplasm but not the nucleolus. SE is a component of the microprocessor in pri-miRNA processing. An mRuby3-tagged SE protein driven by the *SE* promoter[9] co-localized with RBV-eYFP in the nucleoplasm (Fig. 4a).

HYL1 and DCL1 form nuclear foci known as Dicing bodies (D-bodies), which are sites of pri-miRNA processing[21, 32]. To determine the status of D-bodies in the *rbv-1* mutant, we crossed a *HYL1-YFP* transgene[21] into the mutant and obtained plants homozygous for both the transgene and the *rbv-1* mutation. D-body numbers were determined in 420 and 537 root nuclei of wild-type and *rbv-1* plants, respectively. The number of HYL1-YFP D-bodies was significantly decreased in *rbv-1* (Fig. 4b and 4c), suggesting that mutation of RBV leads to a defect in D-body formation or the localization of HYL1 to D-bodies.

Mutation of *RBV* leads to a defect of miRNA loading into AGO1

miRNAs are loaded into AGO1 to form miRISCs that execute the silencing of target RNAs[23, 25, 33]. To determine whether *RBV* affects the formation of RISCs, we performed AGO1 IP followed by sRNA-seq with both input and IP samples from wild-type and *rbv-1* seedlings. Three replicates were performed and showed high reproducibility (Supplementary Fig. 9). From input samples, many miRNAs showed small but statistically significant reduction in abundance in the *rbv-1* mutant (Supplementary Data 3). A few miRNAs, such as miR845a, miR845b and miR843, were elevated in abundance in the mutant (Fig. 5a; Supplementary Data 3). The loading status of each miRNA was expressed as the ratio of miRNA abundance between AGO1 IP and input and differences between wild type and *rbv-1* were evaluated by student's t-test. A global reduction in the AGO1 loading of miRNAs was observed in the *rbv-1* mutant (Fig. 5b; Supplementary Data 4). Notably, miR845a and miR845b, which exhibited elevated levels, were less associated with AGO1 in the mutant, indicating that RBV is crucial for miRISC formation. RNA gel blots were also performed to validate the sRNA-seq results. miR159 and miR166 levels were lower in input and further reduced in AGO1 IP in the mutant (Fig. 5c). miR845a levels were strongly increased in the *rbv-1* input sample as compared with Col input but greatly reduced in the *rbv-1* AGO1 IP sample as compared with Col AGO1 IP (Fig. 5c; Supplementary Data 4). Thus, the *rbv-1* mutant exhibits a global miRNA loading defect.

To determine how RBV might promote the loading of miRNAs into AGO1, we first examined whether RBV interacts with AGO1. Co-IP was performed with *pRBV:RBV-eYFP* plants using anti-GFP and anti-AGO1 antibodies, but no interaction between RBV and AGO1 was detected. We next sought to determine

whether the bi-molecular feature of miRISCs is affected in the mutant. Protein extracts from *pRBV:RBV-eYFP rbv-1* and *rbv-1* plants were subjected to gel filtration followed by western blotting to detect AGO1 and northern blotting to detect miR159. Among the 44 fractions, AGO1 was distributed in both high molecular weight (HMW) (fractions 9-11) and low molecular weight (LMW) (fractions 16-17) complexes in *pRBV:RBV-eYFP rbv-1*, consistent with previous findings[27]. Notably, the LMW complexes, which corresponded to monomeric AGO1 in size, also showed peak levels of miR159, suggesting that they represent miRISCs with one AGO1 protein and one miRNA. However, such miRISCs were greatly reduced in *rbv-1*, with AGO1 being in HMW complexes instead (Fig. 5d). The distribution of miR159 also shifted towards HMW complexes in *rbv-1* (Fig. 5d). Western blot analyses with anti-GFP antibody showed that RBV-eYFP was present in fractions 8-16 with estimated molecular weights much higher than that of an RBV-eYFP monomer, suggesting that RBV itself also resides in protein complexes (Supplementary Fig. 10). Therefore, RBV promotes the formation of miRISCs that contain only AGO1 and miRNAs. The HMW AGO1 complexes may represent intermediates in RISC formation.

AGO1 is known to associate with trans-acting small interfering RNAs (ta-siRNAs) in addition to miRNAs[34, 35]. Contrary to miRNAs, which are loaded into AGO1 in the nucleus, ta-siRNAs are loaded into AGO1 in the cytoplasm[24]. We investigated whether RBV affects the ta-siRNA-AGO1 association. We quantified 21-nt siRNAs that mapped to 100-bp windows that overlapped with *TAS1A*, *TAS1B*, *TAS1C* and *TAS2* loci. The ta-siRNAs were not significantly altered in *rbv-1* input as compared to Col input in each 100-bp window (Supplementary Fig. 11 a, c, e and g). We then analyzed their levels in AGO-IP vs. input, and found that the loading of ta-siRNAs into AGO1 was largely unaffected. Only ta-siRNAs in window 5 of *TAS1A*, window 3 of *TAS2B* and window 3 of *TAS1C* were affected (Supplementary Fig.11 b, d, f and h). The levels of miR173, the trigger of ta-siRNA biogenesis from *TAS1* and *TAS2* loci, were not significantly different between Col and the *rbv-1* mutant (Supplementary Fig. 11 i). The fact that ta-siRNA loading was largely unaffected in the mutant is consistent with RBV being a nuclear protein.

RBV is required for the splicing of short introns in certain pre-mRNAs

In order to determine whether *RBV* affects the expression of protein-coding genes, we performed RNA-seq with 14-day-old seedlings of Col and *rbv-1* in triplicates. The three biological replicates for each genotype were highly reproducible (Supplementary Fig. 12).

Differentially expressed genes (DEGs) were identified between Col and mutant samples with FPKM >1, fold-change >2 and FDR<0.05 as the cutoff. In total, we identified 632 up-regulated (hyper-DEGs) and 363 down-regulated (hypo-DEGs) genes, respectively (Supplementary Data 5 and 6, Supplementary Fig. 13a). Gene Ontology (GO) analyses showed that both the hyper-DEGs and hypo-DEGs were enriched in genes with roles in responses to stimuli (Supplementary Fig. 13b and 13c). We also examined the global transcript levels of miRNA targets from the RNA-seq data. Although several examined miRNA targets were found to be derepressed in the mutant by RT-qPCR, a global trend of increased levels of miRNA target transcripts was not observed (Supplementary Fig. 13d, Supplementary Data 7). It is possible that the *RBV* mutation leads to a defect in the transcriptional regulation of these genes as well.

It has been reported that MOS4-Associated Complex (MAC) components MAC3A, MAC3B, MAC7, and two WD40 repeat proteins (PRL1 and PRL2) affect both miRNA biogenesis and the splicing of protein-coding transcripts[36, 37, 38]. This prompted us to examine whether the *rbv-1* mutant had splicing defects using the RNA-seq data (see Methods). In total, 474 Genes were found to have intron retention defects in the *rbv-1* mutant compared to Col, with a total of 511 intron retention events (Supplementary Data 8). Two examples (*At4g15790* and *At1g03280*) are presented in Fig. 6a. PI (percent of intron reads) was increased in the *rbv-1* mutant (Fig. 6b). No differential expression of the genes with intron retention was observed between the *rbv-1* mutant and wild type (Supplementary Fig. 14a). Next, we examined whether the genes with intron retention or the retained introns in the *rbv-1* mutant had any common features. Intriguingly, the retained introns in the *rbv-1* mutant seemed to be shorter as compared to the overall length distribution of introns (P= 1.984083e-22) (Fig. 6c). Besides, genes affected in *rbv-1* tended to have more introns than all genes (P=2.823e-05) (Fig. 6d). Intron retention defects are also found in *prl1 prl2* and *mac3a mac3b* mutants[36]. We compared the retained introns in *rbv-1*, *prl1 prl2* and *mac3a mac3b* and found little overlap between the introns affected in *rbv-1* with those affected in either *prl1 prl2* or *mac3a mac3b* (Fig. 6d), suggesting that *RBV* affects different introns from the MAC components.

Many *MIR* genes have introns[39, 40, 41]. The RNA-seq did not detect intron retention events in pri-miRNAs, but the low abundance of pri-miRNAs could have precluded the detection of intron retention events. We performed RT-PCR to test the splicing of introns from three miRNA precursors (pri-miR163, pri-miR156 and pri-miR168) with intron-flanking primers. Genomic DNA was amplified with the same primers to indicate the size of the intron-containing fragments. No defects in pri-miRNA splicing were observed in the *rbv-1* mutant (Supplementary Fig. 14b).

