SMA1, a homolog of the splicing factor Prp28, has a multifaceted role in miRNA biogenesis in Arabidopsis

Shengjun Li1,2,3,†, Ran Xu4,‡, Aixia Li4,‡, Kan Liu2,3, Liqing Gu4, Mu Li2,3, Hairui Zhang5, Yueying Zhang4, Shangshang Zhuang4, Quanhui Wang1, Gang Gao5, Na Li4, Chi Zhang2,3, Yunhai Li4,* and Bin Yu2,3,*

1Key Laboratory of Biofuels, Shandong Provincial Key Laboratory of Energy Genetics, Qingdao Engineering Research Center of Biomass Resources and Environment, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao 266101, China, 2Center for Plant Science Innovation, University of Nebraska-Lincoln, Lincoln, NE 68588-0666, USA, 3School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE 68588-0118, USA, 4State Key Laboratory of Plant Cell and Chromosome Engineering, CAS Centre for Excellence in Molecular Plant Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China and 5School of Life Science, Shanxi Normal University, Linfen 041004, China

ABSTRACT
MicroRNAs (miRNAs) are a class of small non-coding RNAs that repress gene expression. In plants, the RNase III enzyme Dicer-like (DCL1) processes primary miRNAs (pri-miRNAs) into miRNAs. Here, we show that SMALL1 (SMA1), a homolog of the DEAD-box pre-mRNA splicing factor Prp28, plays essential roles in miRNA biogenesis in Arabidopsis. A hypomorphic sma1-1 mutation causes growth defects and reduces miRNA accumulation correlated with increased target transcript levels. SMA1 interacts with the DCL1 complex and positively influences pri-miRNA processing. Moreover, SMA1 binds the promoter region of genes encoding pri-miRNAs (MIRs) and is required for MIR transcription. Furthermore, SMA1 also enhances the abundance of the DCL1 protein levels through promoting the splicing of the DCL1 pre-mRNAs. Collectively, our data provide new insights into the function of SMA1/Prp28 in regulating miRNA abundance in plants.

INTRODUCTION
In plants, miRNAs are involved in various biological processes, such as growth and developmental regulation, abiotic and biotic stress responses, by targeting mRNA transcripts for degradation or translational repression (1–4). The biogenesis of most plant miRNAs begins with the transcription of primary miRNAs (pri-miRNAs) by the RNA polymerase II (Pol II) from miRNA encoding genes (MIRs) (5). Following transcription, pri-miRNAs are processed by the DICER-LIKE 1 (DCL1) complex, which is composed of the RNase III enzyme DCL1, the Zinc-finger protein SERRATE (SE) and the double-stranded RNA (dsRNA)-binding protein HYPONASTIC LEAVES 1 (HYL1), to release the miRNA/miRNA* duplexes in the nucleus (2,6–8). Among DCL1 components, DCL1 directly cleaves pri-miRNAs while SE and HYL1 improve the efficiency and accuracy of DCL1 activity (9–11). After processing, miRNAs are methylated by the miRNA methyltransferase HUA ENHANCER 1 (HEN1) to prevent untemplated uridine addition that leads to degradation (12–14), and then are loaded into the effector protein named ARGONAUTE 1 (AGO1) to recognize target mRNA transcripts for cleavage or translational inhibition (15–18).

Recently, many proteins have been shown to act in miRNA biogenesis pathway. At the transcriptional level, the transcriptional co-activator Mediator (19), NAGATIVE ON TATA LESS 2 (NOT2) (20), CELL CYCLIN-DEPENDENT KINASES (CDKs) (23) modulate the transcriptional activity of Pol II. After transcription, the RNA-binding proteins DAWDLE (DDL) (24) and PLEIOTROPIC REGULATORY LOCUS 1 (PRL1) (25) stabilize pri-miRNA transcripts, and then the RNA-binding protein TOUGH (TGH) (26), the THO/TREX complex (27) and the ribosomal protein STV1 (28) promote the recruitment of pri-miRNAs to the DCL1 complex. At the pri-miRNA processing step, the MOS4-associated complex (MAC), NOT2, Elongator, DDL and TGH act as facilitators of the DCL1 activity (20–22,24–26,29). The activity of the DCL1 complex is also subjected to post-translational
regulation. For instance, the phosphorylation status (30–32) and transportation from cytoplasm to nucleus of HYL1 protein is crucial for pri-miRNA processing (30,33). Moreover, the proper level of DCL1 can be transcriptionally or post-transcriptionally regulated by the splicing factor STABILIZED1 (STA1) (34) and the transcription factor XAP5 CIRCADIAN TIMEKEEPER (CMA33/XCT) (35). Interestingly, miRNA biogenesis and pre-mRNA splicing share some protein factors, such as the MAC complex, STA1 (34), CAP-BINDING PROTEINS (CBPs) (36) and SE (37), suggesting the potential interplay between miRNA biogenesis and mRNA metabolism.

Prp28 is a conserved splicing protein in eukaryotes. It encodes a DEAD-box helicase and is a component of U4/U5/U6 tri-small nuclear ribonucleoprotein (snRNP) in yeast and human. Prp28 is required for spliceosome assembly and functions in an adenosine triphosphate (ATP)-independent manner (38–41). Dysfunction of Prp28 in yeast and animals results in disorders of cell proliferation and differentiation (42). Compared with other organisms, plant Prp28 ortholog remains to be identified.

