

Proper regulation of a sperm-specific *cis*-nat-siRNA is essential for double fertilization in *Arabidopsis*

Mily Ron,¹ Monica Alandete Saez,² Leor Eshed Williams,³ Jennifer C. Fletcher,
and Sheila McCormick⁴

Department of Plant and Microbial Biology, Plant Gene Expression Center, U.S. Department of Agriculture/Agricultural Research Service, University of California at Berkeley, Albany, California 94710, USA

Natural *cis*-antisense siRNAs (*cis*-nat-siRNAs) are a recently characterized class of small regulatory RNAs that are widespread in eukaryotes. Despite their abundance, the importance of their regulatory activity is largely unknown. The only functional role for eukaryotic *cis*-nat-siRNAs that has been described to date is in environmental stress responses in plants. Here we demonstrate that *cis*-nat-siRNA-based regulation plays key roles in *Arabidopsis* reproductive function, as it facilitates gametophyte formation and double fertilization, a developmental process of enormous agricultural value. We show that male gametophytic *kokopelli* (*kpl*) mutants display frequent single-fertilization events, and that *KPL* and a inversely transcribed gene, *ARIADNE14* (*ARI14*), which encodes a putative ubiquitin E3 ligase, generate a sperm-specific nat-siRNA pair. In the absence of *KPL*, *ARI14* RNA levels in sperm are increased and fertilization is impaired. Furthermore, *ARI14* transcripts accumulate in several siRNA biogenesis pathway mutants, and overexpression of *ARI14* in sperm phenocopies the reduced seed set of the *kokopelli* mutants. These results extend the regulatory capacity of *cis*-nat-siRNAs to development by identifying a role for *cis*-nat-siRNAs in controlling sperm function during double fertilization.

[*Keywords*: ARIADNE; single fertilization; embryo; endosperm; small RNA; ubiquitin ligase]

Supplemental material is available at <http://www.genesdev.org>.

Received November 5, 2009; revised version accepted March 24, 2010.

Some small RNAs regulate gene expression through post-transcriptional gene silencing, translational inhibition, and heterochromatin modification (Vaucheret 2006). siRNAs, which include *trans*-acting siRNAs (tasiRNAs) and natural *cis*-antisense siRNAs (*cis*-nat-siRNAs), are derived from long, perfectly dsRNAs (Xie and Qi 2008). Although specific siRNAs have been implicated in regulation of development, adaptive responses to biotic and abiotic stresses, genome stability, and transposon movement (Vaucheret 2006), the extent of their regulatory importance remains unclear (Ghildiyal and Zamore 2009).

cis-nat-siRNAs are derived from the overlapping transcripts of two adjacent genes located on opposite strands (Borsani et al. 2005). The overlapping transcripts are thought to form a dsRNA that is cleaved by a DICER endonuclease to produce small 21- to 24-nucleotide (nt) RNAs (Kim 2005). The small RNA then is loaded into

a RNA-induced silencing complex (RISC) and directs cleavage of target RNAs by sequence complementarity (Vazquez 2006). Recent genome-wide analyses demonstrated the widespread existence of *cis*-nat-siRNAs in eukaryotic genomes, which raises the possibility that *cis*-nat-siRNAs could be major effectors of gene regulation (Jin et al. 2008; Ghildiyal and Zamore 2009). Approximately 26% of human genes, 22% of mouse genes, 16% of *Drosophila* genes (Zhang et al. 2006), and 9% of *Arabidopsis* genes (Jen et al. 2005; Jin et al. 2008) are overlapping and may form *cis*-nat-siRNAs. Yet despite their abundance in animal genomes, to date there is no direct evidence showing the gene regulation capacity of *cis*-nat-siRNAs (Okamura and Lai 2008; Werner et al. 2009). In plants, there are two studies demonstrating regulatory roles of *cis*-nat-siRNAs generated from two coding RNAs. In each case, one transcript of the *cis*-nat-siRNA pair was constitutively expressed and the other was induced by stress, either abiotic (Borsani et al. 2005) or biotic (Katiyar-Agarwal et al. 2006). In both cases, down-regulation of the target gene by specific *cis*-nat-siRNAs conferred tolerance to the inductive stress.

In flowering plants, the male gametophyte (pollen grain) contains three cells: a vegetative cell that forms

Present addresses: ¹Department of Plant Biology and Genome Center, University of California at Davis, Davis, CA 95616, USA; ²Public Intellectual Property Resource for Agriculture (PIPRA), University of California at Davis, Davis, CA 95616, USA; ³Faculty of Agriculture, Hebrew University of Jerusalem, Rehovot 76100, Israel.

⁴Corresponding author.

E-MAIL sheilamc@berkeley.edu; FAX (510) 559-5678.

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.1882810>.

the pollen tube, and two sperm enclosed within the vegetative cell cytoplasm (McCormick 2004). The female gametophyte (embryo sac) develops within the ovule and consists of seven cells of four types: three antipodals, two synergids, one egg, and one central cell (Yadegari and Drews 2004). The fertilization process starts when a pollen grain lands on the stigma, rehydrates, and begins germination. The pollen tube grows through female tissue until it reaches an available ovule, enters one of the synergids, and releases the sperm. Double fertilization occurs when one sperm fuses with the egg to produce the zygote and the other sperm fuses with the central cell to form the primary endosperm cell. Double fertilization is one of the most important developmental processes in higher plants, because the embryo and endosperm constitute the edible parts of many seeds and grains used in human and animal nutrition. Understanding the fertilization process therefore has great relevance for genetic engineering aimed at improving the nutritional quality of crop plants. Because they have condensed chromatin, sperm were thought to be relatively transcriptionally quiescent, but several studies showed that sperm express a distinct and diverse set of genes (Engel et al. 2003; Okada et al. 2007; Borges et al. 2008).

Kokopelli is a Native American male fertility deity. Here we describe analyses of the *kokopelli* (*kpl*) mutant of *Arabidopsis thaliana* that implicate a regulatory role for cis-nat-siRNAs in sperm function. We show that *KPL* and an inversely transcribed gene, *ARIADNE14* (*ARI14*), which encodes a putative ubiquitin E3 ligase, have in-

verse RNA expression patterns, and that, in the absence of *KPL* transcripts, *ARI14* RNA levels in sperm were increased and fertilization was impaired. We show that *KPL* and *ARI14* represent a sperm-specific cis-nat-siRNA pair. *ARI14* is transcribed both in the vegetative cell and in sperm, but a *KPL-ARI14*-nat-siRNA is generated only in sperm, resulting in depletion of *ARI14* in sperm. *kpl* plants showed frequent single-fertilization events, in which only the egg or the central cell was fertilized, causing seed abortion. We also show that overexpression of *ARI14* in sperm phenocopies the reduced seed set of *kpl* plants. Our data demonstrate that cell-specific nat-siRNA activity is required for fertilization, thus showing that nat-siRNAs can spatially regulate a key developmental process.

Results

T-DNA insertions in KOKOPELLI affect the male gametophyte

The *kokopelli-1* (*kpl-1*) allele has a T-DNA insertion in the third exon of *At5g63720* (Fig. 1A). Plants heterozygous for *kpl-1* showed atypical segregation of the T-DNA-linked kanamycin resistance gene in self-pollinated progeny (Supplemental Table S1), which caught our interest. The Kan^R to Kan^S segregation ratio of 1.6:1 deviates significantly from the expected 3:1 ratio for a recessive sporophytic mutation ($\chi^2 = 57.6$, $P < 0.0001$) and from the expected 1:1 ratio for a gametophytic mutation ($\chi^2 = 30.9$,

