JIPB Journal of Integrative Plant Biology

Journal of Integrative Plant Biology 2012, **54** (1): 55–64

Research Article

EMBRYONIC FACTOR 19 Encodes a Pentatricopeptide Repeat Protein that is Essential for the Initiation of Zygotic Embryogenesis in *Arabidopsis*

Dali Yu[1,2], Li Jiang[1,2], Huaqin Gong[1] and Chun-Ming Liu[1*]

[1]Key Laboratory of Plant Molecular Physiology, Institute of Botany, the Chinese Academy of Sciences, Beijing 100093, China

[2]Graduate University of Chinese Academy of Sciences, Beijing 100049, China

[*]Corresponding author

Tel: +86 10 6283 6663; Fax: +86 10 6283 6691; E-mail: cmliu@ibcas.ac.cn

Available online on 21 November 2011 at www.jipb.net and www.wileyonlinelibrary.com/journal/jipb

doi: 10.1111/j.1744-7909.2011.01089.x

Abstract

Early embryogenesis is the most fundamental developmental process in biology. Screening of ethyl methanesulfonate (EMS)-mutagenized populations of *Arabidopsis thaliana* led to the identification of a zygote-lethal mutant *embryonic factor 19* (*fac19*) in which embryo development was arrested at the elongated zygote to octant stage. The number of endosperm nuclei decreased significantly in *fac19* embryos. Genetic analysis showed *fac19* was caused by a single recessive mutation with typical mendelian segregation, suggesting equal maternal and paternal contributions of *FAC19* towards zygotic embryogenesis. Positional cloning showed that *FAC19* encodes a putative mitochondrial protein with 16 conserved pentatricopeptide repeat (PPR) motifs. The *fac19* mutation caused a conversion from hydrophilic serine located in a previously unknown domain to hydrophobic leucine. Crosses between *FAC19*/*fac19* and the T-DNA insertion mutants in the same gene failed to complement the *fac19* defects, confirming the identity of the gene. This study revealed the critical importance of a PPR protein-mediated mitochondrial function in early embryogenesis.

Keywords: embryogenesis; mitochondrial proteins; pentatricopeptide repeat motifs; zygote activation.

Yu D, Jiang L, Gong H, Liu CM (2012) *EMBRYONIC FACTOR 19* encodes a pentatricopeptide repeat protein that is essential for the initiation of zygotic embryogenesis in *Arabidopsis*. *J. Integr. Plant Biol*. **54**(1), 55–64.

Introduction

Embryogenesis is a key step in the life cycle of flowering plants, during which the basic body plan and tissue pattern are established. In the model plant *Arabidopsis*, early embryonic development is initiated after fertilization and proceeds through a virtually invariant series of cell divisions, revealing a tightly controlled zygotic activation, cell division, patterning and morphogenesis (Fan et al. 2008). The freshly formed zygote is about 20 μm in length, which then goes through a process of cell wall thickening and cell elongation to expand approximately three times along the future apical-basal axis, followed by the first asymmetric division that yields a smaller embryonic apical cell and a larger extra-embryonic basal cell (Mansfield and Briarty 1991). The apical cell and its daughter cells divide twice vertically and once horizontally, leading to a spherical proembryo of eight cells (the octant stage). In contrast, the basal cell and its derivatives undergo repeatedly horizontal divisions, thus producing a 9-cell suspensor that links the embryo to the maternal tissue. After the octant stage, the uppermost daughter cell of the suspensor divides transversely to form the hypophysis, which contributes to the formation of the root meristem. This marks the end of the early phase of embryogenesis (Jürgens 1998).

Independent studies revealed that maternal genes contribute to both early embryo and endosperm development through

stored maternal mRNAs and/or proteins and selective silencing of paternal alleles (Vielle-Calzada et al. 2000; Grimanelli et al. 2005; Pillot et al. 2010). However, controversial results have been reported, showing immediate activations of zygotic genes (Weijers et al. 2001; Xu et al. 2005; Ning et al. 2006; Meyer and Scholten 2007). In addition, many early embryo defective genes segregate with a typical 3:1 ratio, suggesting no apparent maternal effects (Tzafrir et al. 2004). So the extent of maternal influence on early seed development is still not clear.