In this study, we report that the Prp28 homolog SMALL1 (SMA1) contributes to miRNA biogenesis in Arabidopsis. A partial loss-of-function sma1-1 mutant displays severe growth defects, while a null sma1-2 mutant is embryonic lethal. In sma1-1, the accumulation of miRNAs and pri-miRNAs is reduced, suggesting that SMA1 is required for miRNA biogenesis. SMA1 interacts with the DCL1 complex and influences pri-miRNA processing. In addition, SMA1 is crucial for proper splicing of the DCL1 pre-mRNAs and pri-miRNAs. Moreover, SMA1 binds the MIR promoters and promotes MIR transcription through enhancing the occupancy of Pol II at the MIR promoters. Taken together, these results demonstrated that SMA1 plays a multifaceted role in miRNA biogenesis and revealed a new function for a conserved protein.

MATERIALS AND METHODS

Plant materials and growth condition

All plant materials used in this study are in the Columbia genetic background (Col-0). sma1-1 contains a G-to-T mutation at the nucleotide position 1900-bp, which can be identified through NdeI digestion of polymerase chain reaction (PCR) product amplified with the primer pairs of 730-F and 730-R. sma1-2 (SALK_110742) was obtained from the Nottingham Arabidopsis Stock Center (NASC). sma1-sgR lines were generated by CRISPR/Cas9-induced mutation. Transgenic line containing a single copy of pMIR167a::GUS or p35S::HYL1-YFP were crossed to sma1-1. In the F2 generation, WT plants or sma1-1 harboring pMIR167a::GUS or p35S::HYL1-YFP were selected through PCR-based genotyping for sma1-1, GUS or GFP. Approximately 15 WT or sma1-1 plants were pooled for GUS transcript level analyses. The primers used for genotyping are listed in Supplementary Table S1.

To germinate plants, seeds were sterilized using 50% bleach for 10 min, followed by four-time washes with sterile distilled water. After vernalization at 4°C for 3 days, seeds were sown on medium containing 1/2 Murashige and Skoog (MS), 1% glucose and 0.8% agar. 10-day-old seedlings were transferred to soil. All plants were grown at 22°C with 16-h light and 8-h dark cycles in a growth chamber.

Map-based cloning of SMA1

To map the sma1-1 mutation, sma1-1 was crossed with Landsberg erecta (Ler) to generate F2 population. Map-based cloning were performed with the molecular markers listed in the Supplementary Table S1. sma1-1 mutation was mapped into a 31-kb interval between markers T1B8-9 and T1B8-12 on the chromosome 2. The sma1-1 mutation within this region was then identified through DNA-sequencing.

Plasmid construction

The protein encoding region of SMA1 was PCR-amplified with primers SMA1CDS-F and SMA1CDS-R, cloned into PCR8/GW/TOPO cloning vector (Invitrogen) and subsequently subcloned into pMDC43 binary vector to generate 35S::GFP-SMA1 construct. To generate the bimolecular fluorescence complementation (BiFC) construct for SMA1, SMA1 cDNA was inserted into the multiple cloning sites of pSAT4-cCFP-C vector. Then, the resulting plasmid was digested with the I-sceI restriction enzyme to release the 35S::cCFP-SMA1 fragment, which was then subcloned to pPZP-RC32-ocs-bar-R1 vector. The constructs of nVenus-DCL1, nVenus-SE, nVenus-HYL1 and nVenus-AGO1 were described previously (26). The construct for CRISPR/Cas9-induced small mutation was generated with binary vector pHEE401 using BsaI (NEB) and T4 ligase (NEB) according to the method described previously (43). The primers were listed in Supplementary Table S1.

Morphological analyses and GUS histochemical staining

Morphological and cellular analyses were performed according to the previously reported methods (44). For GUS staining, tissues from sma1-1 or WT plants harboring pMIR167a::GUS were incubated for 5 h in the staining solution (1 mM 5-bromo-4-chloro-3-indoly-beta-D-glucuronic acid, 100 mM NaPO4 buffer, 3 mM each K2Fe(CN)6/K4Fe(CN)6, 10 mM EDTA and 0.1% NP-40) at 37°C. Seventy percent ethanol was used for tissue clearing before imaging.

Detection of DCL1, SE, HYL1 and AGO proteins in Arabidopsis

Protein extracts from inflorescences of Arabidopsis were separated on a 10% Sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and detected with antibodies against HYL1 (Agrisera, AS09527) and HSC70 (Enzo, ADI-SPA-818).

BiFC assay

BiFC assay was performed as described (21). The constructs of cCFP-SMA1, nVENUS-DCL1, nVENUS-SE,
nVENUS-HYLI and nVENUS-AGO1 were transformed into Agrobacterium strains. All these transgenes are under the control of the cauliflower mosaic virus 35S promoter (35S promoter). The transformed Agrobacterium strains were then collected and suspended in infiltration buffer (10 mM MES pH 5.6, 150 mM acetosyringone and 10 mM MgCl2). Paired C-terminal fragment of cyan fluorescent protein (eCFP) and nVENUS fusion proteins were co-expressed in Nicotiana benthamiana leaves. After 40 h expression, a confocal microscopy (Fluoview 500 workstation; Olympus) was used to detect YFP and chlorophyll auto fluorescence signals at 488 nm with a narrow barrier (505–525 nm, BA505–525; Olympus).

Co-IP assay
To examine the interaction of SMA1 with DCL1 and SE, GFP-SMA1 was transiently co-expressed with MYC-DCL1 or MYC-SE in N. benthamiana as described (26). The expression of these transgenes was directed by the 35S promoter. Total protein was extracted with extraction buffer (50 mM Tris–HCl 8.0, 150 mM NaCl, 5% glycerol, 5% Triton X-100, 1 mM EDTA, 1× complete protease inhibitor cocktail and 1 mM phenylmethylsulfonyl fluoride). Immunoprecipitation (IP) was performed on protein extracts using anti-GFP antibodies (GTA020, Bulldog Bio) coupled to protein G agarose beads. After IP, proteins were separated on a 10% SDS-PAGE and detected with western blot using monoclonal antibodies against GFP (Biolegend, 902602) or MYC (Millipore, 06–340).