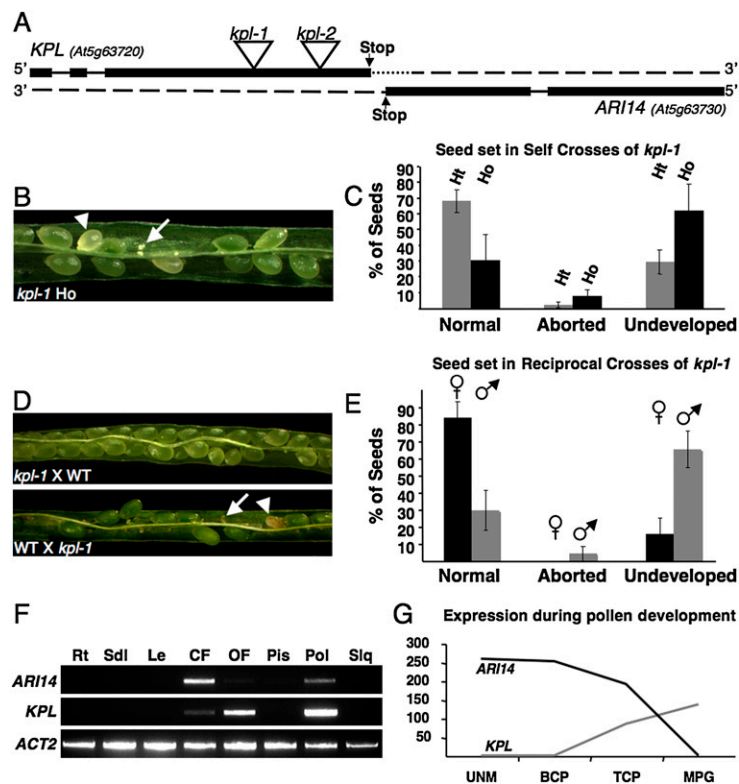


Figure 1. *kokopelli* is a male gametophytic mutant that shows reduced seed set. (A) Schematic representation of the *KPL* and *ARI14* genomic region. Thick and thin solid lines represent exons and introns, respectively. Dotted lines represent 3' UTRs. Dashed lines represent the complementary DNA strand. Positions of T-DNA insertions within the *kpl-1* and *kpl-2* alleles are marked with triangles. (B) Representative silique from a self-pollinated *kpl-1/kpl-1* plant showing a reduced seed set. The arrow points to an undeveloped ovule, and the arrowhead points to an aborted seed. (C) Percentage of normal seeds, aborted seeds, and undeveloped ovules from self-pollinated heterozygous (Ht, gray bars, $n = 1061$) or homozygous (Ho, black bars, $n = 1547$) *kpl-1* plants. Error bars represent standard deviation from the mean. (D) Siliques of reciprocal crosses of *kpl-1/kpl-1* plants with wild type (WS4). The arrow points to an undeveloped ovule, and the arrowhead points to an aborted seed. (E) Percentage of normal seeds, aborted seeds, or undeveloped ovules in reciprocal crosses with wild type, using the *kpl-1/kpl-1* plants as either male (gray bars, $n = 1385$) or female (black bars, $n = 293$). Error bars represent standard deviation from the mean. (F) RT-PCR profile of *KPL* expression in various tissues. (Rt) Roots; (Sdl) 14-d-old seedlings; (Le) rosette leaves; (CF) closed flowers; (OF) open flowers; (Pis) unpollinated pistils; (Pol) pollen; (Slq) siliques. (G) *KPL* and *ARI14* expression profiles throughout male gametophyte development. Microarray data was adapted from Honys and Twell (2004). (UNM) Microspores; (BCP) bicellular pollen; (TCP) tricellular pollen; (MPG) mature pollen.

$P < 0.0001$), indicating abnormal transmission through one or both parents. The reduced seed set seen in self-pollinated plants containing *kpl-1* was variable from plant to plant and from silique to silique, but, on average, *kpl-1/+* plants had ~32% reduced seed set (aborted, $2.4\% \pm 1.9\%$; undeveloped, $29.5\% \pm 7.5\%$), while *kpl-1/kpl-1* plants had ~70% reduced seed set (aborted, $7.7\% \pm 4.2\%$; undeveloped, $61.9\% \pm 17.1\%$) (Fig. 1B,C). Reciprocal crosses between *kpl/+* or *kpl/kpl* with wild type showed reduced transmission through the male (Fig. 1D,E; Supplemental Table S1), indicating that the reduced seed set was due to a defect in the male gametophyte. Note that the *kpl-1* defect is not fully penetrant, since homozygous plants were obtained.

To determine the step causing reduced *kpl-1* transmission, we checked for morphological abnormalities in *kpl-1* pollen. Mature *kpl-1* pollen grains were tricellular when stained with DAPI (Supplemental Fig. S1C), indicating that male gametophyte development was not affected, and *kpl-1* pollen showed no differences from wild-type pollen (Supplemental Fig. S1A,B) when germinated in vitro (Boavida and McCormick 2007). In addition, when we examined young siliques (2–3 d after pollination, DAP) of *kpl-1/kpl-1* plants harboring the *LAT52:GUS* pollen-specific transgene (Twell et al. 1990), we detected GUS in almost all ovules (92%, $n = 531$), indicating that pollen tube guidance was not affected (Supplemental Fig. S1D). Because *kpl-1* pollen development, germination, elongation, and ovule penetration appeared normal, the defect causing the reduced seed set likely occurs during later fertilization events. To confirm the involvement of *KPL*, we obtained a second T-DNA allele, *kpl-2* (Fig. 1A), which also had reduced seed set. On average, *kpl-2/+* plants had ~27% reduced seed set (aborted, $7.2\% \pm 3.5\%$; undeveloped, $20.2\% \pm 8.3\%$), while *kpl-2/kpl-2* plants had ~45% reduced seed set (aborted, $15.4\% \pm 5.9\%$; undeveloped, $29.9\% \pm 4.9\%$) (Supplemental Table S1; Supplemental Fig. S2).

Undeveloped ovules of kpl plants show frequent single-fertilization events

Two different phenotypes together account for the reduced seed set in siliques of self-pollinated *kpl-1/+* and *kpl-1/kpl-1* plants: undeveloped ovules and aborted seeds. To further characterize the undeveloped ovules, we examined 1- to 2-DAP pistils of self-pollinated *kpl-1/kpl-1* plants. We observed ovules that had not been fertilized (Fig. 2A), and ovules that had been fertilized and were developing correctly, with a developing embryo and proliferating endosperm (Fig. 2B). Surprisingly, we also observed ovules in which only one fertilization event had occurred: Either the egg had been fertilized and a developing embryo was evident, but the central cell was unfertilized (Fig. 2C), or the central cell had been fertilized and endosperm nuclear divisions were evident, but the egg was unfertilized (Fig. 2D). Ovules from the *kpl-2/kpl-2* mutant displayed similar phenotypes (Supplemental Fig. S3).

To assess whether both sperm were delivered to ovules, we introgressed the reporter construct *HTR10pro:HTR10-*

mRFP, which is expressed specifically in sperm nuclei (Berger et al. 2008), into *kpl-1/kpl-1* plants. At 1–2 DAP, in ovules receiving wild-type pollen, most will have been double-fertilized, and therefore no RFP signal from sperm nuclei should be observed at the micropylar pole. In wild-type plants harboring *HTR10pro:HTR10-mRFP*, only 14 of 416 (3.3%) ovules had two sperm nuclei visible (Supplemental Fig. S4A–D). These presumably reflect failed fertilization events, because this incidence is similar to the incidence of undeveloped ovules in wild-type plants. In self-pollinated *kpl-1* plants, 59 of 312 ovules (18.9%) had either one or two sperm nuclei visible in the region of the degenerated synergid (Fig. 2E,G, respectively), although all the pollen grains from these plants contained two sperm (Supplemental Fig. S4E). In some ovules in which only one sperm nucleus was detected, there was also a developing embryo (Fig. 2F), indicating that *kpl-1/kpl-1* pollen had delivered two sperm to the embryo sac but double fertilization was impaired. These ratios are different from the expected ~65% ovules with one or no fertilization events, due to technical difficulties in detecting the red sperm nuclei; nonetheless, the difference from wild type is significant ($\chi^2 = 238.2$, $P < 0.0001$).

To confirm that the embryo or endosperm in ovules with a single-fertilization event resulted from fertilization and were not spontaneous, we crossed *kpl-1* with a plant carrying a *FAC1pro:GUS* reporter. *FAC1* (*Embryonic Factor 1*) is expressed from the paternal allele in both the developing embryo and endosperm (Xu et al. 2005). We used pollen from a *kpl-1/+; FAC1pro:GUS/+* F₁ plant to pollinate wild-type plants, and assessed expression of *FAC1pro:GUS* in ovules at 2.5 DAP. We detected three different GUS expression patterns. One had GUS in both the embryo and endosperm, indicating that the ovules had been doubly fertilized (Fig. 2H), while the other two had GUS in ovules containing either only an embryo (Fig. 2I) or only endosperm (Fig. 2J), indicating that these ovules had indeed been fertilized.

The KPL and ARI14 transcripts have the potential to generate nat-siRNAs in sperm

At5g63720 (*KPL*) is in a reverse orientation to the adjacent gene *At5g63730* (*ARI14*) (Fig. 1A), and the distance between their stop codons is 82 base pairs (bp). The *Arabidopsis* Information Resource (TAIR; <http://www.arabidopsis.org>) annotations for *KPL* and *ARI14* lack 3' untranslated regions (UTRs), explaining their absence in the Jin et al. (2008) list of gene pairs potentially generating nat-siRNA pairs. Using BLAST, we identified ESTs (e.g., AYAOO64TF) representing the 3' end of *KPL*, which overlaps 72 bp with the last exon of *ARI14*. *KPL* has no recognizable protein domain signatures, but contains a putative nuclear localization signal. *ARI14* is a member of the gene family encoding ARIADNE 1-like proteins, a family of putative ubiquitin E3 ligases with RING finger domains (Mladek et al. 2003). When we examined *KPL* and *ARI14* expression in different tissues using RT-PCR, both were detected only in closed flower buds, open flowers, and pollen (Fig. 1F). In addition, *KPL* and *ARI14*