Discussion

RBV is involved in *MIR* gene transcription, pri-miRNA processing and miRISC formation

RBV, a WD40 protein, is an evolutionarily conserved protein in plants. However, it has not been studied from any species. In this study, we show that a recessive mutation in *RBV* reduces the levels of many endogenous miRNAs, indicating that mutation of RBV leads to a defect in miRNA biogenesis. Consistent with the known roles of plant miRNAs in various developmental processes, the *rbv-1* mutant exhibits strong developmental defects. How does *RBV* promote miRNA biogenesis? Our findings suggest a role of *RBV* in promoting Pol II transcription of *MIR* genes, which is supported by the reduced levels of pri-miRNAs, compromised *MIR167a* promoter activity, and decreased Pol II occupancy at *MIR* genes in the *rbv-1* mutant. In Arabidopsis, a series of proteins have been found to promote *MIR* transcription and/or pri-miRNA processing, and can be grouped into two main classes[1]. One group contains CBP80 and CBP20[42], STA1[43], SICKLE[40], TOUGH[44], PINP1[45], THO1 and THO2[46, 47], and MOS2[48]. A common feature is that mutants in these genes show reduced levels of miRNAs and increased abundance of pri-miRNAs. A second group of proteins acts in a different manner. In loss-of-function mutants in the genes in this group, the abundances of both pri-miRNAs and mature miRNAs are reduced. Proteins in this group

include DAWDLE[49], CDC5[7], NOT2[6], Elongator[8], PRL1[50], MAC7[36], PP4[38] and THP1[9]. Besides, mutants in *PRL1*, *CDC5*, *MAC7*, *PP4* and *THP1* show a reduced number of HYL1 D-bodies[7, 9, 36, 38, 50]. A number of proteins in this group interact with DCL1, HYL1, or SE and thus are thought to bridge *MIR* gene transcription and pri-miRNA processing[6, 7, 8, 9, 38, 49]. Besides the two main classes, the third group, which includes CHR2, a partner of SE, promotes the transcription of *MIR* genes but represses miRNA accumulation by inhibiting pri-miRNA processing[51]. Our studies show that RBV belongs to the second group of proteins that promotes *MIR* gene transcription and possibly pri-miRNA processing.

RBV also differs from the second group of proteins in that it has a clear role in miRISC formation. In the *rbv-1* mutant, both AGO1 and miR159 shift into complexes with higher molecular weights. RBV itself is also found in high molecular weight complexes. We suspect that the high molecular weight complexes containing AGO1 and miR159 represent intermediates in miRISC formation, and that RBV helps the dissociation of AGO1-miRNA from other proteins to form active miRISCs. Heat Shock Protein 90 (HSP90) is required for sRNA loading into AGO1 in tobacco lysates[52] and for RISC formation in Drosophila and humans[53]. HSP90 proteins are involved in RNA silencing in animals[54] and plants[55, 56]. Consistently, HSP90.4 was found in GFP-AGO1 IP spectrometry experiments[24]. Another protein that has recently been shown to play a role in miRISC formation is CARP9, which interacts with both HYL1 and AGO1 in the nucleus[57]. The relationship between RBV, HSP90 and CARP9 is worth investigating in the future.

In summary, RBV, as a nuclear WD40 protein, is involved in multiple steps in miRNA biogenesis, including *MIR* transcription, pri-miRNA processing and AGO1 loading (Fig.7).

RBV is also required for the splicing of certain mRNAs

Proteins such as MAC7, PP4, MAC3A/3B, and PRL1 act not only in miRNA biogenesis but also in pre-mRNA splicing in *Arabidopsis*[36, 38]. We found that RBV also plays a role in splicing. Similar to mutants in *MAC7*, *PP4*, *MAC3A/3B*, and *PRL1*, no significant correlations between intron retention and changes in gene expression were found in the *rbv-1* mutant. Introns retained in the *rbv-1* mutant had minimal overlap with those in the other mutants, suggesting the presence of different categories of introns, whose splicing requires distinct factors. *RBV* acts on genes with more introns and introns that are shorter in length. Furthermore, intron retention was not found for intron-containing pri-miRNAs in *rbv-1*, suggesting that mutation of *RBV* leads to a defect in miRNA biogenesis and pre-mRNA splicing independently. It is possible that *RBV* function is required for yet other aspects of nuclear RNA metabolism.

Methods

Plant materials and growth conditions

The transgenic *Arabidopsis thaliana* (Col ecotype) line *pSUC2:amiR-SUL* expressing amiR-SUL under the *SUCROSE-PROTON SYMPORTER 2* (*SUC2*) promoter was a kind gift from Dr. Detlef Weigel[28]. Transgenic lines harboring *pSE:SE-mRuby3*, *pMIR167a:GUS*, and *pHYL1:HYL1-YFP*[5, 9, 21, 50] were used in this study.

All seeds were sown on 1/2 Murashige and Skoog Basal Medium (Sigma-Aldrich, M5519) plus 1% Sucrose and 0.8% Agar. All plants were grown at 22°C under 16 h light and 8 h dark cycles.

Mutagenesis and mapping

EMS mutagenesis was performed as described[9]. A mutant, *amiR-SUL rbv-1*, with reduced leaf bleaching, was isolated and then backcrossed with the parental line *pSUC2:amiR-SUL*. In the F2 generation, ~100 plants with the *amiR-SUL rbv-1* phenotype were identified. Genomic DNA was extracted by the CTAB method [58] from individual plants and equal amount of DNA was pooled for genomic DNA library construction. The library was paired-end (PE151bp) sequenced on the Illumina platform HiSeq4000 at 50x coverage at BGI-Shenzhen, China. Focusing only on G-to-A mutations in coding regions, a mutation in At5g64730 was identified in *amiR-SUL rbv-1*. A Derived Cleaved Amplified Polymorphic Sequences (dCAPS) marker was designed to genotype this mutation (Supplementary Table 2). The PCR products from wild type can be digested by *Nco*I, whereas those from *amiR-SUL rbv-1* could not. Genotyping the ~100 individuals showed that this mutation was linked to the plant phenotype.

DNA constructs and complementation

The genomic region of *RBV* (At5g64730) including ~1 kb promoter was amplified with the primers proRBV-F and RBV-R (Supplementary Table 2) and cloned into the pTSK108 vector. The clone was sequenced to ensure the absence of mutations and was recombined with the pGWB640 gateway vector[59] to generate the *pRBV:RBV-eYFP* construct via LR reaction. The *pRBV:RBV-eYFP* plasmid was transformed into the *rbv-1* mutant in both *amiR-SUL* and Col backgrounds through the *Agrobacterium tumefaciens*-mediated floral dip method as described[60]. The T1 transgenic plants were selected by BASTA resistance.

Fluorescence microscopy

For the visualization of RBV-eYFP and HYL1-YFP in transgenic plants, roots of 10-day-old seedlings were observed under a Zeiss LSM 5 Pascal inverted confocal microscope (excitation, 488/512 nm/nm; emission, 520/50 nm/nm). For SE-mRuby3, roots of 10-day-old *pSE:SE-mRuby3* seedlings were observed under a LeicaSP8 confocal microscope (excitation, 514 nm; emission, 540-600 nm).

Small RNA sequencing and data analysis

Total RNAs were extracted from 14-day-old seedlings using TRI reagent (TR118, Molecular Research Centre). Small RNAs in the size range of 15-40 nt were isolated from 30μg total RNAs by resolving total RNAs in denaturing polyacrylamide gels, excising gel pieces containing 15-40 nt RNAs and elution of the small RNAs according to Liu et al[61]. Small RNA libraries were prepared with NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs, E7300), and then sequenced using an Illumina HiSeq2500 platform at BerryGenomics, China. Data analysis was performed with the pRNASeqTools pipeline (https://github.com/grubbybio/pRNASeqTools). The raw reads (SE50) were trimmed using Perl scripts to remove adapters (adaptor: AGATCGGAAGAGC). The clean reads were mapped to the

Arabidopsis thaliana genome using the Bowtie program[62]. For miRNA analysis, the small RNA reads were mapped to miRBase v21. The sequences of miRNA and miRNA* were obtained from the PRIMEN database: http://www.pmiren.com/ftpdownload/Arabidopsis_thaliana_Ath/Arabidopsis_thalianamature.fa. For tasiRNA analysis, the reads that were mapped to 100-bp windows overlapping with *TAS1A*/*B*/*C* and *TAS2* loci were counted. The small RNA reads were normalized by calculating the RPM value (reads per million trimmed reads)[63]. The comparison between genotypes was conducted by the R package "DESeq2" with 1.5-fold change and P<0.01 as the cutoff [64]. For AGO1 IP sRNA-seq analysis, reads were normalized by total mapped reads. The comparison between genotypes was conducted with the two-factor model by DESeq2.

Small RNA gel blotting

RNA gel analysis of small RNAs was performed as described [65]. 10μg of total RNAs was separated on 15% polyacrylamide/8M urea gels. After gel electrophoresis, the RNAs were transferred to a Hybond-NX nylon membrane (GE healthcare). Antisense complementary oligonucleotides (Supplementary Table 2) were synthesized with both 5' and 3' end-labeled biotin. A probe complementary to U6 (5' CATCCTTGCGCAGGGGCCA 3') was used to detect U6 as an internal control. Hybridization was performed for 16 h at 55°C followed by washes. Signals were detected using the chemiluminescent Nucleic Acid Detection Module (Thermo Fisher, 89880).