ChIP assay
Chromatin immunoprecipitation assay (ChIP) was performed using 14-d-old seedlings of Col-0, sma1-1 and sma1-1 containing 35S::GFP or 35S::GFP-SMA1 as described (19). Three biological replicates were performed. Antibodies of C-terminal domain (CTD) (ab817, Abcam), the Ser5P form of Pol II CTD (ab5408, Abcam), the Ser2P form of Pol II CTD (ab5095, Abcam) or GFP (GTA020, Bulldog Bio) were used for IP. Quantitative PCR (qPCR) was performed using primers listed in Supplementary Table S1.

RNA isolation, northern blot and qRT-PCR analyses
Total RNA was extracted with Trizol reagent. Northern blot was performed as described (26). Approximately 15 μg total RNAs extracted from inflorescences were resolved on 16% PAGE gel and transferred to nylon membranes. 32P-labeled antisense DNA oligonucleotides were used to detect small RNAs. Radioactive signals were detected with a Phosphor imager and quantified with ImageQuant.

The levels of pri-miRNAs, miRNA target transcripts and GUS mRNA were determined using qRT-PCR. A total of 2 μg total RNAs from inflorescences were used to generate cDNAs using the SuperScript III reverse transcriptase (Invitrogen) and an oligo dT18 primer. cDNAs were then used as templates for qPCR on an iCycler apparatus (Bio-Rad) with the SYBR Green Master Mix (Bio-Rad). The primers used for PCR are listed on Supplementary Table S1.

Small RNA sequencing
Inflorescence tissues of Col-0 and sma1-1 were harvested as two biological replicates and used for small RNA sequencing. The generated sequencing data were analyzed for miRNA abundance according to Ren et al. (26). The dataset was deposited into National Center for Biotechnology Information Gene Expression Omnibus (Col-0 accession numbers: GSM3132081 and GSM3132082; sma1-1 accession numbers: GSM3132083 and GSM3132084).

RESULTS
SMA1 is required for development and miRNA accumulation
From a Col-0 (wild-type, WT) population harboring a proMED8::GUS transgene, we isolated a spontaneous mutant with tiny plant size that we named as small1-1 (sma1-1). sma1-1 was backcrossed to Col-0 for five times to remove the proMED8::GUS transgene and other potential mutations before further investigation. Compared with WT, sma1-1 displayed growth defects with reduced organ size, dwarf and short roots (Figure 1A–F). We suspected that sma1-1 may contain less amounts of miRNAs because its pleiotropic growth defects agree with the role of miRNA in controlling plant development. To test this hypothesis, we compared miRNA levels in sma1-1 with those in Col-0 by Northern blot. All 11 tested miRNAs were decreased in abundance in sma1-1 relative to Col-0 (Figure 1G). qRT-PCR analyses further confirmed that the accumulation of miRNAs was reduced in sma1-1 relative to Col-0 (Figure 1H). These results demonstrate that SMA1 is required for miRNA accumulation. To determine the effect of SMA1 on miRNA expression at global levels, we further analyzed miRNA profile from inflorescences of sma1-1 and WT through illumina deep sequencing. Many miRNAs were reduced in abundance in sma1-1 relative to WT (Figure 1I and Supplementary dataset 1), revealing that SMA1 has a general effect on miRNA accumulation. We also examined the effect of sma1-1 on the accurate cleavage of pri-miRNAs, meaning that sequenced miRNAs or miRNA* should fall within ±2 bases of the annotated mature miRNA(s) or miRNA*(s) positions (26). The ratio of imprecise miRNAs in sma1-1 was similar with that of WT (Supplementary dataset 2), suggesting that SMA1 did not affect the accurate cleavage of pri-miRNAs.

We next tested the effect of SMA1 on the accumulation of miRNA target transcripts, since miRNAs mainly repress the transcript levels of their targets. qRT-PCR analyses showed that relative to Col-0, sma1-1 increased the levels of several examined target transcripts, including CUC1, MYB65, PHV, APS3, PHO2 and SPL10, which are targets of miR164, miR159, miR166, miR395, miR399 and miR156, respectively (Figure 1J). This result is consistent with the decreased levels of miRNAs in sma1-1.

SMA1 encodes a pre-mRNA splicing factor
To identify how SMA1 affects miRNA accumulation, we used a map-based cloning approach to identify the sma1-1 mutation. The mutation was mapped to a 31-kb region of chromosome 2 (Figure 2A). DNA sequencing analysis of
Figure 1. sma1-1 causes pleiotropic developmental defects and reduces the accumulation of miRNAs. (A) Four-week-old plants of Col-0 and sma1-1. (B) Opening flowers of Col-0 and sma1-1 (from left to right). (C) The first-pair of leaves of Col-0 and sma1-1. (D) The mature siliques of Col-0 and sma1-1 (from left to right). (E) Seven-week-old plants of Col-0 and sma1-1. (F) Seven-day-old seedlings of Col-0 and sma1-1. (G) The accumulation of miRNAs in Col-0 and sma1-1 detected by northern blot. The U6 RNA was shown as a loading control. The numbers below the picture indicate the relative amount of small RNAs in sma1-1 and Col-0 and represent mean of three replicates (P < 0.05). In the miR159/miR319 blot, the upper band is miR159 whereas the lower band represents miR319. (H) The levels of miRNAs detected by RT-qPCR. miRNA levels in sma1-1 were normalized to those of U6 RNAs and compared with Col-0 (set as 1). (I) Small RNA sequencing analysis of miRNAs in sma1-1 and Col-0. The miRNA abundance was calculated as reads per million, and a log2-transformed ratio of sma1-1/Col-0 was plotted. Each circle represents one miRNA. Think lines indicate median values. (J) The levels of miRNA target transcripts in Col-0 and sma1-1 detected by RT-qPCR. The levels of miRNA target transcripts were normalized to those of UBIQUITIN 5 (UBQ5) and compared with Col-0 (set as 1). Error bars in (H) and (J): standard deviations (SD) of three replicates (*: P < 0.05 with one-sample t-test).