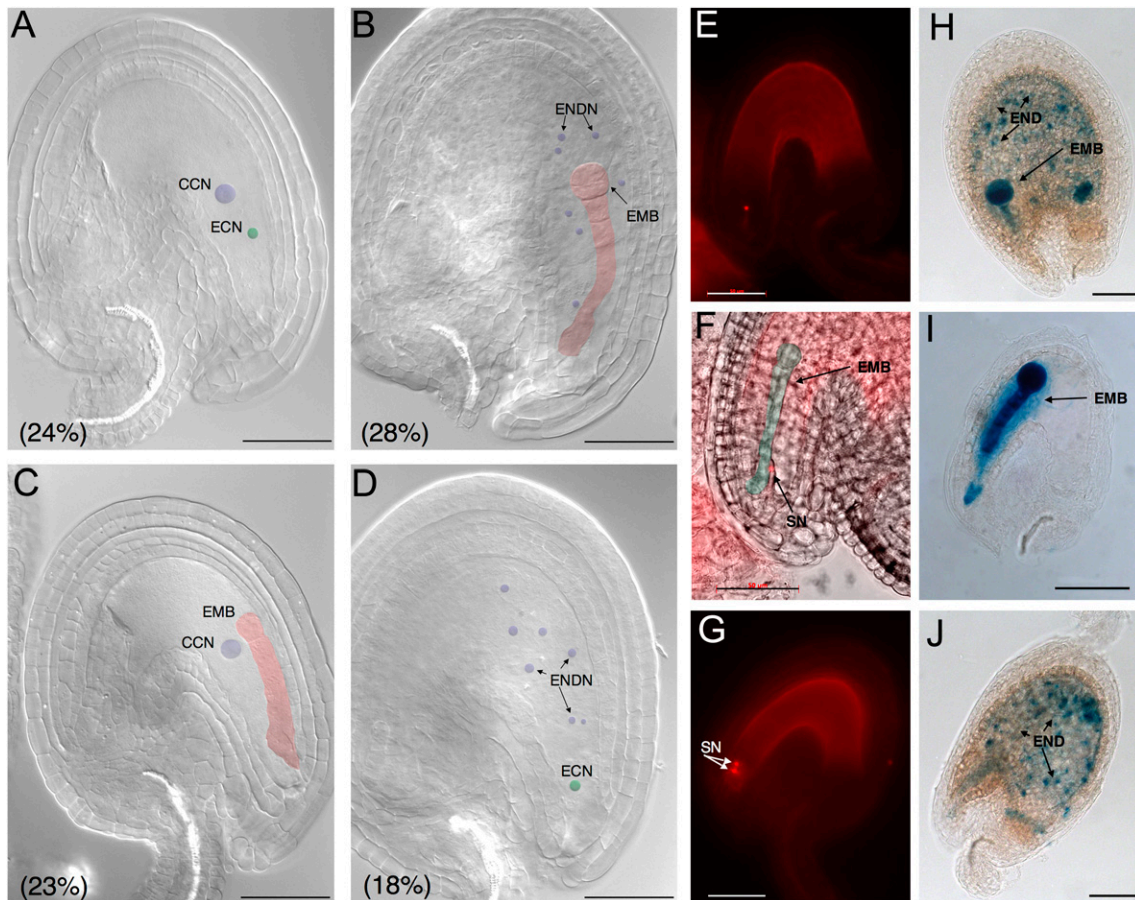


Figure 2. Defects in ovules receiving *kpl-1* pollen. (A–D) Siliques from a *kpl-1/kpl-1* plant were dissected 1–2 DAP, and then the ovules were cleared and examined using differential interference contrast (DIC) microscopy. The frequency of each phenotype is given in parentheses ($n = 764$ ovules; an additional 7% had been double-fertilized, but either the embryo or endosperm had arrested). Bars, 50 μm . (A) Unfertilized ovule with central cell nucleus (CCN; false-colored blue) and egg cell nucleus (ECN; false-colored green). (B) Fertilized ovule with embryo (EMB; false-colored red) and developing endosperm (endosperm nuclei; ENDN, false-colored blue). (C) Representative ovule with a two-celled embryo and unfertilized central cell. (D) Representative ovule with endosperm nuclei and an unfertilized egg cell. (E–G) Undeveloped ovules from a self-fertilized *kpl-1/kpl-1* plant harboring the *HTR10pro:HTR10-mRFP* transgene. Sperm nuclei (SN) are bright red dots. Embryo sacs are red due to autofluorescence. Bars, 50 μm . (E) Fluorescent image of an ovule with one sperm (SN). (F) Overlay of fluorescent and bright-field images of the ovule in E, showing a developing embryo (EMB, false-colored green) and the sperm nucleus (SN). (G) Fluorescent image of an ovule with two sperm nuclei (SN). (H–J) Patterns of GUS expression observed after wild-type plants were pollinated using *kpl* plants carrying *FAC1pro:GUS* as pollen donor. (H) Doubly fertilized ovule. (I) Embryo (EMB) only. (J) Endosperm (END) only.

expression throughout male gametophyte development (Hony and Twell 2004) showed inverse patterns (Fig. 1G). According to microarray data (Borges et al. 2008), *ARI14* transcripts are called “present” in pollen and “absent” in sperm, while *KPL* transcripts are sperm-enriched; they are called “present” in both pollen and sperm, but the signal in pollen is coming from the sperm. Given these observations, we hypothesized that *KPL* and *ARI14* generate a pair of *cis*-nat-siRNA, and that the transcripts of *KPL* post-transcriptionally down-regulate the level of *ARI14* in sperm.

KPL and *ARI14* expression patterns

To test our hypothesis, we used RT-PCR and quantitative real-time PCR (qPCR) assays to determine relative *KPL*

and *ARI14* expression levels. *KPL* transcripts were readily detected in wild-type open flowers, but the level of *ARI14* transcripts was very low (Fig. 3A,B). In homozygous *kpl* mutants, the level of *KPL* transcripts was reduced, while the level of *ARI14* transcripts was substantially increased (Fig. 3A,B), suggesting inverse regulation of the two transcripts. To test their cell specificity, we generated transgenic reporter lines harboring constructs in which expression of the tdTomato fluorescent protein was driven by either the *KPL* or *ARI14* promoters. In all lines examined, the *KPL* promoter was active specifically in sperm (Fig. 3C), while the *ARI14* promoter was active in both the vegetative cell and the sperm (Fig. 3C; Supplemental Fig. S5). Because *ARI14* transcripts were called “absent” from sperm (Borges et al. 2008), we determined the expression pattern of a translational fusion, in which

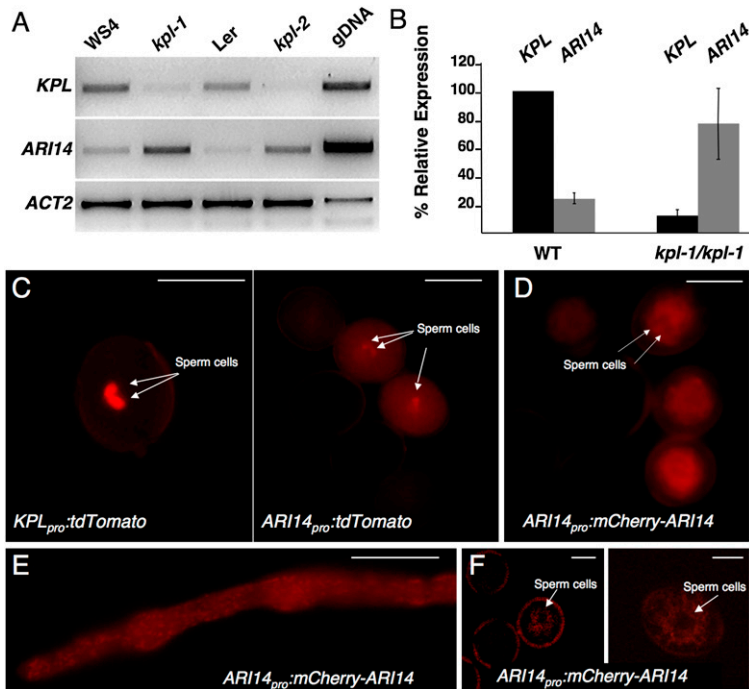


Figure 3. Gene expression of *KPL* and *ARI14* in wild-type and *kpl* pollen. (A,B) Relative expression levels of *KPL* and *ARI14*. Total RNA was extracted from 1-d-old open flowers of *kpl* homozygous mutants and their corresponding wild-type ecotypes. (A) RT-PCR analyses, with *ACT2* used as a control. (B) qPCR analysis. The RNA level of *KPL* in wild type was set at 100%. Each expression level was normalized to that of *IPP2*. Standard deviations were calculated from two biological replicates. (C) Fluorescence microscopy images of pollen grains from transgenic plants harboring promoter fusion constructs of *KPL* and *ARI14* to *tdTomato*. (D–F) Fluorescence microscopy images of pollen from transgenic plants harboring the *ARI14*:*pro:mCherry-ARI14* translational fusion construct. Mature pollen grains are visualized with epifluorescence (D) or CLSM (F), and a germinated pollen tube is visualized with epifluorescence (E). Arrows point to the shadows where the sperm are located. Bars: C–E, 20 μ m; F, 10 μ m.

the entire upstream region and coding sequence of *ARI14* was fused to the mCherry fluorescent protein. In all lines examined, the mCherry-*ARI14* fusion protein was absent from sperm and was in speckles within the vegetative cell cytoplasm (Fig. 3D–F; Supplemental Fig. S6). These results support the hypothesis that *KPL* and *ARI14* mRNAs are both transcribed in sperm, but that *ARI14* transcript levels are unstable in the presence of *KPL* transcripts.