Reverse transcription quantitative PCR (RT-qPCR)

RT-qPCR was performed to quantify mRNA and pri-miRNAs levels. 1μg total RNA was reverse transcribed with oligo (dT) using the PrimeScript™ 1st Strand cDNA Synthesis Kit (TAKARA, 6110A) according to the manufacturer's instructions. RT-qPCR was performed in the 96-well StepOnePlus real-time system (ABI) using the SYBR premix ExTaq II kit (TAKARA, RR820A). The following scheme was used: 95°C for 30s and 40 cycles of 95°C for 5s and 60°C for 30s. The levels of transcripts were normalized to the level of the internal standard *UBQ5* and the $2^{-\triangle\triangle}$CT values of control samples were set to 1. The primers for pri-miRNAs are listed in Supplementary Table 2. Three independent biological replicates were carried out for each genotype. Student's *t* test was used for the evaluation of statistical significance.

Histochemical GUS staining

14-day-old seedlings of *pMIR167a:GUS* and *pMIR167a:GUS rbv-1* (homozygous for both transgene and the *rbv-1* mutation) were subjected to histochemical GUS staining according to the standard protocol[66].

Chromatin Immunoprecipitation (ChIP) assay

RNA Polymerase II ChIP was performed as described (Kim et al., 2011) with 14-day-old Col and *rbv-1* seedlings using an antibody against RPB1 (Abcam, ab817). RT-qPCR was performed with co-

immunoprecipitated DNA, using primers listed in Supplementary Table 2. Relative enrichment was calculated by normalizing the amount of ChIP-ed DNA to the corresponding amount in the input.

Western blot Analysis

Western blots were performed as described [65]. Proteins from 14-day-old seedlings were extracted, resolved in 12% (v/v) SDS-polyacrylamide gels, and transferred to Hybond C-Extra membranes (Amersham Biosciences). The membranes were blocked with 5% (w/v) non-fat milk in Tris buffered saline Tween (TBST) buffer and then probed with specific antibodies. Antibodies used included anti-GAPDH (Santa Cruz Biotechnology, sc-365062, dilution, 1:1000), anti-SUL (dilution, 1:1000) [36], anti-HYL1 (Agrisera, AS06136, dilution, 1:2000), anti-AGO1 (Agrisera, AS09527, dilution, 1:2000), anti-SE (Agrisera, AS09532A, dilution, 1:1000), and anti-GFP (Abcam, ab290, dilution, 1:3000). After three washes, the membranes were probed with horseradish peroxidase-conjugated, goat-anti-rabbit IgG (Bio-Rad, cat.172-1019) (dilution, 1:2000) or goat-anti-mouse IgG (Bio-Rad, cat.170-6516, dilution, 1:2000). The protein signals were detected with the Amersham^{TM} ECL^{TM} Prime Western Blotting Detection Reagent (GE healthcare, RPN2232) and visualized with the Chemiluminescence imaging system (Clinx Science Instruments Co. Ltd, China).

RNA-Seq data analysis

Total RNAs extracted from 14-day-old seedlings were sent for library construction at Novogene, China and libraries were sequenced on an Illumina Hiseq 4000 platform to generate paired-end reads of 150bp in length. The data analysis of RNA-seq libraries was carried out as described[67]. The clean reads were collapsed into nonredundant ones and mapped to the *Arabidopsis* genome (ARAPORT11) using STAR, allowing a maximum of eight mismatches per paired-end read[68]. Differentially expressed genes (DEGs) were identified between Col and *rbv-1* using cuffdiff[69] with FPKM >1, fold-change >2 and FDR<0.05 as filters. Expression levels of all genes and DEGs were plotted using the value of log2 (FPKM+1). GO enrichment analysis was performed with agriGO[70] using a webtool (http://bioinfo.cau.edu.cn/agriGO/). Only the top 20 terms were presented in this paper. The analysis of splicing defects was carried out using Araport 11 intron annotation and a previously developed pipeline known as SQUID (https://github.com/Xinglab/SQUID). In brief, the level of retained introns was calculated using two methods: PI_Junction (*intron - exon junction reads*/[*intron - exon junction reads* + *exon-exon reads*]) and PI_density (normalized intronic reads/normalized exonic reads). The differentially spliced introns were defined using a stringent cutoff: combined_FDR < 0.1, Diff_PI_Junction > 0.05, Diff_PI_Density > 0.05. PI_Junction was used to represent the levels of retained introns.

AGO1 IP assay

1g of 14-day-old seedlings was ground in liquid nitrogen, and IP buffer (50 mM Tris pH 7.5, 150 mM NaCl, 10% Glycerol, 0.1% NP-40, 1mM PMSF) and EDTA-free protease inhibitor mixture (Roche) were added to the powder, which was followed by 20 min incubation with gentle shaking at 4°C. The supernatant was

incubated with 100μL of dynabeads (Invitrogen, 10002D) for 2h at 4°C. After centrifugation, the supernatant was used for IP. 50 ul was saved as input, and the rest was incubated with anti-AGO1 antibody (Agrisera, AS09527, dilution, 1:2000) for 2 h at 4 ℃. The beads were washed with wash buffer (IP buffer with 0.5% NP-40), and 1/10 (v/v) was added to 2x SDS-loading buffer for western blot analysis and 9/10 (v/v) was used for RNA isolation. Small RNA libraries were prepared using the NEBNext Multiplex Small RNA Library Prep Set for Illumina (New England Biolabs, E7300), and sequenced on an Illumina HiSeq2500 platform at BerryGenomics, China. The proteins were separated by SDS-PAGE and protein gel blot analysis was performed using anti-AGO1 antibody (Agrisera, AS09527, dilution, 1:2000).

Gel filtration assay

1g of 14-day-old T3 *pRBV: RBV-eYFP rbv-1* seedlings and wild type was collected and ground in liquid nitrogen. Then 1.5 ml phosphate buffer supplemented with 1mM PMSF, 1% EDTA-free protease inhibitor mixture, RNase Inhibitor (TAKARA) and 0.4% CA630 was added to the powder. The homogenized crude extracts were kept on ice for 20 min. After two rounds of centrifugation for 20 min (12000 g) each at 4℃, 1.5ml of the cleared crude extract was immediately used for sample injection. Gel-filtration chromatography was carried out as described[27] using a Superdex 200 Increase 10/300GL column. In total, 44 fractions (each 800 μL) were collected. 300 μL of each fraction was used for protein extraction for detecting AGO1 and 500 μL for total RNA purification for sRNA detection. Markers of known sizes (Aprotinin, 6.5kDa; Ribonuclease A, 13.7kDa; Carbonic Anhydrase, 29 kDa; Conalbumin, 75kDa; Aldolase, 158 kDa; Ferritin, 440 kDa; Thyroglobulin, 669 kDa; blue dextran, 2000kDa) were used for column calibration.

RBV conservation analysis

hmmsearch (HMMER 3.3) was employed to search for RBV homologs in a plant genome database (Supplementary Data 2) with a seed file created by the amino acid sequence of RBV. Close homologs of RBV were retained by a primary Neighbor-Joining tree based on MUSCLE (v3.8.1551)-aligned sequences and manual selection. The coding sequences of these genes were aligned by MUSCLE and manually corrected. The alignment was used to create the final Maximun Likelihood tree by RAxML v8.2.4 with parameters "-f a -m GTRCAT -x 12345 -p 1234567 -# 1000 -T 12". The output tree was visualized by MEGAX v10.1.8.

Data availability

All raw and processed data were deposited at NCBI GEO (http://www.ncbi.nlm.nih.gov/geo/) with the accession number GSE152911.

Accession Numbers

Genes referred to in this study correspond to the following *Arabidopsis* Genome Initiative locus identifiers: *RBV* (AT5G64730), *SUL*/*CHL42* (AT4G18480), *SPL3* (AT2G33810), *SPL5* (AT3G15270), *SPL10*

(AT1G27370), *PHB* (AT2G34710), *PHV* (AT1G30490), *REV* (AT5G60690), *AGO1* (AT1G48410), *MYB33* (AT5G06100), *MYB65* (AT3G11440), *ARF8* (AT5G37020), *CSD2* (AT2G28190), *CUC2* (AT5G53950)、*HYL1* (AT1G09700), *SE* (AT2G27100)、*DCL1* (AT1G01040), *NRPB1* (AT4G35800), *RPB2/NRPB2* (AT4G21710), *HEN1* (AT4G20910), *DDL* (AT3G20550), *CBP20* (AT5G44200), *CBP80* (AT2G13540), *CPL1* (AT4G21670), *MOS2* (AT1G33520), *HASTY* (AT3G05040), *NOT2A* (AT1G07705), *NOT2B* (AT5G18230), *CDC5* (AT1G09770), *TGH* (AT5G23080), *PRL1* (AT4G15900), *MAC7* (AT2G38770).

Declarations

Acknowledgments

We thank Dr. Detlef Weigel for sharing the *pSUC2:amiR-SUL* transgenic line. This work was funded by Guangdong Innovation Team Project (2014ZT05S078), National Natural Science Foundation of China (31701112, 31870287), China Postdoctoral Science Foundation (2015M582402) and National Institutes of Health (GM129373).