this region revealed a G-to-T transversion at nucleotide position 1900 in the AT2G33730 gene (Figure 2A). To confirm if this mutation is responsible for the observed phenotypes of sma1-1, the protein-encoding region of AT2G33730 was cloned into pMDC43 vector to generate 35S::GFP-SMA1 construct and was introduced into sma1-1. The expression of the 35S::GFP-SMA1 transgene rescued the growth defects in the sma1-1 mutant, including plant size and height (Figure 2C–E and Supplementary Figure S1). Moreover, in transgenic plants, the levels of miRNA and pri-miRNA were comparable to those in Col (Figure 2G and H). These results demonstrate that AT2G33730 indeed encodes SMA1.

SMA1 is a conserved protein in plants (Supplementary Figure S2). It encodes a peptide of 733 amino acids, which shares considerable similarity (42%) with yeast pre-mRNA splicing factor Prp28 (Supplementary Figure S3). Like Prp28, SMA1 contains a putative DEAD-box domain with nine conserved motifs, which are required for ATPase and RNA unwinding activity (Figure 2B and Supplementary Figure S3) (45). The sma1-1 mutation caused the conversion of an (conserved/invariant) Aspartic acid (D) in the conserved DEAD-box domain to Tyrosine (Y) at the amino acid position 634 (Figure 2B and Supplementary Figure S3).
Identification of $SMA1$ gene and complementation of $sma1-1$. (A) Map-based cloning of the $sma1-1$ mutation and diagram of the $SMA1$ gene structure. The $sma1-1$ mutation was mapped to the region between markers T1B8-12 and T1B8-9 on the chromosome 2, and co-segregated with marker 730-CAPS. ATG: the start codon. TGA: the stop codon. Blackbox: the coding sequence. Whiteboxes: the 5′ or 3′ untranslated regions. The mutationsites of $sma1-1$ and $sma1-2$ (SALK 110742) are shown above the $SMA1$ gene structure. (B) The predicted domains of the $SMA1$ protein. The nine conserved motifs in the DEAD-box domain are indicated. The G-to-T mutation in $sma1-1$ causes a D(Asp)-to-Y(Tyr) change at the amino acid position 634. (C–F) The $35S::GFP-SMA1$ transgene rescues the growth defects of $sma1-1$ in 4-week-old seedlings (C), opening flowers (D), mature siliques (E) and the fifth leaves (F). Scale bars: 1 cm (C), 1 mm (D) and 1 cm (E and F). (G and H) The levels of pri-miRNAs (H) and mature miRNAs (H) in $35S::GFP-SMA1$ transgenic plants were detected by RT-qPCR. The levels of pri-miRNAs and miRNAs were normalized to $UBQ5$ and $U6$, respectively. The numbers represent mean of three replicates relative to Col-0 (set as 1). Error bars in (G) and (H): standard deviations (SD) of three replicates. Means with different letters are significantly different determined by ANOVA ($P < 0.05$).

We also obtained a second $sma1$ allele (SALK 110742, $sma1-2$) with a T-DNA insertion (Supplementary Figure S4B). However, no homozygous $sma1-2$ plant was identified from self-crossed population of $sma1-2$/+. Moreover, the immature siliques from heterozygous plants contained ~25% of aborted seeds (Supplementary Figure S4C), suggesting the arrest of embryo development. We crossed $sma1-1$ with the heterozygous plants of $sma1-2$. The F1 plants with $sma1-1$/+ genotype developed normally as the WT, while F1 plants with $sma1-1$/sma1-2 genotype exhibited the $sma1-1$ phenotypes (Supplementary Figure S4D). These results reveal that $sma1-2$ is likely a null mutation and causes embryo lethality when becoming homozygous. We next used a CRISPR/Cas9 system to generate two additional alleles of $sma1$, $sma1-3$ and $sma1-4$, which contain a nucleotide insertion and a deletion of 32 bp nucleotides at the 3′ end of the $SMA1$ gene, respectively (Supplementary Figure S5A). Compared with WT, these two mutants were defective in development and displayed reduced miRNA abundance (Supplementary Figure S5C), which further support that $SMA1$ functions in miRNA biogenesis.