Coexpression of *ARI14* and *KPL* induces *ARI14*-specific small RNAs

To test the hypothesis that *ARI14* is regulated post-transcriptionally, and to be able to detect *ARI14*-specific small RNAs, we generated stable transgenic *Arabidopsis* plants harboring either a *35Spro:GFP-ARI14-30C* fusion construct or a *35Spro:KPL* construct. The expression levels in parental plants containing these constructs were verified by GFP imaging or RT-PCR, respectively. We generated F_1 plants that coexpressed both *GFP-ARI14-30C* and *KPL* and compared their GFP expression levels with those of parental plants that contained only the *GFP-ARI14-30C* fusion (Fig. 4). The F_1 plants had less GFP than the *GFP-ARI14-30C* parent (Fig. 4A,B), indicating that the *GFP-ARI14-30C* fusion protein was reduced. Due to chlorophyll autofluorescence and reduced GFP, these F_1 plants looked red when imaged, and are referred to as “Red” plants (Fig. 4B). F_2 seeds from self-pollinated individual Red plants were doubly selected for seedlings harboring both constructs, and then *KPL* and GFP expression was evaluated by qPCR (Fig. 4C,D, respectively). This analysis confirmed lower GFP expression levels in plants coexpressing both constructs than in plants expressing only *GFP-ARI14-30C*. The reduction of *KPL* expression in Red plants (32%–36%) (Fig. 4C) compared

with the parent is likely explained by a dosage effect of different copy numbers, as seeds from several resistant plants were pooled together and were likely a mix of heterozygous and homozygous plants. Nevertheless, the reduction of GFP expression (68%–82%) (Fig. 4D) was greater and cannot be explained solely by a copy number difference, therefore indicating that the overlapping region of *ARI14* and *KPL* is sufficient to reduce *GFP-ARI14-30C* transcript levels. As a control, we also crossed *35Spro:GFP-ARI14-30C* plants with *35Spro:PIF3* plants. When imaged, the *GFP-ARI14-30C* parent and the F_1 plants showed equivalent fluorescence intensities (Supplemental Fig. S7), indicating that the reduction of GFP expression in the Red plants was due to the presence of the *KPL* transcript and not due to GFP copy number or cosuppression of the two *35Spro*-containing transgenes.

To further characterize the interaction between *GFP-ARI14-30C* and *KPL* transcripts, we isolated total RNA, including small RNA, from pooled F_2 seedlings that contained both constructs. We used a mixture of six *ARI14*-specific oligomers as probes (Supplemental Fig. S8A) and detected a 21-nt RNA species only in F_2 plants that contained both *KPL* and *GFP-ARI14-30C*, while no similar RNA species were detected in parental plants containing only the *GFP-ARI14-30C* fusion (Fig. 4E). These probes are from the *ARI14* sequence and can recognize small RNAs representing the *KPL* sequence, indicating that *ARI14* was being down-regulated post-transcriptionally. We repeated this experiment using subsets of these six oligomers as probes, and were able to delimit the hybridizing sequences to a 60-nt region of *ARI14* (probe set 2 + 3) (Supplemental Fig. S8B), which indicated that there was a specific target in the *ARI14* mRNA.

To test whether the 21-nt species possibly targeted *ARI14* mRNA for endonucleolytic cleavage in wild-type

sperm, we performed a 5' rapid amplification of cDNA ends (RACE) assay using polyadenylated RNA isolated from wild-type pollen to map potential cleavage sites. Among the cleavage sites found in the *ARI14* mRNA, 14 out of 28 mapped to one site at the overlap between the *KPL* 3'UTR and the *ARI14* exon (at position 1455 of the *ARI14* mRNA). This site is within the 60-nt region that detected a 21-nt species (Supplemental Fig. S8A). Two additional cleavage sites were mapped upstream (at positions 1393 and 1203 of *ARI14* mRNA) (Fig. 4F). Furthermore, in a database of small RNAs from wild-type *Arabidopsis* sperm (Slotkin et al. 2009), we identified one 21-nt RNA that could target *ARI14*. This small RNA represents position 1426–1446 of the *KPL* mRNA, and could potentially target cleavage of *ARI14* at position 1647 (Fig. 4F). These positions are not in the predicted overlap between *KPL* and *ARI14* mRNAs, suggesting that the primary siRNAs are generated from a longer overlap between the two transcripts, probably in the pre-mRNA, before polyA addition.

To confirm that this *KPL* small RNA species indeed targeted *ARI14*, we first used 3' RACE to show that the polyA addition site was at position 1604 in both the F₂

Red plants and wild-type pollen (Fig. 4F), at the STOP codon of *KPL*. To detect *ARI14* small RNAs, we used low-molecular-weight-enriched RNA from wild-type pollen and Red plants, probes based on the 21-nt RNA sequence from the sperm small RNA library (probe 1647), and probes 2, 4, 5, and 6 used in Figure 4E. Probe 2 detected a 21-nt RNA only in RNA from Red plants, while probes 4, 5, and 6 did not detect any small RNA species (as in Supplemental Fig. S8B). In addition, probe 1647 detected a 22-nt siRNA (Supplemental Fig. S8C). These results confirm that the pre-mRNA of *KPL* and *ARI14* form dsRNA in the nucleus, and that this dsRNA is processed to generate small RNAs that target *ARI14*.

Biogenesis pathway of *KPL-ARI14* nat-siRNA

We wanted to determine which RNAi pathway components might be involved in biogenesis of the *KPL-ARI14* nat-siRNA. We did not detect any *KPL-ARI14* nat-siRNA species with small RNA isolated from pollen (Supplemental Fig. S8), probably because such a species is expected to be sperm-specific and would be diluted in pollen RNA. Therefore, we used qPCR to examine the *ARI14* expression level in several small RNA biogenesis mutants and in their corresponding wild-type ecotypes (Fig. 5). *ARI14* transcript levels were increased in *dcl1*, indicating that DCL1 was the only DICER-like protein that had any effect. Higher levels of *ARI14* RNA were also detected in the dsRNA-binding protein mutant *hyl1*, in the RNA methyltransferase mutant *hen1*, in the RNA polymerase IV mutant *sde4*, and in *sgs3*. A mutation in *RDR6* (*SDE1*) had no effect on *ARI14* transcript levels, while a mutation in *RDR2* showed a slight increase. These results implicate a biogenesis pathway in which

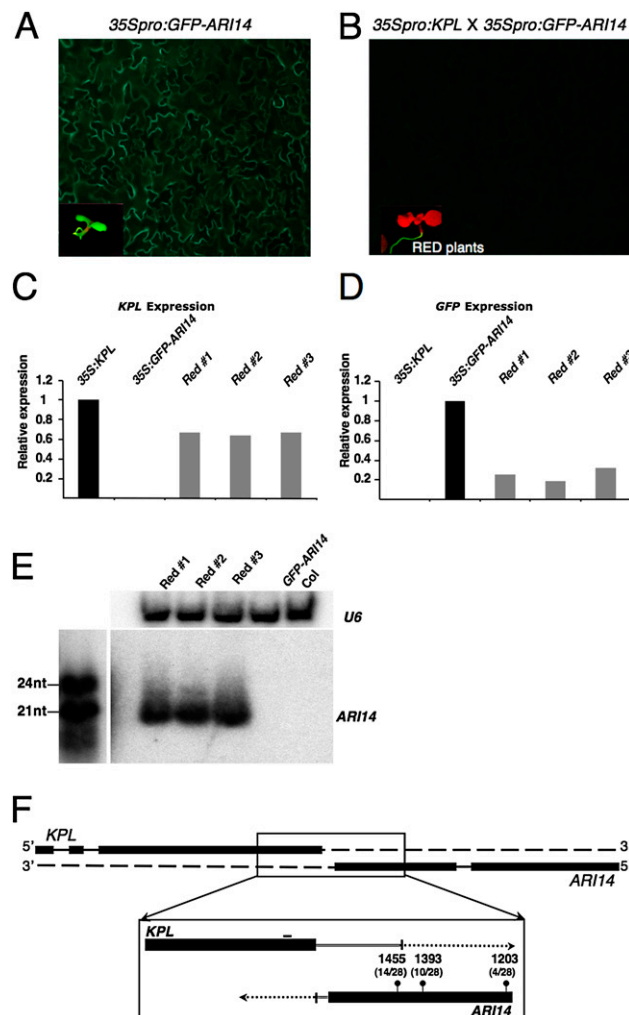


Figure 4. *ARI14* is regulated post-transcriptionally by *KPL*. (A,B) GFP expression in leaf epidermal cells of transgenic plants expressing the *35Spro:GFP-ARI14* construct (A) or both the *35Spro:KPL* and *35Spro:GFP-ARI14* constructs (B). (C,D) qPCR analyses of *KPL* (C) and *GFP* (D) levels in plants expressing *35Spro:GFP-ARI14* or both *35Spro:KPL* and *35Spro:GFP-ARI14* (Red plants). Total RNA was extracted from seedlings of F₂ plants selected for the presence of either one or both constructs. The RNA level of the relevant parent was set at 1. Each expression level was normalized to that of *IPP2*. (E) Northern blot analysis of small RNAs from F₂ progeny of plants carrying both the *35Spro:KPL* and *35Spro:GFP-ARI14* constructs, the *35Spro:GFP-ARI14* construct alone, or wild-type (Col) plants. The hybridization probes were a mixture of *ARI14*-specific oligomers; a U6 probe was used as a loading control for small RNAs. (F) Schematic of the overlapping region of *KPL* and *ARI14*. Thick and thin solid lines represent exons and introns, respectively. Dashed lines represent the complementary DNA strand. In the expanded area, double lines represent 3' UTR, vertical lines represent polyA addition sites, and dotted lines represent putative pre-mRNA downstream from the polyA addition site. Pinhead vertical lines represent the position of the *ARI14* cleavage site detected by RLM-5' RACE. The numbers above each point indicate the position of RLM-5' RACE products along the *ARI14* cDNA sequence, and the numbers in parentheses indicate the frequency of such products at each site. The short line above the *KPL* exon represents a *KPL* sequence signature obtained from the sperm siRNA of wild type (Slotkin et al. 2009).