Through screenings of embryo defective mutants and reverse genetic approaches, several genes have been identified to be essential for zygotic activation in *Arabidopsis*. These genes function in various processes such as protein degradation, energy metabolism, DNA replication and RNA processing. *AtCUL1*, whose product is a subunit of the E3-unbiquitin ligase complex called the Skp1-Cul1/Cdc53-F-box (SCF), functions probably in the first divisions of the zygote and the endosperm in controlling chromosome segregations (Shen et al. 2002). *EMBRYONIC FACTOR 1* (*FAC1*), which is essential for the zygote-to-embryo transition, encodes an adenosine monophosphate (AMP) deaminase and may act through AMP depletion to provide more energy for zygote activation (Xu et al. 2005). *ZEU1* encodes a thynidylate kinase (*AtTMPK*), which synthesizes dTDP, the limiting component of dTTP biosynthesis, involved in the G1/S-phase transition at fertilization by regulating DNA replication (Ronceret et al. 2008). *AtCDC5*, a homolog of *Schizosaccharomyces pombe CDC5* gene in *Arabidopsis*, plays an important role in the first division of embryo and endosperm through cell cycle regulation in the G2/M phase transition (Lin et al. 2007). Recently, it has been reported that RNA processing plays an important role in early embryonic development. Mutations of *GAMETOPHYTIC FACTOR1* (*GFA1*) and *YAOZHE* (*YAO*), two genes encoding nuclear spliceosome proteins, lead to embryo arrest at a very early stage (Liu et al. 2009; Li et al. 2010). Interestingly, *GLUTAMINE-RICH PROTEIN23 (GRP23)*, which encodes a pentatricopeptide repeat (PPR) protein, can interact with RNA polymerase II subunit III and functions as a potential regulator of gene expression during early embryogenesis (Ding et al. 2006). PPR genes, encoding proteins characterized by the tandem repeats of the degenerate PPR motif, are found in all sequenced genomes but greatly expand in plants and their functions remain largely unknown (Small and Peeters 2000; O'Toole et al. 2008).

In this study, we describe a novel early embryo-lethal mutant *embryonic factor 19* (*fac19)* in which the majority of mutant embryos were arrested from zygote to octant stage. The mutation was inherited in a typical Mendelian manner, suggesting no apparent parental effect. Map-based cloning revealed that *FAC19* encodes a predicted mitochondrial PPR protein with 16 conserved PPR motifs and is expressed preferentially in open flowers. Through identification of additional two T-DNA insertion mutants and complementation analyses, the identity of the *FAC19* gene was confirmed. Our results established the crucial role of *FAC19* in zygotic activation, most likely in mitochondria RNA processing.

Results

fac19 mutant exhibits zygote arrest phenotype

The fourth and fifth (in order of appearance) silliques from the primary inflorescence of 4–5-week-old plants were selected for phenotype examination. After splitting with fine forceps and examination under a dissecting microscope, full seed set was observed in wild type (**Figure 1A**), while under the same conditions heterozygous *FAC19*/*fac19* plants segregated small, shrunken and light brown aborted seeds that were randomly distributed among normal green seeds (**Figure 1B**, **Table 1**). To analyze the mutant phenotype in detail, immature seeds from heterozygous *fac19* plants were excised before being shrunken and cleared by a whole-mount clearing method with HCG solution containing 80% (W/V) chloral hydrate and 10% (V/V) glycerol. Cleared seeds were examined with a Leica florescence microscope using differential interference contrast (DIC) optics. The result showed that when the wild type embryos reached globular stage (**Figure 1G**), among the aborted seeds 22.9% were arrested at the semi-elongated and elongated zygote stage (**Figure 1H, I**), 45.9% were at the first division stage (**Figure 1J**), and 31.2% were at 2- to 4-cell embryo stage (**Figure 1K**, $n = 157$). Section and staining analysis confirmed the phenotype observed under the DIC microscope and revealed no additional cellular defects (**Figure 1C–F**). Endosperm development was also affected in the *fac19* mutant. Compared to the wild type, the numbers of endosperm nuclei were decreased significantly when the embryos were at the same developmental stage (**Figure 2**).

An equal parental contribution of *FAC19* towards the zygote embryogenesis

Microscopic examination showed that 24.2% of seeds produced by *FAC19*/*fac19* plants were abnormal and shriveled quickly after fertilization (**Figure 1B**, **Table 1**). This suggested *fac19* embryo-lethal phenotype was caused by a single recessive mutation. When reciprocal crosses were carried out between wild type and *FAC19*/*fac19* plants, regardless of whether the mutant was used as male or female, nearly all F1 seeds developed normally (**Table 1**). This suggested *FAC19* had no parental effect, where either the maternally or paternally delivered wild type copy of *FAC19* gene was