$SMA1$ regulates $DCL1$ pre-mRNA splicing

Because $SMA1$ is a Prp28 homolog, we reasoned that it might function as a splicing factor to control the expression levels of genes involved in miRNA biogenesis. We therefore examined the transcription levels of the well-known key factors involved in miRNA biogenesis including $CBP20$, $CBP80$, $DCL1$, $DDL$, $HEN1$, $HST$, $HYL1$ and $SE$ through qRT-PCR. The transcript levels of these genes were indistinguishable between Col-0 and $sma1-1$ mutant (Figure 3A). To confirm the result, we examined the protein levels by western blot using the antibodies against DCL1, HYL1, AGO1 and SE. The results showed that the protein levels of HYL1, AGO1, and SE were not changed in $sma1-1$ (Figure
3B). In contrast, DCL1 protein levels were decreased in the smal-1 mutant compared with Col-0 (Figure 3B). We next tested if the splicing of DCL1 pre-mRNAs was impaired in smal-1. Several pairs of primers spanning various introns (Figure 3C) were used to examine the splicing of the DCL1 pre-mRNA through RT-PCR analysis in smal-1 and Col-0. Compared with Col, smal-1 did not show obvious effect on the splicing of the first, second and eighth introns of the DCL1 pre-mRNAs (Figure 3D). However, the retention of the ninth intron was increased in smal-1 relative to Col-0 (Figure 3D), suggesting that SMA1 may regulate the splicing of DCL1 pre-mRNA, which is likely attribute to the decreased DCL1 protein levels in smal-1. The effect of smal-1 on the splicing of DCL1 pre-mRNA suggests that SMA1 may have a general role in splicing of protein-coding genes.

To test this possibility, we examined the splicing of additional 18 pre-mRNAs includingHEN1, HYL1, SE, DDL and others (Supplementary Figure S6A). The result showed that smal-1 increased intron retention of four pre-mRNAs (Supplementary Figure S6B), indicating SMA1 may have a general role in intron splicing.

SMA1 affects pri-miRNA splicing

The defection of DCL1 pre-mRNA splicing in smal-1 suggests that like Prp28 in yeast and animals, SMA1 likely participates in splicing. Thus, we sought to examine if smal-1 affected the splicing of intron-containing pri-miRNAs through RT-PCR using primers spanning the intron region of pri-miRNAs (Figure 4A). Increased intron retention was detected in all five examined pri-miRNAs in smal-1 compared with Col-0 (Figure 4B). This result reveals that SMA1 affects the splicing of some pri-miRNAs.

SMA1 promotes pri-miRNA transcription

Because DCL1 is the enzyme processing pri-miRNAs, we reasoned that reduced DCL1 protein levels may impair pri-miRNA processing, resulting in increased pri-miRNA levels. To test this hypothesis, we examined pri-miRNA levels in smal-1 and Col-0. However, qRT-PCR analyses showed that the levels of eight examined pri-miRNAs were decreased in smal-1 relative to Col (Figure 5A). This result suggests that SMA1 may positively regulate pri-miRNA accumulation. Given the fact that the abundance of intronless pri-miRNAs such as pri-miR159a, pri-miR165a, pri-miR167a, pri-miR172e, pri-miR173 and pri-miR319b was also reduced in smal-1 (Figure 5A), we suspected that SMA1 might have other roles in promoting pri-miRNA accumulation besides splicing. Because pri-miRNA levels are partially determined by Pol II-dependent transcription, we hypothesized that SMA1 may positively regulate MIR transcription. We used a pMIR167a::GUS reporter line, which has been used to evaluate the function of several genes in controlling MIR transcription (19, 25, 29), to test this hypothesis. We crossed the pMIR167a::GUS reporter line to smal-1. In F2 population, WT plants (SMA1/SMA1, SMA1/smal-1) and smal-1 containing the pMIR167a::GUS transgene were selected to examine GUS expression. GUS staining of 5-day-old seedlings and inflorescences showed that GUS protein levels were decreased in smal-1 background compared with those in WT (Figure 5B). RT-PCR analysis of RNAs isolated from inflorescences further confirmed that the GUS transcript levels were reduced in smal-1 (Figure 5C). These results indicate that SMA1 may play a role in promoting MIR transcription. To confirm that smal-1 indeed affects MIR transcription, we examined the occupancy of Pol II at the MIR promoters in smal-1 and Col through the ChiP using an antibody. Af-
Figure 4. Intron-retention analysis of pri-miRNA transcripts. (A) Diagram showing the partial regions of several intron-containing pri-miRNAs. Arrows indicate primers used for PCR. The numbers in red indicate the length of PCR products without intron splicing. The numbers in black show the intron length. (B) Intron-retention of pri-miRNAs analyzed by RT-PCR. Upper arrows indicate the unspliced transcripts. Lower arrows indicate the intron-spliced transcripts. UBQ5 served as the loading control. Three replicates were performed and one representative figure was shown for each PCR amplification. The numbers shown below the figures indicate the ratio of spliced transcripts to total transcripts.

between AT2G17470 and AT2G17460 was not detected in SMA1 IPs. Taken together, our results suggest that SMA1 may have a direct role in promoting MIR transcription. We also examined if SMA1 could bind protein-coding genes and found that SMA1 occupied at the promoter of one out of five examined genes (Supplementary Figure S7C), suggesting that SMA1 may also recognize some protein coding genes.