Author contributions

C.L., Q.C., and X.C. designed the project and experiments. C.L., Q.C., C.X., F.W. and D.C., performed the experiments. C.L, S.L., C.Y. and L.G. analyzed the data. C.L., Q.C., B.M., and X.C. wrote the manuscript.

Competing interests: The authors declare no conflict of interest.

References

1. Yu Y, Jia T, Chen X. The 'how' and 'where' of plant microRNAs. *The New phytologist* **216**, 1002–1017 (2017).
2. Song XW, Li Y, Cao XF, Qi YJ. MicroRNAs and Their Regulatory Roles in Plant-Environment Interactions. *Annu Rev Plant Biol* **70**, 489–525 (2019).
3. Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC. Expression of Arabidopsis MIRNA genes. *Plant physiology* **138**, 2145–2154 (2005).
4. Zheng B, Wang Z, Li S, Yu B, Liu JY, Chen X. Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. *Genes Dev* **23**, 2850–2860 (2009).
5. Kim YJ, Zheng B, Yu Y, Won SY, Mo B, Chen X. The role of Mediator in small and long noncoding RNA production in Arabidopsis thaliana. *Embo J* **30**, 814–822 (2011).
6. 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).
7. Zhang S, Xie M, Ren G, Yu B. CDC5, a DNA binding protein, positively regulates posttranscriptional processing and/or transcription of primary microRNA transcripts. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 17588–17593 (2013).

8. Fang X, Cui Y, Li Y, Qi Y. Transcription and processing of primary microRNAs are coupled by Elongator complex in Arabidopsis. *Nature plants* **1**, 15075 (2015).
9. Zhang B, *et al.* Linking key steps of microRNA biogenesis by TREX-2 and the nuclear pore complex in Arabidopsis. *Nature plants* **6**, 957–969 (2020).
10. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. MicroRNAs in plants. *Genes Dev* **16**, 1616–1626 (2002).
11. Park W, Li J, Song R, Messing J, Chen X. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. *Current biology: CB* **12**, 1484–1495 (2002).
12. Fukudome A, Fukuhara T. Plant dicer-like proteins: double-stranded RNA-cleaving enzymes for small RNA biogenesis. *J Plant Res* **130**, 33–44 (2017).
13. Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ. Primary microRNA transcripts are processed co-transcriptionally. *Nat Struct Mol Biol* **15**, 902–909 (2008).
14. Pawlicki JM, Steitz JA. Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J Cell Biol* **182**, 61–76 (2008).
15. Vazquez F, Gasciolli V, Crete P, Vaucheret H. The nuclear dsRNA binding protein HYL1 is required for MicroRNA accumulation and plant development, but not posttranscriptional transgene silencing. *Current Biology* **14**, 346–351 (2004).
16. Han MH, Goud S, Song L, Fedoroff N. The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 1093–1098 (2004).
17. Yang L, Liu ZQ, Lu F, Dong AW, Huang H. SERRATE is a novel nuclear regulator in primary microRNA processing in Arabidopsis. *Plant J* **47**, 841–850 (2006).
18. Lobbes D, Rallapalli G, Schmidt DD, Martin C, Clarke J. SERRATE: a new player on the plant microRNA scene. *EMBO Rep* **7**, 1052–1058 (2006).
19. Song L, Han MH, Lesicka J, Fedoroff N. Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 5437–5442 (2007).
20. Iwata Y, Takahashi M, Fedoroff NV, Hamdan SM. Dissecting the interactions of SERRATE with RNA and DICER-LIKE 1 in Arabidopsis microRNA precursor processing. *Nucleic Acids Res* **41**, 9129–9140 (2013).
21. Fang Y, Spector DL. Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. *Current biology: CB* **17**, 818–823 (2007).
22. Yu B, *et al.* Methylation as a crucial step in plant microRNA biogenesis. *Science* **307**, 932–935 (2005).
23. Baumberger N, Baulcombe DC. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits rnicroRNAs and short interfering RNAs. *Proceedings of the National Academy of Sciences of the*

United States of America **102**, 11928–11933 (2005).

24. Bologna NG, *et al.* Nucleo-cytosolic Shuttling of ARGONAUTE1 Prompts a Revised Model of the Plant MicroRNA Pathway. *Mol Cell* **69**, 709–719 e705 (2018).
25. Qi Y, Denli AM, Hannon GJ. Biochemical specialization within Arabidopsis RNA silencing pathways. *Mol Cell* **19**, 421–428 (2005).
26. Wang H, *et al.* Deep sequencing of small RNAs specifically associated with Arabidopsis AGO1 and AGO4 uncovers new AGO functions. *Plant J* **67**, 292–304 (2011).
27. Dalmadi A, Gyula P, Balint J, Szittya G, Havelda Z. AGO-unbound cytosolic pool of mature miRNAs in plant cells reveals a novel regulatory step at AGO1 loading. *Nucleic Acids Res* **47**, 9803–9817 (2019).
28. de Felippes FF, Ott F, Weigel D. Comparative analysis of non-autonomous effects of tasiRNAs and miRNAs in Arabidopsis thaliana. *Nucleic Acids Res* **39**, 2880–2889 (2011).
29. Morel JB, *et al.* Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell* **14**, 629–639 (2002).
30. Miller JC, Chezem WR, Clay NK. Ternary WD40 Repeat-Containing Protein Complexes: Evolution, Composition and Roles in Plant Immunity. *Frontiers in plant science* **6**, 1108 (2015).
31. Xu C, Min J. Structure and function of WD40 domain proteins. *Protein Cell* **2**, 202–214 (2011).
32. Xie D, *et al.* Phase separation of SERRATE drives dicing body assembly and promotes miRNA processing in Arabidopsis. *Nat Cell Biol* **23**, 32–39 (2021).
33. Vaucheret H, Vazquez F, Crete P, Bartel DP. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev* **18**, 1187–1197 (2004).
34. Fei Q, Xia R, Meyers BC. Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell* **25**, 2400–2415 (2013).
35. You C, *et al.* FIERY1 promotes microRNA accumulation by suppressing rRNA-derived small interfering RNAs in Arabidopsis. *Nature communications* **10**, 4424 (2019).
36. Jia T, *et al.* The Arabidopsis MOS4-associated Complex Promotes MicroRNA Biogenesis and Precursor Messenger RNA Splicing. *Plant Cell*, (2017).
37. 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).
38. Wang S, *et al.* The PROTEIN PHOSPHATASE4 Complex Promotes Transcription and Processing of Primary microRNAs in Arabidopsis. *Plant Cell* **31**, 486–501 (2019).
39. Laubinger S, *et al.* Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 8795–8800 (2008).
40. Zhan X, *et al.* Arabidopsis proline-rich protein important for development and abiotic stress tolerance is involved in microRNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 18198–18203 (2012).

41. Zielezinski A, *et al.* mirEX 2.0 - an integrated environment for expression profiling of plant microRNAs. *Bmc Plant Biol* **15**, 144 (2015).
42. Kim S, Yang JY, Xu J, Jang IC, Prigge MJ, Chua NH. Two cap-binding proteins CBP20 and CBP80 are involved in processing primary MicroRNAs. *Plant & cell physiology* **49**, 1634–1644 (2008).
43. Ben Chaabane S, *et al.* STA1, an Arabidopsis pre-mRNA processing factor 6 homolog, is a new player involved in miRNA biogenesis. *Nucleic Acids Res* **41**, 1984–1997 (2013).
44. Ren G, Xie M, Dou Y, Zhang S, Zhang C, Yu B. Regulation of miRNA abundance by RNA binding protein TOUGH in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 12817–12821 (2012).
45. Qiao Y, Shi J, Zhai Y, Hou Y, Ma W. Phytophthora effector targets a novel component of small RNA pathway in plants to promote infection. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 5850–5855 (2015).
46. Furumizu C, Tsukaya H, Komeda Y. Characterization of EMU, the Arabidopsis homolog of the yeast THO complex member HPR1. *Rna* **16**, 1809–1817 (2010).
47. Francisco-Mangilet AG, *et al.* THO2, a core member of the THO/TREX complex, is required for microRNA production in Arabidopsis. *Plant J* **82**, 1018–1029 (2015).
48. Wu X, *et al.* A role for the RNA-binding protein MOS2 in microRNA maturation in Arabidopsis. *Cell Res* **23**, 645–657 (2013).
49. Yu B, *et al.* The FHA domain proteins DAWDLE in Arabidopsis and SNIP1 in humans act in small RNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 10073–10078 (2008).
50. Zhang SX, Liu YH, Yu B. PRL1, an RNA-Binding Protein, Positively Regulates the Accumulation of miRNAs and siRNAs in Arabidopsis. *PLoS genetics* **10**, (2014).
51. Wang Z, *et al.* SWI2/SNF2 ATPase CHR2 remodels pri-miRNAs via Serrate to impede miRNA production. *Nature* **557**, 516–521 (2018).
52. Iki T, *et al.* In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90. *Mol Cell* **39**, 282–291 (2010).
53. Iwasaki S, *et al.* Hsc70/Hsp90 Chaperone Machinery Mediates ATP-Dependent RISC Loading of Small RNA Duplexes. *Mol Cell* **39**, 292–299 (2010).
54. Johnston M, Geoffroy MC, Sobala A, Hay R, Hutvagner G. HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells. *Mol Biol Cell* **21**, 1462–1469 (2010).
55. Iki T, Yoshikawa M, Meshi T, Ishikawa M. Cyclophilin 40 facilitates HSP90-mediated RISC assembly in plants. *Embo J* **31**, 267–278 (2012).
56. Smith MR, *et al.* Cyclophilin 40 is required for microRNA activity in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 5424–5429 (2009).
57. Tomassi AH, *et al.* The Intrinsically Disordered Protein CARP9 Bridges HYL1 to AGO1 in the Nucleus to Promote MicroRNA Activity. *Plant physiology* **184**, 316–329 (2020).