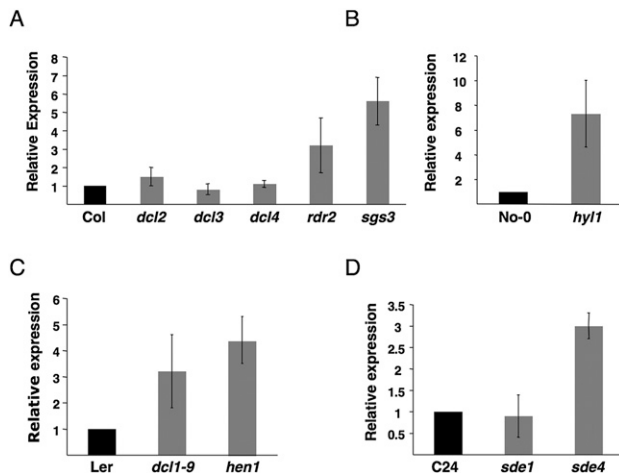


Figure 5. Accumulation of *ARI14* mRNA in different siRNA biogenesis mutants. (A–D) qPCR analysis of *ARI14* expression in 1-d-old open flowers of various small RNA biosynthesis mutants and their corresponding wild-type controls. The RNA level of the wild-type control in each group was set at 1. Each expression level was normalized to that of *IPP2*. Standard deviations were calculated from two to three biological replicas.

KPL-ARI14 nat-siRNA might be processed by the DCL1-HYL1 complex; stabilized by HEN1-mediated methylation; and amplified by RDR2, SGS3-mediated, and RNA polymerase IV-mediated reactions; and are consistent with findings for the salt-induced nat-siRNAs (Borsani et al. 2005) and pathogen-induced nat-siRNAs (Katiyar-Agarwal et al. 2006).

Overexpression of *ARI14* in sperm results in a reduced seed set phenotype

We hypothesized that the reduced seed set phenotype in the *kpl* mutants was a result of the failure to clear *ARI14* transcripts, leading to higher levels of *ARI14* in sperm. To test this hypothesis, we generated transgenic plants overexpressing *ARI14* under the control of the sperm-specific *KPL* promoter. Because endogenous *KPL* transcripts

might overlap with the *ARI14* transgene, which could result in suppression of the transgene, we constructed a siRNA-resistant form of *ARI14*. We obtained 20 independent T₁ lines, and verified by RT-PCR that they overexpressed the resistant form of *ARI14*. The siliques of these lines showed a reduced seed set similar to that of the *kpl* mutants (data not shown). We obtained T₂ progenies of six T₁ lines and grew them without selection, allowing the T-DNA to segregate. The T-DNA cassette carrying the resistant form of *ARI14* also contained a pollen-specific marker (*LAT52:tdTomato*), which we used to identify heterozygous plants. From each progeny, we analyzed two heterozygous plants and one plant from which the T-DNA had segregated away. We measured the relative expression levels of *ARI14* RNA in these plants by qPCR, and compared those levels with those in wild-type and *kpl-1/kpl-1* plants (Fig. 6A). We also scored the seed set (Fig. 6B,C). In all T₂ progenies tested, only the heterozygous plants carrying the transgene showed higher *ARI14* expression levels (4.2-fold to 19.7-fold, compared with 3.7-fold in *kpl-1/kpl-1*) (Fig. 6A). These plants also had, on average, ~50% reduced seed set (Fig. 6B), indicating that high *ARI14* levels in sperm impair fertilization. Notably, no single-fertilization events were observed in these plants, but only unfertilized ovules.

The *KPL-ARI14* nat-siRNA pair appears to be of recent origin

The ARIADNE proteins are members of an ancient RING protein family, with homologs in animals, fungi, and plants (Eisenhaber et al. 2007). In contrast, we could identify only *KPL* homologs from *Arabidopsis lyrata*, *Arabidopsis arenosa*, *Capsella rubella*, and *Brassica rapa*. *KPL* and *ARI14* are adjacent only in *Arabidopsis* and *C. rubella* (Supplemental Fig. S9). Most sequenced genomes have only a few *ARI* genes, but *Arabidopsis* has 16 members; probably as a result of several genome duplications (Mladek et al. 2003). In *Brassica*, there is only one gene representing the *ARI13/14/15* cluster (Supplemental Fig. S10). Alignments of a bacterial artificial

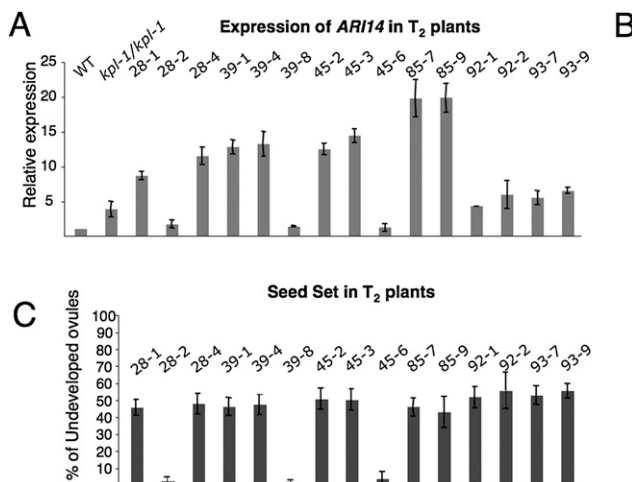


Figure 6. Overexpression of *ARI14* in sperm phenocopies the reduced seed set observed in *kpl* mutants. (A) Total RNA was extracted from 1-d-old open flowers of T₂ progenies from six independent T₁ lines (nos. 28, 39, 45, 85, 92, and 93) that ectopically expressed *ARI14* driven by the *KPL* promoter, and was then used for qPCR analysis. The expression level of *ARI14* in wild type was set to 1. Standard deviations were plotted from three replicates. The expression level was normalized to that of *IPP2*. (B) A representative silique of a self-pollinated T₂ plant overexpressing *ARI14*. The arrow indicates an undeveloped ovule. (C) Seed set analysis of the plants shown in A. Standard deviations were calculated from seed counts of 10 siliques for each plant.

chromosome (BAC) of *B. rapa* subsp. *pekinesis* (accession no. AC189659.2) with *A. thaliana* and *A. lyrata*, using the CoGe server (Lyons and Freeling 2008), show full synteny in the region (Supplemental Fig. S10) except for *KPL* and three upstream genes, although presumably *KPL* exists elsewhere in *Brassica* genomes, as there are homologs of *KPL* in *Brassica* ESTs (EST identifier JCVI_14548 and EE509369, from *Brassica oleracea* and *Brassica napus*, respectively). *KPL* was likely inserted into this genomic region in the *Brassicaceae* lineage, because no homologs are found in the syntenic regions of papaya, the most closely related outgroup in the *Order Brassicales*, or in grape, a more distantly related outgroup of basal eurosids. Thus, the *KPL-ARI14* nat-siRNA is likely to have originated recently within the *Brassicaceae* lineage of eurosid plants.

Discussion

KPL plays a key role in double fertilization

We provide evidence that post-transcriptional regulation, via a *cis*-nat-siRNA corresponding to the *KPL* and *ARI14* transcript pair, plays an important role in fertilization. Although double fertilization in angiosperms was described more than a century ago (McCormick 2004), the molecular basis underlying the mechanism that ensures double fertilization is still largely unknown. It is still unresolved as to whether both sperm are equally capable of fertilizing either the egg or central cell. Plants employ a rapid polyspermy block on the egg, but not on the central cell, to prevent both sperm from fertilizing the egg (Scott et al. 2008). Several *Arabidopsis* mutants have been described that have a single sperm-like cell, rather than two sperm, but there are differing findings as to whether such single sperm-like cells have preferred fusion partners (Iwakawa et al. 2006; Nowack et al. 2006; Chen et al. 2008; Kim et al. 2008; Frank and Johnson 2009; Gusti et al. 2009). Furthermore, when two wild-type sperm are delivered to a mutant female gametophyte that produces two eggs rather than one, both eggs can be fertilized, suggesting that both sperm have the capacity to fuse with an egg (Pagnussat et al. 2007; Ingouff et al. 2009). Interestingly, *Plumbago zeylanica* has morphologically distinct sperm that preferentially fertilize either the egg or the central cell (Russell 1985), and expression studies in these dimorphic sperm indicate that there might be promoters that are specific to one of the two sperm (Gou et al. 2009). However, the *KPL* promoter and all other promoters thus far tested in *Arabidopsis* (Alandete-Saez et al. 2008; Berger et al. 2008) drive expression in both sperm.