SMA1 associates with the DCL1 complex

Recent studies have revealed that several splicing-related proteins such as MAC3, CDC5 and PRL1 associate with the DCL1 complex to modulate its activity (21,25,49). By analogy, we sought to test if SMA1 interacts with the DCL1 complex. We performed a BiFC assay to determine the interaction of SMA1 with DCL1, SE and HYL1. In this assay, SMA1 protein was fused with the cCFP at its C-terminus (cCFP-SMA1), while the DCL1, SE and HYL1 proteins were fused to nVENUS at their C-terminus (nVENUS-DCL1, nVENUS-SE and nVENUS-HYL1, respectively). After 2 days co-infiltration of paired proteins (cCFP-SMA1/nVENUS-DCL1, cCFP-SMA1/nVENUS-SE and cCFP-SMA1/nVENUS-HYL1), yellow fluorescence (YFP, showing in green) signals were examined with confocal microscopy. Co-expression of cCFP-SMA1 with nVENUS-DCL1 or nVENUS-SE produced strong YFP signals (Figure 6A). In contrast, the combination of cCFP-SMA1 and nVENUS-HYL1 or the negative control nVENUS-AGO1 did not generate fluorescence signal (Figure 6A). These results suggest a potential interaction of SMA1 with the DCL1 complex. We next performed a co-IP assay to confirm the SMA1-DCL1 and SMA1-SE interactions. In this experiment, we transiently co-expressed GFP-
Figure 5. SMA1 is required for MIR transcription. (A) Pri-miRNA levels in Col-0 and sma1-1 detected by RT-qPCR. Pri-miRNA levels in sma1-1 were normalized to those of UBQ5 and compared with Col-0 (set as 1). (B) Histochemical GUS staining of five-day-old seedlings and inflorescences in Col-0 and sma1-1 harboring the pMIR167a::GUS transgene. Fifteen plants containing GUS were analyzed for each genotype and one image for each genotype is shown. (C) GUS transcript levels in Col-0 and sma1-1 harboring the pMIR167a::GUS transgene. GUS transcript levels detected by RT-qPCR were normalized to those of UBQ5 and compared with Col-0 (set as 1). (D) The occupancy of Pol II at MIR promoters in Col-0 and sma1-1 detected by ChIP followed by qPCR. The intergenic region between At2g17470 and At2g17460 (POL II C1) was used as a negative control. IP was performed using anti-RPB1 antibodies. (E) The association of GFP and GFP-SMA1 with MIR promoters in transgenic plants containing 35S::GFP or 35S::GFP-SMA1. The occupancy of GFP-SMA1 and GFP at MIR promoters was detected by qPCR following ChIP. The intergenic region between At2g17470 and At2g17460 (POL II C1) was used as a negative control. IP was performed using anti-GFP antibodies. Error bars (A–E) indicate standard deviations (SD) of three replicates. * in A and C: P < 0.05 with one-sample t-test. Means with different letters are significantly different determined by ANOVA (P < 0.05).

SMa1/MYC-DCL1, GFP-SMA1/MYC-SE, GFP/MYC-DCL1 or GFP/MYC-SE in N. benthamiana, and performed IP using anti-GFP antibodies. After IP, both DCL1 and SE were detected in the GFP-SMA1 precipitates but not in the GFP precipitates (Figure 6B and C). In addition, RNase A treatment did not disrupt the associations of SMA1 with DCL1 and SE (Figure 6B and C), suggesting that their interactions are independent of RNAs. The association of SMA1 with the DCL1 complex led us to test if sma1-1 affected the assembly of the DCL1 complex (D-body). We crossed a transgenic line harboring the HYL1-YFP, which has been used as a marker for D-body (20,50,51), into sma1-1 and examined the numbers of D-bodies. sma1-1 did not show obvious effect on the formation of D-bodies (Supplementary Figure S8).

DISCUSSION

Prp28 is a conserved protein in eukaryotes. Its roles in splicing have been documented in yeast and human (52,53). In yeast and human, PRP28 promotes the transition from a pre-catalytic spliceosome to an activated spliceosome by displacing U1 snRNP from the 5'-splice site (54,55). However, whether Prp28 has functions other than splicing is not clear. Moreover, little is known about the activity of plant Prp28 orthologs due to the lack of corresponding mutants. In this study, we find that sma1-2, a null mutant of the Arabidopsis ortholog of Prp28, is embryo lethal and that a hypomorphic sma1-1 mutant displays pleiotropic developmental defects. These results suggest that SMA1 is essential for plant development. In sma1-1, the accumulation of miRNAs is reduced. This observation together with the fact that SMA1 binds the MIR loci and interacts with the DCL1 complex suggests that SMA1 is a player in miRNA biogenesis. The impaired miRNA biogenesis shall partially contribute to the developmental defects of sma1-1. However, the roles of SMA1 in splicing and transcription shall also be important for plant development.

How does SMA1 function in miRNA biogenesis? In sma1-1, pri-miRNA accumulation is reduced. One possible explanation is that SMA1 affects pri-miRNA splicing. This opinion is evidenced by the fact that the splicing of some pri-miRNAs is impaired in sma1-1. Recent studies have suggested the crosstalk between pri-miRNA processing and splicing (56,57). Some slicing related factors are required for miRNA biogenesis (27,34,37,58–61). Splicing itself also regulates the processing of some pri-miRNAs (62–71). The observation that the accumulation of two miR-
Figure 6. SMA1 associates with the DCL1 complex. (A) The interaction of SMA1 with DCL1, SE, HYL1 and AGO1 detected by BiFC. Paired cCFP- and nVenus-fusion proteins were co-expressed in Nicotiana benthamiana leaves. The combination of cCFP-SMA1 and nVenus-AGO1 was used as a negative control. Green: BiFC signal (originally yellow fluorescence). Red: autofluorescence of chlorophyll. (B and C) Co-IP between SMA1 and DCL1 (C), SMA1 and SE (D). IP was performed using anti-GFP antibodies. Proteins of MYC-DCL1, MYC-SE, GFP-SMA1 and GFP were detected by immunoblot with anti-MYC or anti-GFP antibodies. Inputs show the total protein before IP. RNase A was used to digest RNAs.