58. Murray MG, Thompson WF. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* **8**, 4321–4325 (1980).
59. Nakamura S, *et al.* Gateway binary vectors with the bialaphos resistance gene, bar, as a selection marker for plant transformation. *Biosci Biotechnol Biochem* **74**, 1315–1319 (2010).
60. Clough SJ, Bent AF. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* **16**, 735–743 (1998).
61. Liu C, Axtell MJ, Fedoroff NV. The helicase and RNaseIIIa domains of Arabidopsis Dicer-Like1 modulate catalytic parameters during microRNA biogenesis. *Plant physiology* **159**, 748–758 (2012).
62. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25 (2009).
63. Nobuta K, McCormick K, Nakano M, Meyers BC. Bioinformatics analysis of small RNAs in plants using next generation sequencing technologies. *Methods Mol Biol* **592**, 89–106 (2010).
64. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550 (2014).
65. Cai Q, *et al.* The disease resistance protein SNC1 represses the biogenesis of microRNAs and phased siRNAs. *Nature communications* **9**, 5080 (2018).
66. Jefferson RA, Kavanagh TA, Bevan MW. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *Embo J* **6**, 3901–3907 (1987).
67. Li S, Wang Y, Zhao Y, Zhao X, Chen X, Gong Z. Global Co-transcriptional Splicing in Arabidopsis and the Correlation with Splicing Regulation in Mature RNAs. *Mol Plant* **13**, 266–277 (2020).
68. Dobin A, *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
69. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-sEq. *Nat Biotechnol* **31**, 46-+ (2013).
70. Tian T, *et al.* agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res* **45**, W122-W129 (2017).

Figures

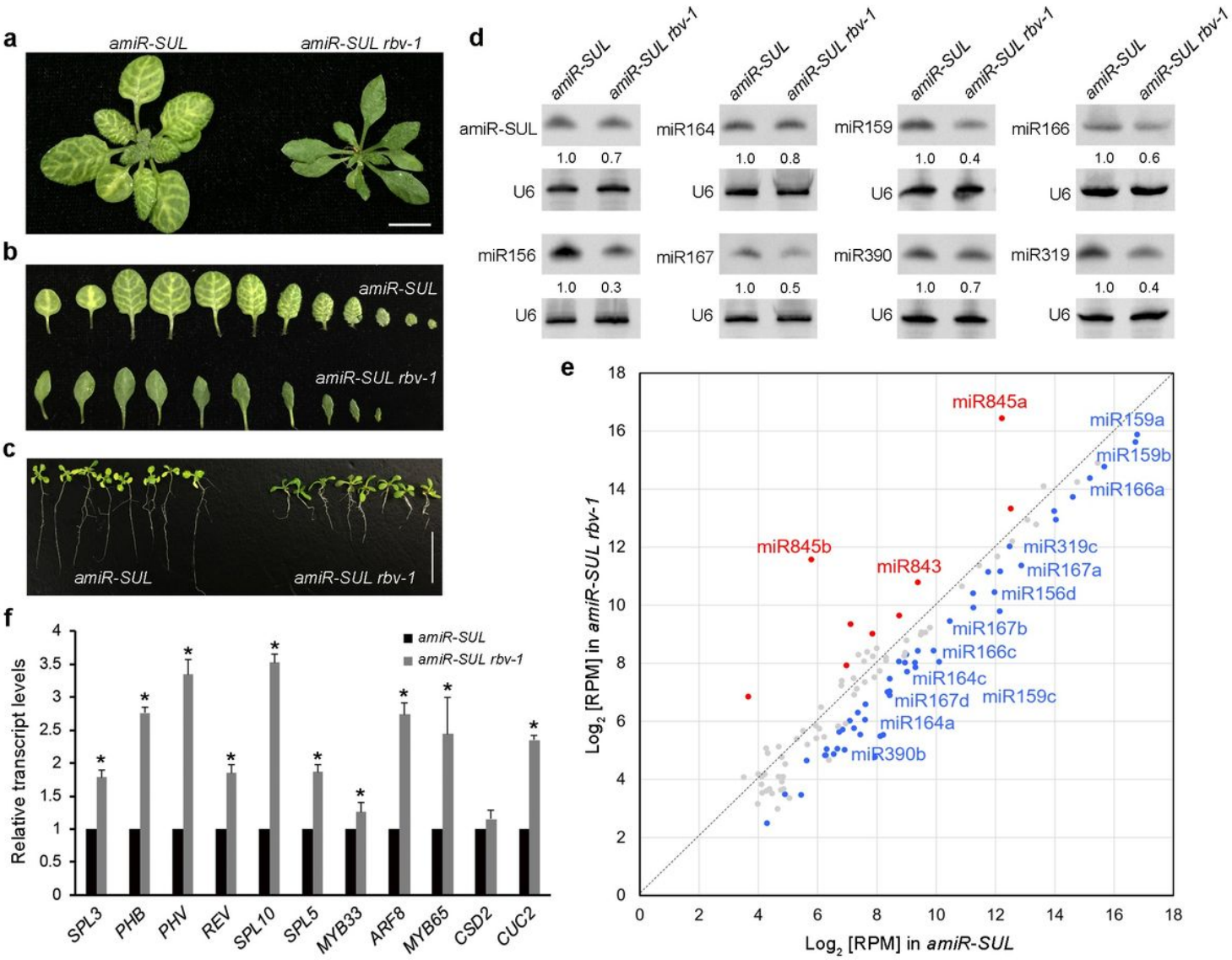

Figure 1

Isolation of a silencing suppressor mutant rbv-1 from an amiR-SUL line. a Phenotypes of 1-month-old amiR-SUL rbv-1 and amiR-SUL plants. Bar=1cm. b Images of rosette leaves from 1-month-old plants grown under long-day conditions. c 14-day-old seedlings showing reduced root length in amiR-SUL rbv-1. Bar=1cm. d RNA gel analysis showing reduced accumulation of amiR-SUL and endogenous miRNAs in the amiR-SUL rbv-1 mutant. U6 was used as an internal control. e A scatter plot showing the abundance of miRNAs in amiR-SUL rbv-1 and amiR-SUL as determined by small RNA-seq with 14-day-old seedlings. miRNA abundance was calculated as reads per million mapped reads (RPM) and miRNAs with RPM >10 in either genotype are shown. The red dots indicate miRNAs with higher levels in amiR-SUL rbv-1, and the blue dots indicate miRNAs with lower levels in amiR-SUL rbv-1. (*Student's t test: $P < 0.05$). f Determination of miRNA target mRNA levels in amiR-SUL and amiR-SUL rbv-1 in 14-day-old seedlings by RT-qPCR. UBQUITIN5 (UBQ5) was used as the internal control. The values were relative to those in amiR-SUL. Error bars represent standard deviation from three technical replicates. Asterisks indicate significant difference between amiR-SUL and amiR-SUL rbv-1 (*Student's t test, $P < 0.05$)