To date, *kpl* is the only mutant that shows prevalent single fertilizations, in which ~40% of the ovules show either only embryo or only endosperm development, even though the pollen tubes of *kpl* deliver two normal-looking sperm into the embryo sac. These data suggest that the *KPL-ARI* *cis*-nat-siRNA regulates an intrinsic component in the sperm that is essential for fertilization. For this reason, *kpl*, in combination with recently developed tools to image the fertilization process (Berger

et al. 2008), offers a unique opportunity to further explore the molecular mechanisms of double fertilization.

KPL and *ARI14* generate a sperm-specific *cis*-nat-siRNA pair

Our data indicate that *KPL* and *ARI14* generate a previously undiscovered *cis*-nat-siRNA pair. We showed that *KPL* and *ARI14* have anti-correlated expression patterns throughout pollen development (Fig. 1), and when comparing wild-type and *kpl* pollen (Fig. 3). We also demonstrated that coexpression of *ARI14* and *KPL* in transgenic plants induces the formation of *ARI14*-specific small RNAs (Fig. 4), and that the *Arabidopsis KPL-ARI14* nat-siRNAs are generated by small RNA biogenesis pathway components (Fig. 5). Genes involved in plant siRNA biosynthesis have pleiotropic phenotypes that have been studied extensively (Chen et al. 2002; Han et al. 2004; Peragine et al. 2004; Mosher et al. 2009). Homozygous mutants of *dcl1*, *hyl1*, and *hen1* show severe fertility defects among other developmental phenotypes. Male fertility is reduced in *hen1-1* and *hen1-2* plants (Chen et al. 2002). Although we are not aware of other published studies that directly address this issue, male transmission defects could certainly contribute to the overall reduced fertility of the *dcl1* and *hyl1* mutants. Other homozygous mutants like *sde4*, *rdr2*, and *sgs3* show either very subtle or no morphological or fertility differences, although their roles in siRNA biosynthesis are already well established (Peragine et al. 2004; Xie et al. 2004; Mosher et al. 2009). These observations suggest the possibility of redundancy between different proteins in the siRNA biosynthesis pathway, which might mask any effects on male transmission in the single mutants.

The reduced seed set phenotype seen when *ARI14* is overexpressed in sperm supports our hypothesis that high levels of *ARI14* transcripts in sperm impair fertilization (Fig. 6). In addition, the single fertilization events observed in *kpl* plants (Fig. 2) indicate that each sperm operates independently, depending on its level of *KPL* and *ARI14* transcripts. This interpretation is consistent with the variable (i.e., from silique to silique) seed set phenotypes (Fig. 1).

High-throughput sequencing of small RNAs (Backman et al. 2008) and genome-wide analyses of *Arabidopsis* potential *cis*-NAT pairs (Henz et al. 2007; Jin et al. 2008) have identified small RNAs that potentially target many protein-coding genes in *Arabidopsis*. However, genes that are expressed at very low levels, in a specific tissue, or under specific conditions, as well as genes that are not accurately annotated might have been missed in these studies. *KPL* and *ARI14* could be such genes due to their restricted expression within the male gametophyte. In animals, *cis*-nat-siRNAs are specially abundant in testis, and large-scale genomic approaches suggest that nat-siRNA overexpression in testis is a general phenomenon and not restricted to specific genes (Okada et al. 2008; Werner et al. 2009). This may suggest that *cis*-nat-siRNAs are a common regulatory mechanism for gene expression in gametes.

KPL can probably be defined as a new gene because its expression is sperm-specific, it is found only in some species of the *Brassicaceae*, and there are no homologs outside of this family. Sperm specificity and restricted species distributions are characteristic of some new genes in *Drosophila* species (Dorus et al. 2008). Furthermore, it appears that *KPL* was inserted recently in the genomic region adjacent to *ARI14* [Supplemental Fig. S10], suggesting that regulation of *ARI14* via a *KPL-ARI14* nat-siRNA is a novel evolutionary feature. Because *KPL* encodes a sperm-specific protein, it is plausible that it might play a role in fertilization, and that down-regulation of *KPL* in *kpl* mutants also might contribute to the reduced seed set phenotype.

Model for regulation of fertilization by *KPL-ARI14* nat-siRNA

Ubiquitination has been implicated in pollen development, gametogenesis, and seed development, and our results suggest that it is also important for proper fertilization. Ubiquitination of a protein requires a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2), and a ubiquitin protein ligase (E3) (Hershko and Ciechanover 1998); specificity is largely determined by the interactions between different E2s and E3s and their target proteins (Kumar et al. 1997). Transcripts encoding proteins involved in the ubiquitin-mediated degradation pathway were highly represented in sperm (Singh et al. 2002; Engel et al. 2003; Okada et al. 2007; Borges et al. 2008), suggesting that ubiquitin-mediated proteolysis might play an important role in male gametogenesis and fertilization.

ARI14 is a member of the *ARIADNE* gene family that encodes putative E3 ubiquitin ligases (Tan et al. 2000; Qiu and Fay 2006). *Arabidopsis* *ARI8* was shown to exhibit E3 ubiquitin ligase activity (Stone et al. 2005), thus it was

presumed that other members of the *ARIADNE*-like family in *Arabidopsis* had similar activities. *ARI14* is highly similar to *ARI13* and *ARI15*, and they, together with *ARI16*, are in the same subgroup of *ARIADNE*-like proteins in *Arabidopsis* (Mladek et al. 2003). *ARI13* is expressed in sperm and pollen, *ARI15* is ubiquitously expressed, including in sperm, and *ARI16* is expressed in siliques (Zimmermann et al. 2004; Borges et al. 2008). *ARI13/14/15* are located in a cluster on chromosome 5—probably a result of a late duplication occurring ~50 million years ago (Mladek et al. 2003). *ARI14* is missing four conserved cysteines in the first RING domain (Supplemental Fig. S11) that were shown to be crucial for binding to E2 ubiquitin-conjugating enzymes (Aguilera et al. 2000), so it is likely that *ARI14* lacks E3 ubiquitin ligase activity. It is plausible that *ARI13* and/or *ARI15* are active E3 ubiquitin ligases in sperm, because they have the required cysteines (Supplemental Fig. S11).

We propose a model for the generation of the *KPL-ARI14* nat-siRNA pair and its role in regulating fertilization (Fig. 7). Our model is based on the premise that *ARI14* is not an active E3 ubiquitin ligase and that it might interfere with the activity of bona fide E3 ubiquitin ligases in sperm, such as *ARI13* (used as an example since it is most similar to *ARI14*). In wild type, *ARI14* is expressed in both the vegetative and the sperm cells, while in sperm, *ARI13* and *KPL* are also expressed (Fig. 3). In the presence of *KPL* transcripts (i.e., in sperm), *ARI14* and *KPL* pre-mRNAs generate dsRNAs due to their transcript complementarity (Supplemental Fig. S8). These dsRNAs are recognized and processed by DCL1 and HYL1 to generate several small RNA products. Strand-specific small RNAs representing *KPL* sequences are incorporated into a RISC complex, which targets and cleaves *ARI14* transcripts in sperm, resulting in down-regulation of *ARI14*. In the absence of *ARI14*, *ARI13*, which is a putative active E3 ubiquitin ligase, can bind to

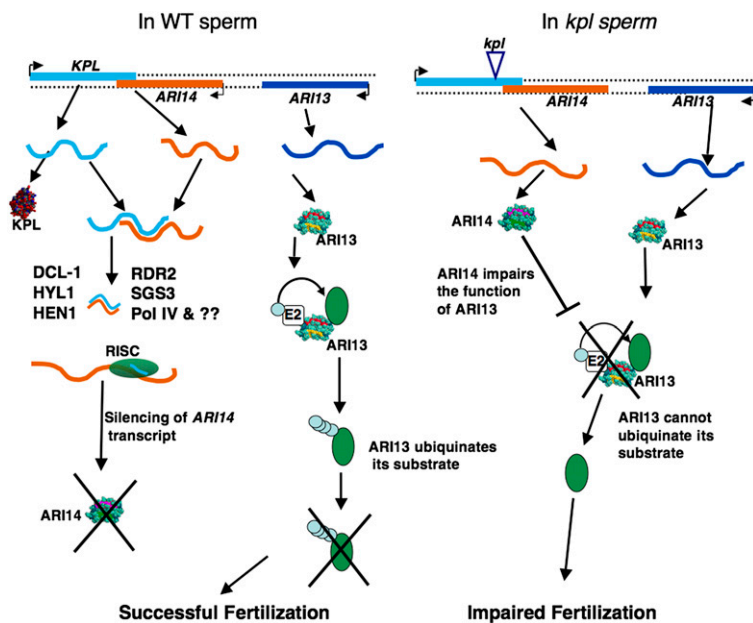


Figure 7. Model for biogenesis of *KPL-ARI14* nat-siRNAs and their proposed role in fertilization. In wild-type sperm, the *KPL-ARI14* nat-siRNAs lead to down-regulation of *ARI14*. Reduced *KPL* transcripts in *kpl* mutants result in accumulation of *ARI14* in the sperm. By binding to an unknown substrate in sperm, *ARI14*, which is likely an inactive ubiquitin E3 ligase, might prevent active *ARI* family E3 ligases (e.g., *ARI13*) in sperm from tagging that substrate for degradation, resulting in impaired fertilization.