NAs with defective pri-miRNA splicing is reduced in smal-1 suggests that SMA1 may promote pri-miRNA accumulation by modulating splicing. However, SMA1 must have other roles in promoting pri-miRNA levels, because some pri-miRNAs do not show splicing defects in smal-1 and the intronless pri-miRNAs also show reduced abundance in smal-1. Indeed, smal-1 impaired the MIR promoter activities (Figure 7). Moreover, SMA1 is required for Pol II occupancy at the MIR loci. These results reveal that SMA1 may function as a positive regulator of MIR transcription (Figure 7). It is not clear how SMA1 influences MIR transcription, as we did not observe the interaction of SMA1 with Pol II. Since SMA1 is a helicase, we speculate that it may function in transcription by modulating chromatin structures to facilitate the access of Pol II or other transcription factors. Consistent with this notion, SMA1 interacts with the MIR loci and the smal-1 mutation within the helicase domain reduces MIR transcription.

SMA1 also enhances the accumulation of the DCL1 protein (Figure 7). The retention of the ninth intron of DCL1 pre-mRNA transcript is increased in smal-1, suggesting that SMA1 controls DCL1 expression through promoting proper DCL1 pre-mRNA splicing (Figure 7). Since DCL1 is the enzyme catalyzing pri-miRNA processing, we propose that SMA1 indirectly influences pri-miRNA processing through modulating DCL1 protein levels. It is not clear why SMA1 only affects the splicing the ninth intron of DCL1. An explanation is that SMA1 may regulate some specific introns. Alternatively, this maybe due to that smal-1 is a weak allele.

Moreover, SMA1 associates with the DCL1 complex. In metazoans, several DEAD-box RNA helicases including p68 and p72 function as accessory components of the Drosa complex to enhance miRNA production (72–74). We envision that the interaction of SMA1 with the DCL1 complex may also facilitate pri-miRNA processing (Figure 7). Besides SMA1, several splicing related factors such as CDC5, MAC7 and MAC3A/3B also promote MIR transcription and function as accessory components of the DCL1 complex (21,29,49). Given the fact that pri-miRNAs are co-transcriptionally processed by DCL1, we suspect that a common theme for these proteins is to bridge the transcription, splicing and processing of pri-miRNAs.

In summary, we identify SMA1 as a new player in miRNA biogenesis. It positively influences miRNA accumulation through enhancing pri-miRNA transcription and/or splicing, modulating the DCL1 protein levels and facilitating the DCL1 activities (Figure 7). However, we cannot completely rule out the possibility that the observed miRNA biogenesis defects maybe caused by the developmental roadblock of smal that does not allow progression of the normal gene expression program. The splicing of some pri-miRNAs and some pre-mRNAs is defective in smal-1, suggesting that SMA1 plays essential roles in splic-
Figure 7. A model proposing the multiple roles of SMA1 in miRNA biogenesis and mRNA metabolism. SMA1 associates with the MIR promoters to promote Pol II-dependent transcription. In addition, it can modulate the splicing of some pri-miRNAs and may function as an accessory factor of the DCL1 complex to facilitate pri-miRNA processing. Moreover, SMA1 influences pri-miRNA processing through enhancing the proper splicing of the DCL1 pre-mRNAs. SMA1 also promotes the splicing of some protein-coding genes including DCL1, and, by analogy, may regulate the transcription of some protein-coding genes.

ing of a subset of Pol-II dependent RNAs. Notably, we also uncovered a role of SMA1 in MIR transcription. Since MIRs share similar gene structures with other Pol II dependent genes, we propose that SMA1 may affect the transcription of a subset of protein-coding genes. Clearly, the availability of sma1-1 provides an opportunity to further dissect the role of SMA1 in miRNA biogenesis, splicing and transcription in the near future. The discovery that SMA1 functions in miRNA biogenesis and transcription has expanded the knowledge on the function of Prp28. It will be worth to test if Prp28 has the similar roles in other organisms.

DATA AVAILABILITY
Multiple sequence alignment was performed with Clustal X (75) (http://www.clustal.org). Phylogenetic tree was constructed using MEGA 7 (76) (http://www.megasoftware.net). The statistically significant differences were analyzed with STATA software (https://www.stata.com).

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We acknowledge the NASC for providing the sma1-2 (SALK_110742) seeds.

FUNDING
Nebraska Soybean Board [1727 to B.Y.]; National Science Foundation [MCB1808182 to B.Y.]; National Natural Science Foundation of China [31425004, 91417304, 91017014 to Y.L.; 31400249 to R.X.]; Ministry of Agriculture of China [2016ZX08009-003 to N.L.]; Strategic Priority Research Program ‘Molecular Mechanism of Plant Growth and Development’ [XDBP401 to Y.L.]; Pioneer Hundred Program of Chinese Academy of Sciences (to S.L.). Funding for open access charge: Pioneer Hundred Program of Chinese Academy of Sciences.

Conflict of interest statement. None declared.

REFERENCES
1. Chen,X. (2009) Small RNAs and their roles in plant development. Annu. Rev. Cell Dev. Biol., 25, 21–44.
2. Axtell,M.J., Snyder,J.A. and Bartel,D.P. (2007) Common functions for diverse small RNAs of land plants. Plant Cell, 19, 1750–1769.
3. Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell, 116, 281–297.
4. Li,S., Castillo-Gonzalez,C., Yu,B. and Zhang,X. (2017) The functions of plant small RNAs in development and stress responses. Plant J., 90, 654–670.
5. Xie,Z., Allen,E., Fahlgren,N., Calamar,A., Givan,S.A. and Carrington,J.C. (2005) Expression of Arabidopsis MIRNA genes. Plant Physiol., 138, 2145–2154.
6. Baulcombe,D. (2004) RNA silencing in plants. Nature, 431, 356–363.
7. Kurihara,Y. and Watanabe,Y. (2006) The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA, 12, 206–212.
12. Yu,B., Yang,Z., Li,J., Minakshina,S., Yang,M., Padgett,R.W., Steward,R. and Chen,X. (2005) Methylation as a crucial step in plant microRNA biogenesis. Science, 307, 912–915.