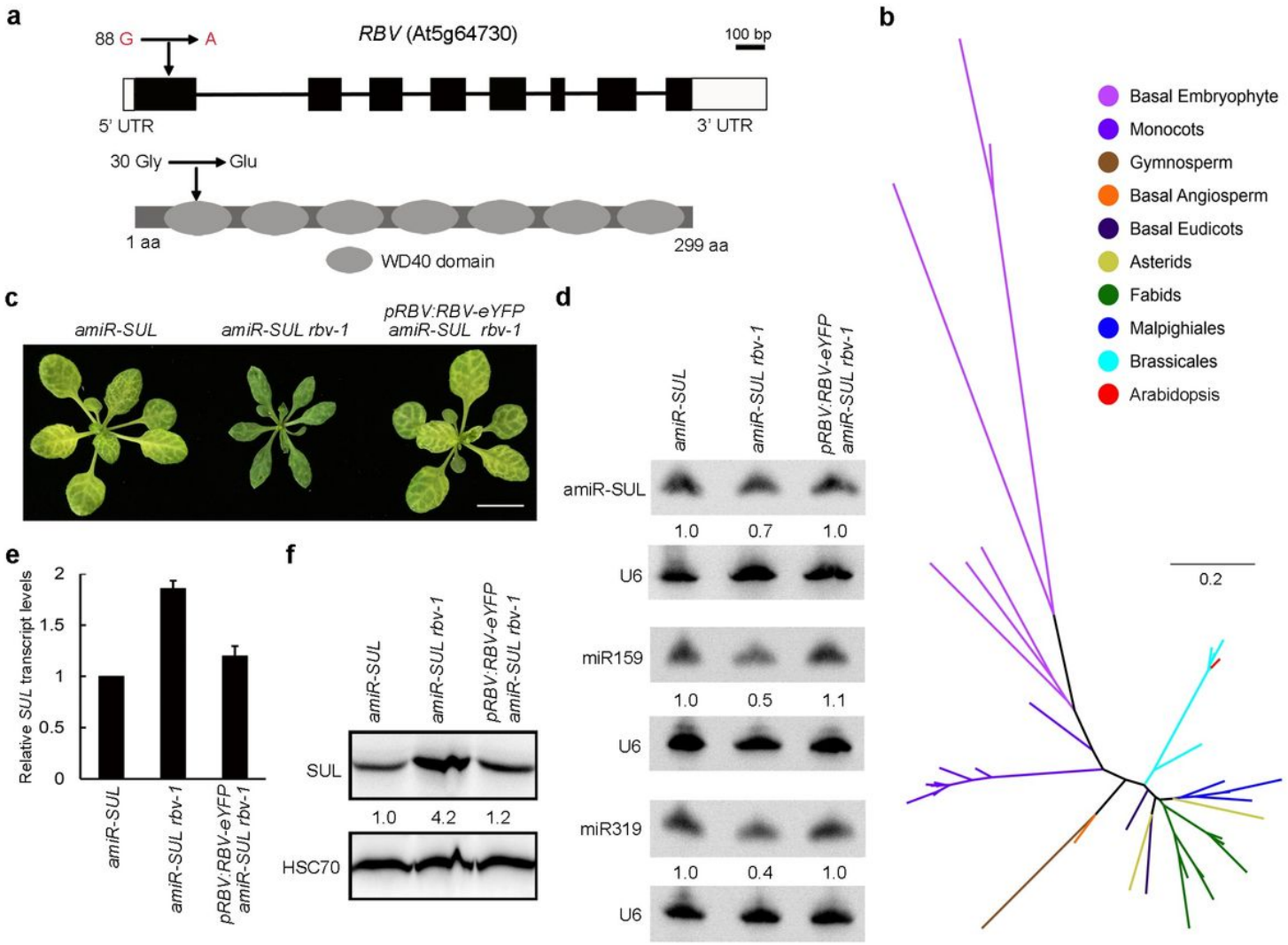

Figure 2

Identification of AT5G64730 as RBV. a Diagrams of the RBV (AT5G64730) gene (upper panel) and protein (lower panel). Rectangles and lines represent exons and introns, respectively. Black and white rectangles represent the coding region and the UTRs, respectively. The point mutation in rbv-1 and the corresponding change at the amino acid level are indicated (arrows). The protein domains were predicted (http://smart.embl-heidelberg.de/). b A phylogenetic tree of RBV and its orthologs in plants. Colors represent different lineages of plant species. c 3-week-old plants of the indicated genotypes. pRBV:RBV-eYFP was introduced into amiR-SUL rbv-1. Bar = 1cm. d RNA gel blot analysis of miRNAs from amiR-SUL, amiR-SUL rbv-1, and the complementation line pRBV:RBV-eYFP amiR-SUL rbv-1 using 14-day-old seedlings. U6 was used as an internal control. e RT-qPCR to determine RNA levels of the amiR-SUL target gene SUL in the indicated genotypes. Three independent biological replicates were used for the calculation of standard deviation. f Protein gel blot analysis to determine the protein levels of the amiR-SUL target gene SUL in the indicated genotypes.

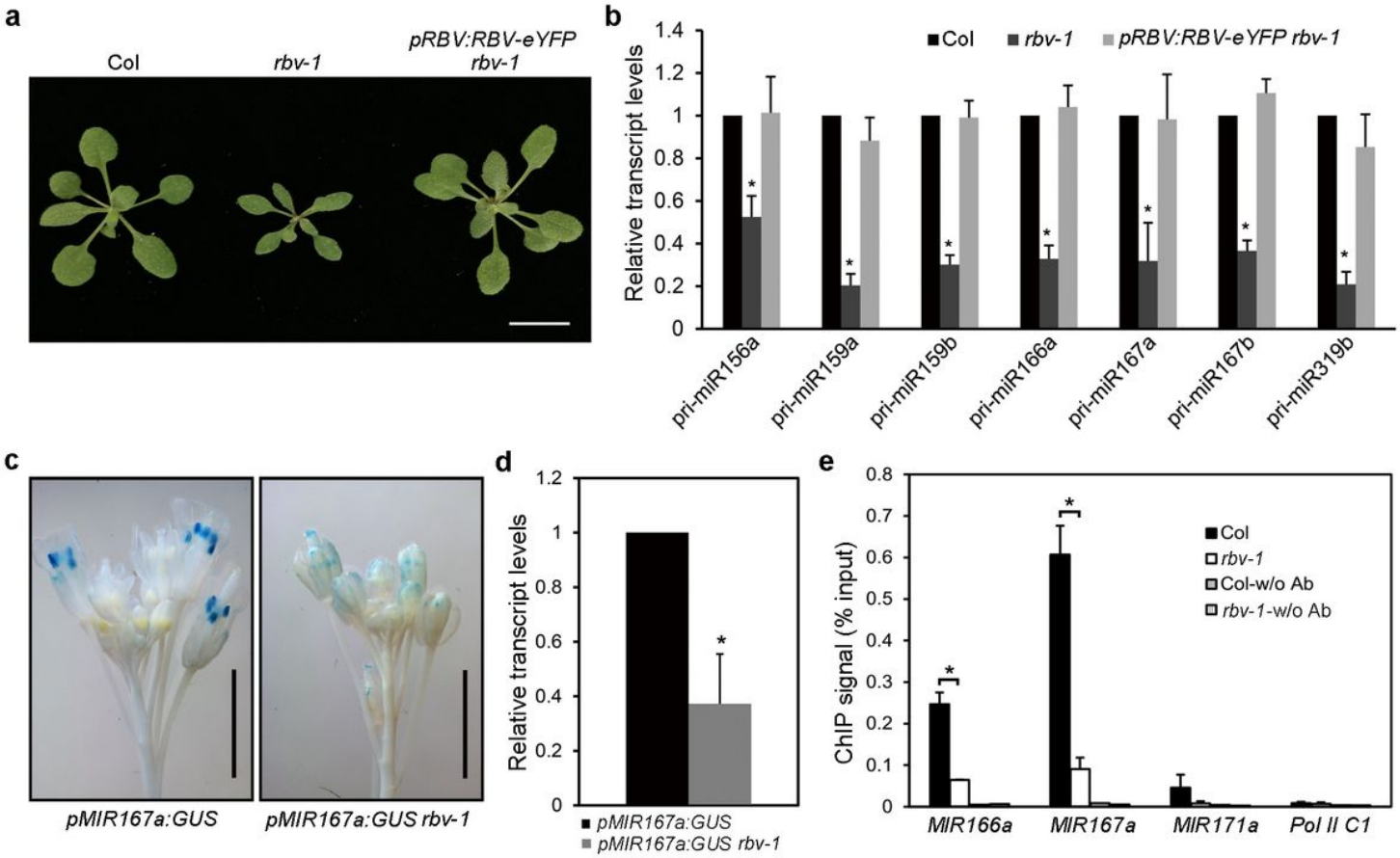

Figure 3

RBV promotes the transcription of MIR genes. a 3-week-old plants of the indicated genotypes. Bar = 1cm. b Levels of seven pri-miRNAs in 14-day-old seedlings of Col, rbv-1 and the complementation line pRBV:RBV-eYFP rbv-1 as determined by RT-qPCR. UBQ5 was used as the internal control. The expression values were relative to those in Col, which were set to 1. Error bars represent standard deviation calculated from three independent replicates. (*Student's t test, $P < 0.05$). c Representative images of GUS staining of pMIR167a:GUS and pMIR167a:GUS rbv-1 inflorescences. Bars = 2 mm. d The transcript levels of GUS from the two genotypes as determined by RT-qPCR. The expression values were relative to pMIR167a:GUS. Error bars represent standard deviation calculated from three independent replicates. (*Student's t test, $P < 0.05$). e RBV is required for the recruitment of Pol II to MIR166a and MIR167a promoters. The occupancy of Pol II at various regions was determined by ChIP with rbv-1 and Col using an antibody that recognizes the C-terminal repeat (YSPTSPS) of the largest subunit of Pol II. ChIP performed without the antibody served as a negative control. A genomic region between the genes AT2G17460 and AT2G17470 named Pol II C1 was also used as a negative control. Means and standard deviation from three independent replicates are presented. (*Student's t test, $P < 0.05$)