its substrate and target it for degradation, leading to proper fertilization. In the *kpl* mutants, the level of full-length *KPL* transcripts is reduced, so that less dsRNA is produced, allowing some expression of *ARI14* in sperm. Although *ARI14* is unlikely to have enzymatic activity, it is possible that it can bind to a substrate without tagging it with ubiquitin. Thus, we propose that, in *kpl* mutants, *ARI14* can bind to a substrate and prevent it from being tagged for degradation, resulting in a higher level of the substrate in sperm, and that the activity of this substrate impairs fertilization.

In this scenario, one might expect *kpl* mutants that misexpress *ARI14* in sperm to have only unfertilized ovules, since both sperm are affected. However, they show frequent single-fertilization events. This could be due to a dosage effect; that is, in *kpl* sperm, there could be competition for the substrate between the inactive *ARI14* and the active *ARI13*, thus affecting the level of the substrate and therefore determining the fate of each sperm. It will be important to identify proteins with which *ARI14* interacts to test this hypothesis. Nonetheless, because each sperm regulates its *KPL* and *ARI14* expression levels independently of the other sperm, single fertilizations are possible. The idea of a dosage effect is also consistent with the phenotypes caused by overexpression of *ARI14* in sperm (Fig. 6); *ARI14* expression was driven by the strong *KPL* promoter, which resulted in higher levels of *ARI14* in sperm than found in the *kpl* mutants. Thus, the balance is shifted toward a situation in which *ARI14* protects the substrate from being degraded. In these transgenic lines, reduced seed sets were associated only with unfertilized ovules, and no single-fertilization events were observed. Taken together, our results demonstrate a key role for a *cis*-nat-siRNA in the regulation of double fertilization in plants, and open new avenues for investigating the molecular mechanisms that control this critical developmental process.

Materials and methods

Plant material and growth conditions

The *A. thaliana* mutants *rdr2-1*, *dcl2-1*, *dcl3-1*, and *dcl4-2* were provided by Jim Carrington (Oregon State University), *dcl1-9* and *hen1-1* were provided by Xuemei Chen (University of California at Riverside), *sde1*(*rdr6*) and *sde4*/*nrd1a* was provided by David Baulcombe (Sainsbury Laboratory, Norwich, United Kingdom), *sgs3* was provided by Herve Vaucheret (Institut National de la Recherche Agronomique, Versailles, France), and *hyl1* was provided by Nina Fedoroff (Pennsylvania State University). These mutants were in the Columbia (Col-0), Landsberg *erecta* (*Ler*), No-0, or C24 genetic backgrounds, as indicated in the text and figures. *FAC1pro:eGFP-GUS* seeds were from Chun-Ming Liu (Institute of Botany, Chinese Academy of Sciences, Beijing, China). Growth conditions are described in the Supplemental Material.

T-DNA lines

The *kpl-1* mutant, carrying a T-DNA inserted in the third exon of *At5g63720* (FST 184H02), was obtained from the INRA Versailles T-DNA mutant collection (WS4 ecotype) (Samson et al. 2002). The *kpl-2* mutant (*Ler* ecotype), carrying a gene trap transposon

insertion (GT_5_38351) was obtained from the *Arabidopsis* Biological Resource Center (ABRC) Stock Center. We used a PCR-based approach to confirm both T-DNA insertions. See the Supplemental Material for technical details and primer information.

Phenotypic analysis

Seed set analyses, GUS and DAPI staining, and imaging of unfertilized and fertilized ovules were performed as described in Alandete-Saez et al. (2008). See the Supplemental Material for details.

Statistical analysis

χ^2 was calculated using GraphPad QuickCalcs (<http://www.graphpad.com/quickcalcs/chisquared1.cfm>).

RNA analyses

High-molecular-weight RNA was extracted using RNeasy Kits (Qiagen). Low-molecular-weight RNA was extracted using TRIzol reagent (Invitrogen) or a PureLink miRNA isolation kit (Invitrogen). See the Supplemental Material for details and primers.

RLM 5' RACE and 3' RACE

A modified procedure for RNA ligase-mediated RACE (RLM 5' RACE) and 3' RACE were performed using the GeneRacer kit (Invitrogen). See the Supplemental Material for details and primers.

Acknowledgments

We thank Feng Li for assistance and advice on small RNA analyses, Leonor Boavida for discussions and advice on phenotypic characterizations, Eric Lyons for advice on evolution and synteny, and Jay Hollick, Jorge Muschietti, Binglian Zheng, Miltos Tsiantis, and Mona Monfared for comments on the manuscript. We thank Roger Tsien for tdTomato and mCherry, Fred Berger for *HTR10pro:HTR10-mRFP* seeds, Luca Comai for *A. arenosa* sequences, Peter Quail for *35S:PIF3* seeds, David Hantz and Julie Calfas for excellent greenhouse assistance, and undergraduates Sam Leiboff, Marian Wahba, Jamison Smith, Artur Karasyov, Chih-Han Lee, Frazier Philips, Laura Mahoney, and Allison Clark for technical assistance. This work was supported by NSF Plant Genome program grant 0211742, and by U.S. Department of Agriculture Current Research Information System 5335-21000-030-00D. M.R. was partially supported by a Vaadia-BARD Post-doctoral Fellowship Award number FI-391-2006 from BARD, the United States–Israel Binational Agricultural Research and Development Fund.

References

- Aguilera M, Oliveros M, Martinez-Padron M, Barbas JA, Ferrus A. 2000. Ariadne-1: A vital *Drosophila* gene is required in development and defines a new conserved family of RING-finger proteins. *Genetics* **155**: 1231–1244.
- Alandete-Saez M, Ron M, McCormick S. 2008. GEX3, expressed in the male gametophyte and in the egg cell of *Arabidopsis thaliana*, is essential for micropylar pollen tube guidance and plays a role during early embryogenesis. *Molecular Plant* **1**: 586–598.
- Backman TWH, Sullivan CM, Cumbie JS, Miller ZA, Chapman EJ, Fahlgren N, Givan SA, Carrington JC, Kasschau KD.