13. Zhai,J., Zhao,Y., Simon,S.A., Huang,S., Petsch,K., Arikit,S., Pillay,M., Ji,L., Xie,M., Cao,X. et al. (2013) Plant microRNAs display differential 3′ truncation and tailing modifications that are ARGO U A N E T I dependent and conserved across species. Plant Cell, 25, 2417–2428.

14. Ren,G., Xie,M., Zhang,S., Vinovskis,C., Chen,X. and Yu,B. (2014) Methylation protects microRNAs from an AGO1-associated activity that uridylates 5′ RNA fragments generated by AGO1 cleavage. PNAS, 111, 6365–6370.

15. Park,M.Y., Wu,G., Gonzalez-Salsero,A., Vaucheret,H. and Poethig,R.S. (2005) Nuclear processing and export of microRNAs in Arabidopsis. PNAS, 102, 3691–3696.

16. Earley,K.W. and Poethig,R.S. (2011) Binding of the cyclophilin 40 ortholog SQUINT to Hsp90 protein is required for SQUINT function in Arabidopsis. J. Biol. Chem., 286, 38184–38189.

17. Cui,Y., Fang,X. and Qi,Y. (2015) Transcription and processing of primary microRNAs are coupled byelongon complex in Arabidopsis. Nat. Plants, 1, 15075.

18. Hajheidari,M., Farrona,S., Huettel,B., Koncz,C. and Koncz,C. (2012) CDC5, a DNA binding protein, positively regulates posttranscriptional processing and/or transcription of primary microRNA transcripts. PNAS, 110, 17588–17593.

19. Fang,X., Cui,Y., Li,Y. and Qi,Y. (2015) Transcription and processing of primary microRNAs are coupled by elongon complex in Arabidopsis. Nat. Plants, 1, 15075.

20. Wang,X., Li,J., Li,X., Kong,Y., Ren,G. and Yu,B. (2013) Regulation of miRNA abundance by RNA binding protein TOUGH in Arabidopsis. PNAS, 109, 12817–12821.

21. Francisco-Manglet,G.A., Carlsson,P., Kim,M.H., Eo,H.J., Oh,S.A., Kim,J.H., Kulcheski,F.R., Park,S.K. and Manavella,P.A. (2015) THO2, a core member of the THO/TREX complex, is required for microRNA production in Arabidopsis. Plant J., 82, 1018–1029.

22. Li,S., Liu,K., Zhang,S., Wang,X., Rogers,K., Ren,G., Zhang,C. and Yu,B. (2017) Regulation of miRNA abundance by RNA binding protein TOUGH in Arabidopsis. PNAS, 109, 12817–12821.

23. Bowman,E.A. and Kelly,W.G. (2014) RNA polymerase II CTD in Arabidopsis. Trends Plant Sci., 19, 633–643.

24. Chen,C., Yu,B., Ren,G. and Yu,B. (2018) Mac3A and Mac3B, two core subunits of the MOS4-associated complex, positively influence miRNA biogenesis. Plant Cell, 30, 481–494.

25. Wu,X., Shi,Y., Li,J., Xu,L., Fang,Y. and Qi,Y. (2013) A role for the RNA-binding protein MOS2 in microRNA maturation in Arabidopsis. Cell Res., 23, 645–657.

26. Li,S., Liu,K., Zhou,B., Liu,M., Zhang,S., Zeng,L., Zhang,C. and Yu,B. (2018) Mac3A and Mac3B, two core subunits of the MOS4-associated complex, positively influence miRNA biogenesis. Plant Cell, 30, 481–494.

27. Straus,E.J. and Guthrie,C. (1997) A DEAD-box protein, required for the first step of mRNA splicing in vitro. Nucleic Acids Res., 25, 3187–3193.

28. Konishi,T., Udozome,N. and Sugimoto,A. (2006) The Caenorhabditis elegans DDX-23, a homolog of yeast splicing factor Prp28, is required for the sperm-oocyte switch and differentiation of various cell types. Dev. Dyn., 237, 2367–2377.

29. Wang,Z.P., Xing,H.L., Dong,L., Zhang,Y.H., Han,C.Y., Wang,X.C. and Chen,Q.J. (2015) Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in Arabidopsis in a single generation. Genome Biol., 16, 144.

30. Li,S., Liu,Y., Zheng,L., Chen,L., Li,N., Corke,F., Fu,X., Zhu,Z., Bevan,M.W. et al. (2012) The plant-specific G protein subunit AGG3 influences organ size and shape in Arabidopsis thaliana. Nat. Plant, 194, 690–703.

31. Mollmann,S., Mathew,R., Neumann,P., Schmitt,A., Luhrmann,R. and Ficner,R. (2014) Structural and functional analysis of the human spliceosomal DEAD-box helicase Prp28. EMBO J., 33, 1626–1642.

32. Zhang,Z., Guo,X., Ge,C., Ma,Z., Jiang,M., Li,T., Koiwa,H., Zhu,J.K., Yang,S.W. and Lee,B.H. (2013) STA1, an Arabidopsis pre-mRNA processing factor 6 homolog, is a new player involved in miRNA biogenesis. Nucleic Acids Res., 41, 1984–1997.