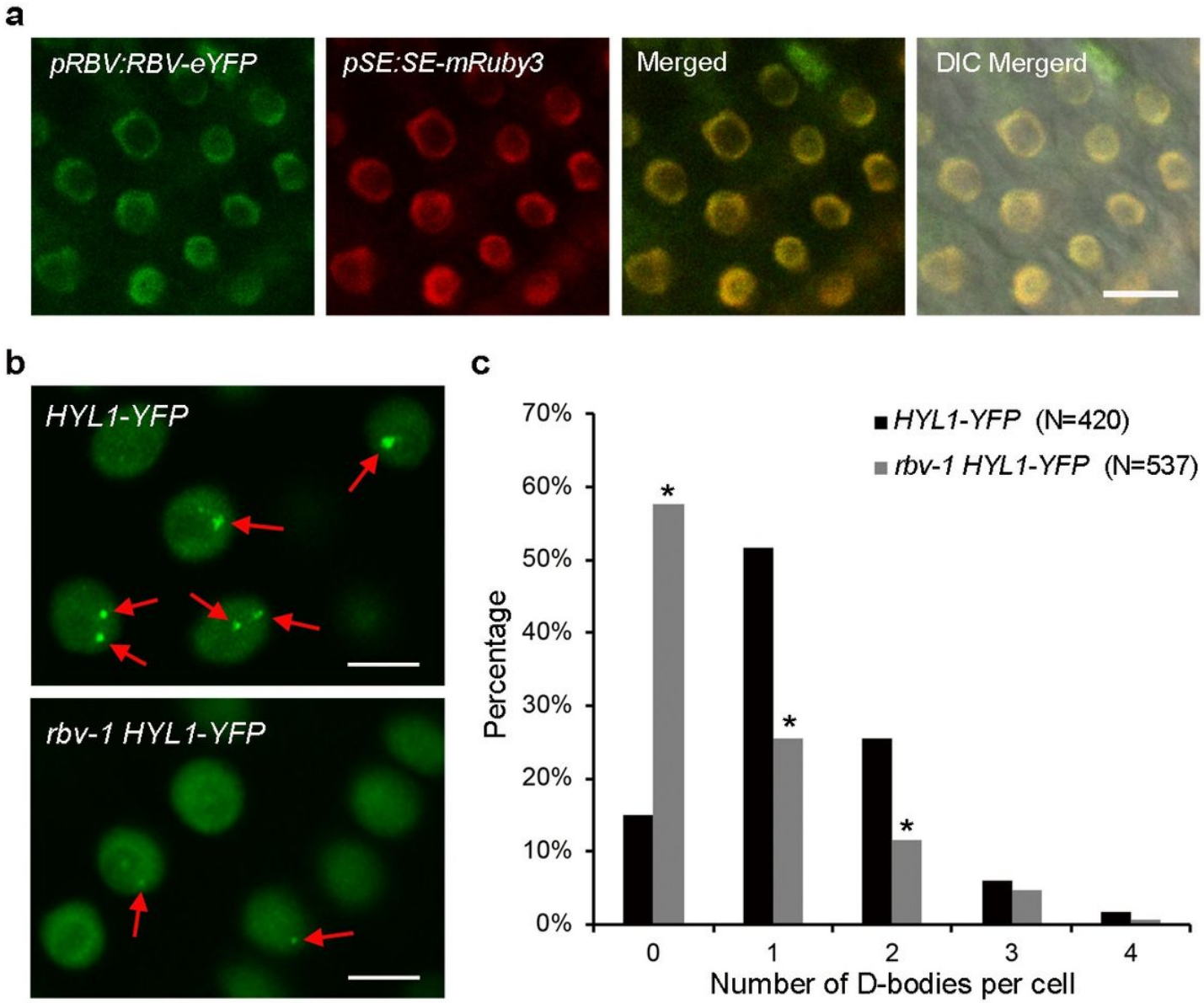

Figure 4

RBV is required for the proper localization of HYL1 in D-bodies. a RBV is localized in the nucleoplasm. eYFP and mRuby3 signals were detected in root cells from pRBV:RBV-eYFP and pSE:SE-mRuby3 transgenic plants, respectively. Bar = 10 μm. b Representative images of HYL1-YFP signals in root cells from the meristematic zone in the two genotypes. Arrows indicate D-bodies. Bars = 5 μm. c The percentage of cells with 1-4 D-bodies per cell in wild type and rbv-1. The X axis represents the number of D-bodies per cell, and the Y axis represents the percentage of cells with the corresponding number of D-bodies. “N” means the numbers of total root cells examined. * $P<0.05$, which was calculated by the Wilcoxon-Matt-Whitney test.

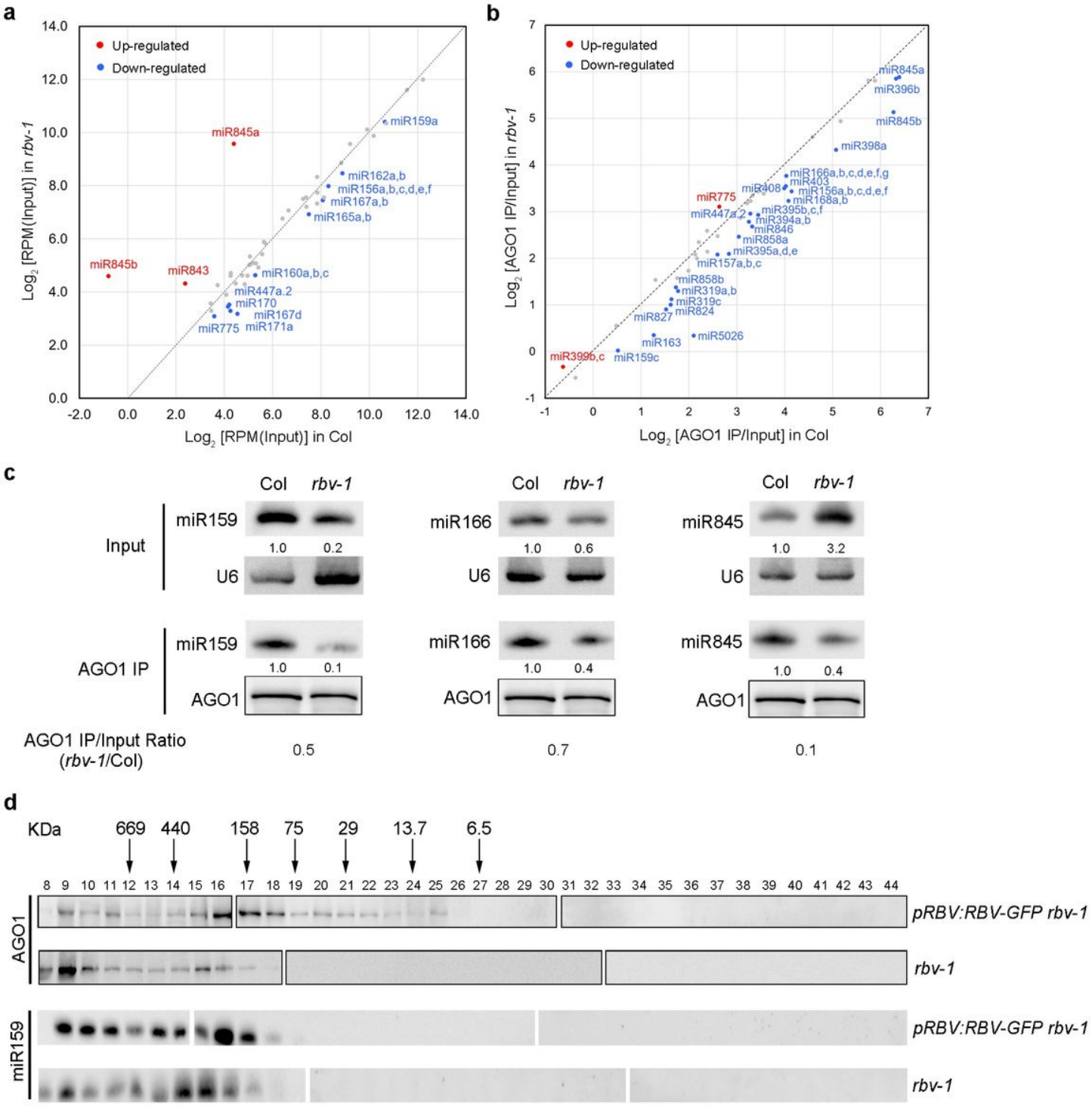

Figure 5

Mutation of RBV leads to a defect of miRNA loading into AGO1. a A scatter plot of miRNA abundance in rbv-1 input vs. Col input. All miRNAs were normalized by total trimmed reads, and those with RPM >10 in either genotype are shown. The red dots indicate miRNAs showing increased abundance in rbv-1, and the blue dots indicate miRNAs with reduced abundance in rbv-1. (*Student's t test: $P < 0.05$). b A scatter plot showing the AGO1 loading efficiency of miRNAs in rbv-1 vs. Col as determined by AGO1 IP small RNA-seq. AGO1 loading efficiency is represented by the ratio of miRNA abundance in AGO1 IP vs input. All

miRNAs with RPM value >10 in either genotype in the input samples (as in a) are shown here. The red dots indicate miRNAs with increased AGO1 association in rbv-1, and the blue dots indicate miRNAs with reduced AGO1 association in rbv-1. c RNA gel blot analysis of three miRNAs before (input) and after AGO1 IP. U6 was used as an internal control for the input samples. For the IP samples, a portion was used for protein gel blot analysis to quantify AGO1 protein levels. The levels of miRNAs in the IP samples were normalized against AGO1 protein levels. No matter whether the assayed miRNAs were increased or reduced in abundance in input samples, they all showed reduced AGO1 association. d Size exclusion chromatography with pRBV:RBV-eYFP rbv-1 and rbv-1 samples followed by western blotting to detect AGO1 and northern blotting to detect miR159. The upper panel indicates the distribution of AGO1 while the lower panel represents the distribution of miR159 among the fractions. The numbers above the AGO1 blots indicate those of the fractions. Note that no AGO1 or miR159 was detected in fractions 1-7 (not shown). The positions of the molecular weight standards are shown above the AGO1 blots.