2008. Update of ASRP: The *Arabidopsis* Small RNA Project database. *Nucleic Acids Res* **36**: D982–D985. doi: 10.1093/nar/gkml1997.
- Berger F, Hamamura Y, Ingouff M, Higashiyama T. 2008. Double fertilization—Caught in the act. *Trends Plant Sci* **13**: 437–443.
- Boavida LC, McCormick S. 2007. Temperature as a determinant factor for increased and reproducible in vitro pollen germination in *Arabidopsis thaliana*. *Plant J* **52**: 570–582.
- Borges F, Gomes G, Gardner R, Moreno N, McCormick S, Feijo JA, Becker JD. 2008. Comparative transcriptomics of *Arabidopsis* sperm cells. *Plant Physiol* **148**: 1168–1181.
- Borsani O, Zhu JH, Verslues PE, Sunkar R, Zhu JK. 2005. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in *Arabidopsis*. *Cell* **123**: 1279–1291.
- Chen X, Liu J, Cheng Y, Jia D. 2002. HEN1 functions pleiotropically in *Arabidopsis* development and acts in C function in the flower. *Development* **129**: 1085–1094.
- Chen Z, Tan JL, Ingouff M, Sundaresan V, Berger F. 2008. Chromatin assembly factor 1 regulates the cell cycle but not cell fate during male gametogenesis in *Arabidopsis thaliana*. *Development* **135**: 65–73.
- Dorus S, Freeman ZN, Parker ER, Heath BD, Karr TL. 2008. Recent origins of sperm genes in *Drosophila*. *Mol Biol Evol* **25**: 2157–2166.
- Eisenhaber B, Chumak N, Eisenhaber F, Hauser MT. 2007. The ring between ring fingers (RBR) protein family. *Genome Biol* **8**: 209. doi: 10.1186/gb-2007-1188-1183-1209.
- Engel ML, Chaboud A, Dumas C, McCormick S. 2003. Sperm cells of *Zea mays* have a complex complement of mRNAs. *Plant J* **34**: 697–707.
- Frank AC, Johnson MA. 2009. Expressing the diphtheria toxin A subunit from the HAP2(GCS1) promoter blocks sperm maturation and produces single sperm-like cells capable of fertilization. *Plant Physiol* **151**: 1390–1400.
- Ghildiyal M, Zamore PD. 2009. Small silencing RNAs: An expanding universe. *Nat Rev Genet* **10**: 94–108.
- Gou X, Yuan T, Wei X, Russell SD. 2009. Gene expression in the dimorphic sperm cells of *Plumbago zeylanica*: Transcript profiling, diversity, and relationship to cell type. *Plant J* **60**: 33–47.
- Gusti A, Baumberger N, Nowack M, Pusch S, Eisler H, Potuschak T, De Veylder L. 2009. The *Arabidopsis thaliana* F-box protein FBL17 is essential for progression through the second mitosis during pollen development. *PLoS ONE* **4**: e4780. doi: 10.1371/journal.pone.0004780.
- Han MH, Goud S, Song L, Fedoroff N. 2004. The *Arabidopsis* double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proc Natl Acad Sci* **101**: 1093–1098.
- Henz SR, Cumbie JS, Kasschau KD, Lohmann JU, Carrington JC, Weigel D, Schmid M. 2007. Distinct expression patterns of natural antisense transcripts in *Arabidopsis*. *Plant Physiol* **144**: 1247–1255.
- Hershko A, Ciechanover A. 1998. The ubiquitin system. *Annu Rev Biochem* **67**: 425–479.
- Hony D, Twell D. 2004. Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol* **5**: R85. doi: 10.1186/gb-2004-1185-1111-r1185.
- Ingouff M, Sakata T, Sprunck S, Dresselhaus T, Berger F. 2009. The two male gametes share equal ability to fertilize the egg cell in *Arabidopsis thaliana*. *Curr Biol* **19**: R19–R20. doi: 10.1016/j.cub.2008.1011.1025.
- Iwakawa H, Shinmyo A, Sekine M. 2006. *Arabidopsis* CDKA1, a cdc2 homologue, controls proliferation of generative cells in male gametogenesis. *Plant J* **45**: 819–831.
- Jen CH, Michalopoulos I, Westhead DR, Meyer P. 2005. Natural antisense transcripts with coding capacity in *Arabidopsis* may have a regulatory role that is not linked to double-stranded RNA degradation. *Genome Biol* **6**: R51. doi: 10.1186/gb-2005-1186-1186-r1151.
- Jin H, Vacic V, Girke T, Lonardi S, Zhu JK. 2008. Small RNAs and the regulation of cis-natural antisense transcripts in *Arabidopsis*. *BMC Mol Biol* **9**: 6. doi: 10.1186/1471-2199-1189-1186.
- Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A Jr, Zhu JK, Staskawicz BJ, Jin H. 2006. A pathogen-inducible endogenous siRNA in plant immunity. *Proc Natl Acad Sci* **103**: 18002–18007.
- Kim VN. 2005. Small RNAs: Classification, biogenesis, and function. *Mol Cells* **19**: 1–15.
- Kim H, Oh S, Brownfield L, Hong S, Ryu H, Hwang I, Twell D, Nam H. 2008. Control of plant germline proliferation by SCF(FBL17) degradation of cell cycle inhibitors. *Nature* **455**: 1134–1137.
- Kumar S, Kao WH, Howley PM. 1997. Physical interaction between specific E2 and Hect E3 enzymes determines functional cooperativity. *J Biol Chem* **272**: 13548–13554.
- Lyons E, Freeling M. 2008. How to usefully compare homologous plant genes and chromosomes as DNA sequences. *Plant J* **53**: 661–673.
- McCormick S. 2004. Control of male gametophyte development. *Plant Cell* **16**: S142–S153. doi: 10.1105/tpc.016659.
- Mladek C, Guger K, Hauser MT. 2003. Identification and characterization of the ARIADNE gene family in *Arabidopsis*. A group of putative E3 ligases. *Plant Physiol* **131**: 27–40.
- Mosher RA, Melnyk CW, Kelly KA, Dunn RM, Studholme DJ, Baulcombe DC. 2009. Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. *Nature* **460**: 283–286.
- Nowack M, Grini P, Jakoby M, Lafos M, Koncz C, Schnittger A. 2006. A positive signal from the fertilization of the egg cell sets off endosperm proliferation in angiosperm embryogenesis. *Nat Genet* **38**: 63–67.
- Okada T, Singh MB, Bhalla PL. 2007. Transcriptome profiling of *Lilium longiflorum* generative cells by cDNA microarray. *Plant Cell Rep* **26**: 1045–1052.
- Okada Y, Tashiro C, Numata K, Watanabe K, Nakaoka H, Yamamoto N, Okubo K, Ikeda R, Saito R, Kanai A, et al. 2008. Comparative expression analysis uncovers novel features of endogenous antisense transcription. *Hum Mol Genet* **17**: 1631–1640.
- Okamura K, Lai EC. 2008. Endogenous small interfering RNAs in animals. *Nat Rev Mol Cell Biol* **9**: 673–678.
- Pagnussat GC, Yu HJ, Sundaresan V. 2007. Cell-fate switch of synergid to egg cell in *Arabidopsis eostre* mutant embryo sacs arises from misexpression of the BEL1-like homeodomain gene BLH1. *Plant Cell* **19**: 3578–3592.
- Peragine A, Yoshikawa M, Wu G, Albrecht HL, Poethig RS. 2004. *SGS3* and *SGS2/SDE1/RDR6* are required for juvenile development and the production of *trans*-acting siRNAs in *Arabidopsis*. *Genes & Dev* **18**: 2368–2379.
- Qiu XH, Fay DS. 2006. ARI-1, an RBR family ubiquitin-ligase, functions with UBC-18 to regulate pharyngeal development in *C. elegans*. *Dev Biol* **291**: 239–252.
- Russell SD. 1985. Preferential fertilization in *Plumbago*: Ultrastructural evidence for gamete-level recognition in an angiosperm. *Proc Natl Acad Sci* **82**: 6129–6132.
- Samson F, Brunaud V, Balzergue S, Dubreucq B, Lepiniec L, Pelletier G, Caboche M, Lecharny A. 2002. FLAGdb/FST: A database of mapped flanking insertion sites (FSTs) of *Arabidopsis thaliana* T-DNA transformants. *Nucleic Acids Res* **30**: 94–97.

- Scott RJ, Armstrong SJ, Doughty J, Spielman M. 2008. Double fertilization in *Arabidopsis thaliana* involves a polyspermy block on the egg but not the central cell. *Molecular Plant* **1**: 611–619.
- Singh MB, Xu H, Bhalla PL, Zhang ZJ, Swoboda I, Russell SD. 2002. Developmental expression of polyubiquitin genes and distribution of ubiquitinated proteins in generative and sperm cells. *Sex Plant Reprod* **14**: 325–329.
- Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, Feijo JA, Martienssen RA. 2009. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136**: 461–472.
- Stone SL, Hauksdottir H, Troy A, Herschleb J, Kraft E, Callis J. 2005. Functional analysis of the RING-type ubiquitin ligase family of *Arabidopsis*. *Plant Physiol* **137**: 13–30.
- Tan NGS, Ardley HC, Rose SA, Leek JP, Markham AF, Robinson PA. 2000. Characterisation of the human and mouse orthologues of the *Drosophila ariadne* gene. *Cytogenet Cell Genet* **90**: 242–245.
- Twell D, Yamaguchi J, McCormick S. 1990. Pollen-specific gene expression in transgenic plants: Coordinate regulation of two different tomato gene promoters during microsporogenesis. *Development* **109**: 705–713.
- Vaucheret H. 2006. Post-transcriptional small RNA pathways in plants: Mechanisms and regulations. *Genes & Dev* **20**: 759–771.
- Vazquez F. 2006. *Arabidopsis* endogenous small RNAs: Highways and byways. *Trends Plant Sci* **11**: 460–468.
- Werner A, Carlile M, Swan D. 2009. What do natural antisense transcripts regulate? *RNA Biol* **6**: 43–48.
- Xie ZX, Qi XP. 2008. Diverse small RNA-directed silencing pathways in plants. *Biochim Biophys Acta Gene Regul Mech* **1779**: 720–724.
- Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC. 2004. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* **2**: E104. doi: 10.1371/journal.pbio.0020104.
- Xu J, Zhang HY, Xie CH, Xue HW, Dijkhuis P, Liu CM. 2005. EMBRYONIC FACTOR 1 encodes an AMP deaminase and is essential for the zygote to embryo transition in *Arabidopsis*. *Plant J* **42**: 743–756.
- Yadegari R, Drews GN. 2004. Female gametophyte development. *Plant Cell* **16**: S133–S141. doi: 10.1105/tpc.018192.
- Zhang Y, Liu XS, Liu QR, Wei LP. 2006. Genome-wide in silico identification and analysis of cis natural antisense transcripts (cis-NATs) in 10 species. *Nucleic Acids Res* **34**: 3465–3475.
- Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W. 2004. GENEVESTIGATOR. *Arabidopsis* microarray database and analysis toolbox. *Plant Physiol* **136**: 2621–2632.