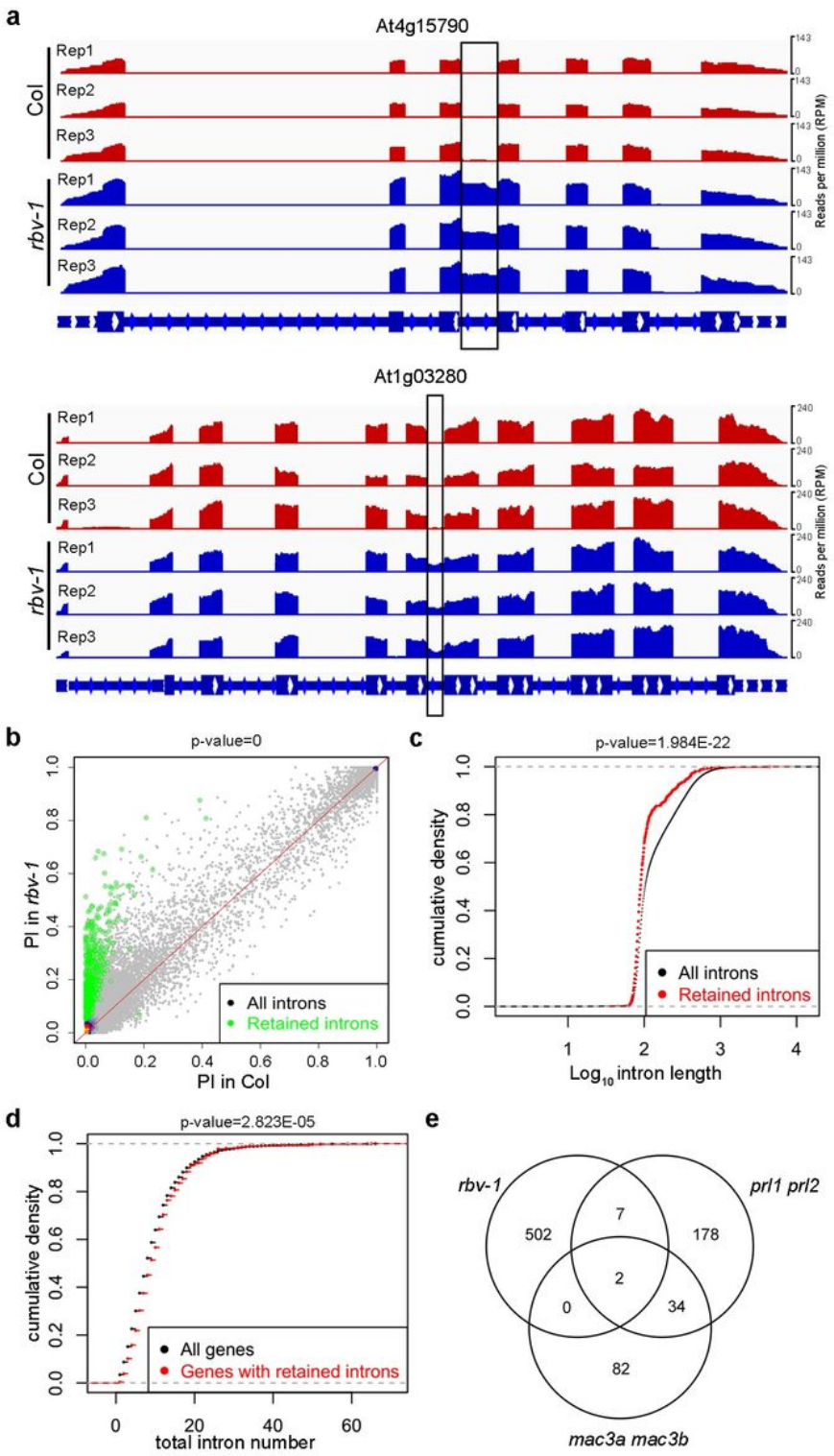

Figure 6

RBV function is required in the splicing of short introns in certain pre-mRNAs. a Examples of two genes with intron retention defects in the rbv-1 mutant. RNA-seq reads are shown against the gene models below. In the gene models, rectangles and lines represent exons and introns, respectively. The black rectangles indicate retained introns in the rbv-1 mutant. b A scatter plot showing percent retained introns (PI) in wild type and the rbv-1 mutant. The green dots represent introns with statistically significant

retention defects in the mutant (Wilcoxon test, P = 0). c Cumulative density plots of intron length for all introns and for retained introns in the rbv-1 mutant (Wilcoxon test, P=1.984E-22). d Cumulative density plots of intron number in all genes and for genes with retained introns. (Wilcoxon test, P =2.823E-05). e Venn diagrams showing the numbers of retained introns in rbv-1, prl1 prl2 and mac3a mac3b mutants, and the overlaps among the introns retained in these mutants.

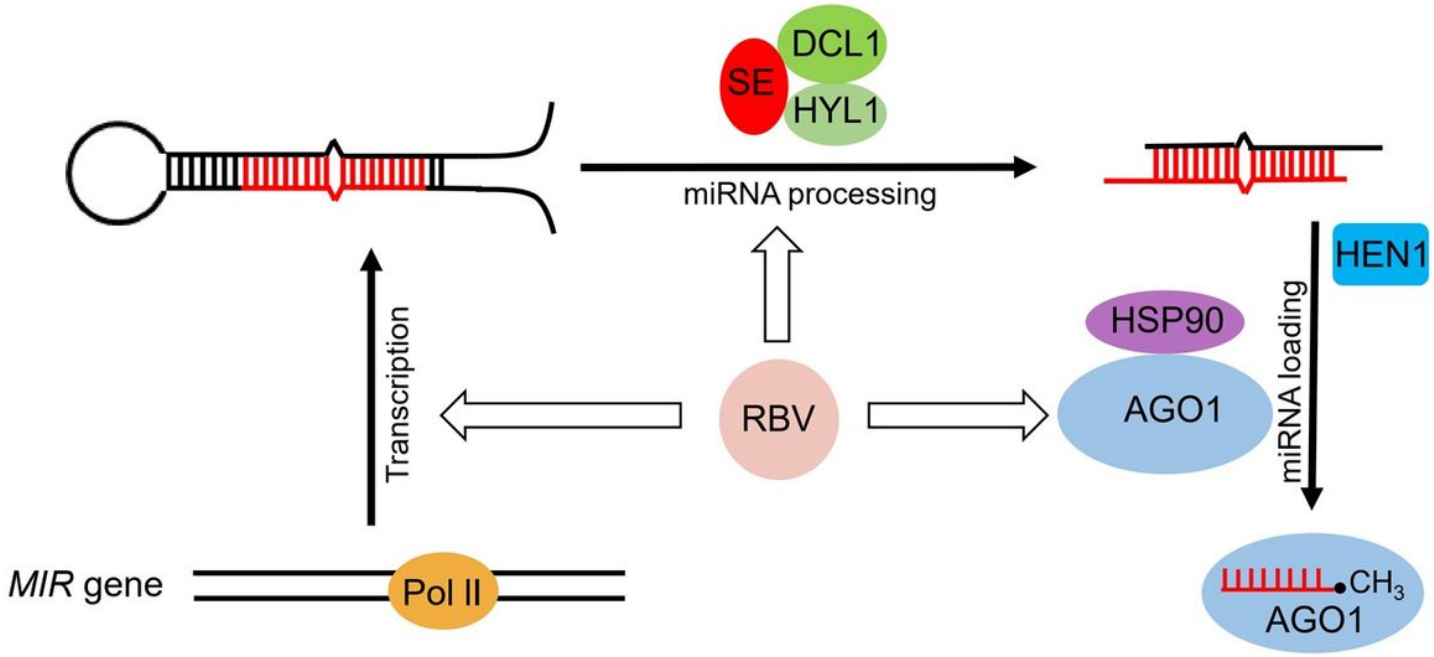

Figure 7

A model for the role of RBV in MIR gene transcription, pri-mRNA processing and AGO1 loading.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryData1.xlsx
- SupplementaryData2.xlsx
- SupplementaryData3.xlsx
- SupplementaryData4.xlsx
- SupplementaryData5.xlsx
- SupplementaryData6.xlsx
- SupplementaryData7.xlsx
- SupplementaryData8.xlsx
- Supplementaryinformation.docx