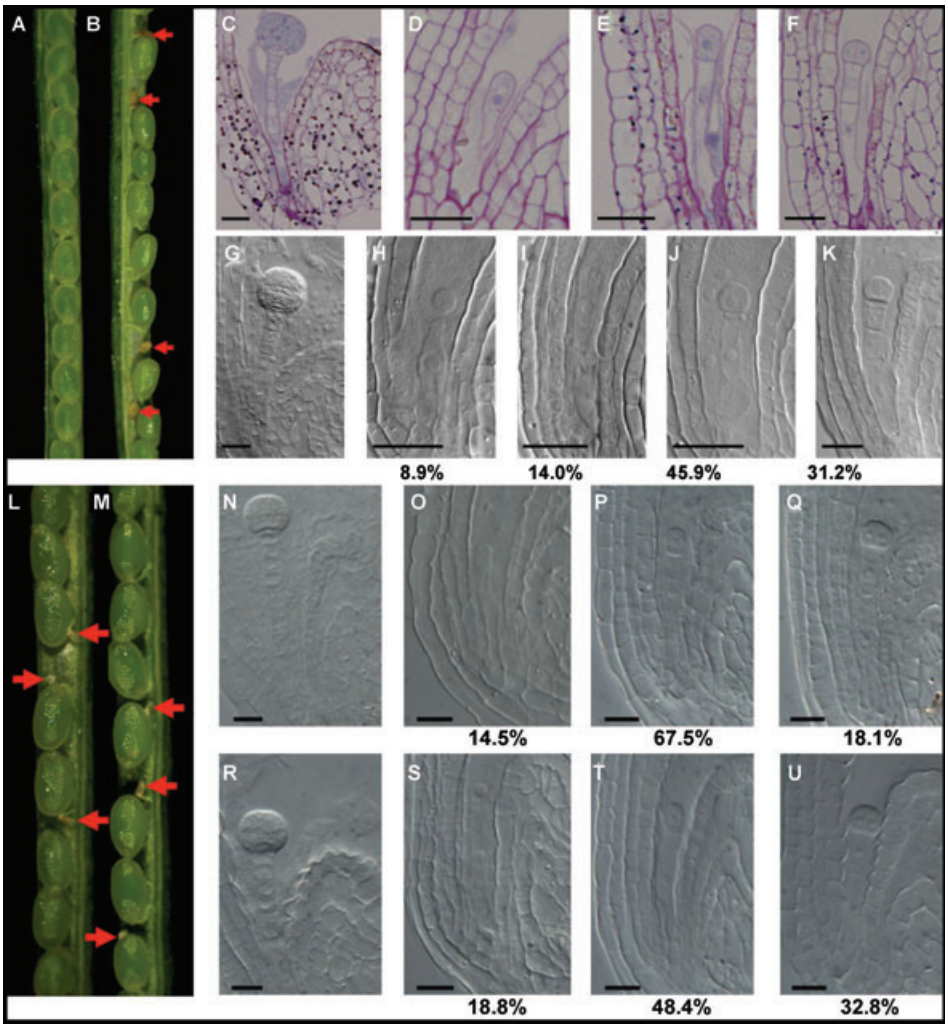

Figure 1. The *fac19* phenotype.

(A) Seeds in an open sillique from a wild type plant.
(B) Seeds in an open sillique from a *FAC19*/*fac19* plant (aborted seeds are indicated by red arrows).
(C–F) Sections of wild type (C) and the arrested *fac19* seeds (D–F). Bar = 20 μm.
(G–K) Wild type (G) and arrested *fac19* seeds (H–K) observed under Nomarski optics. Bar = 25 μm.
(L) Seeds in an open sillique from a *SALK_012298* heterozygous plant (aborted seeds are indicated by red arrows).
(M) Seeds in an open sillique from a *SALK_143867* heterozygous plant (aborted seeds are indicated by red arrows).
(N–Q) Wild type (C) and arrested (D–F) seeds from *SALK_143867* heterozygous plants observed under Nomarski optics. Bar = 25 μm.
(R–U) Wild type (G) and arrested (S–U) seeds from *SALK_012298* heterozygous plants observed under Nomarski optics. Bar = 25 μm.
The numbers underneath indicate percentages of defected embryos at a particular stage. For *fac19*, $n = 157$; *SALK_012298*, $n = 83$; For *SALK_143867*, $n = 64$. Bar = 25 μm.

sufficient to rescue the *fac19* mutation. The segregation of wild type and *FAC19*/*fac19* plants produced after these reciprocal crosses was consistent with a 1:1 ratio (**Table 2**), suggesting both the male and female gametophytes functioned normally and gametes carrying the mutation could be fully transmitted to the next generation. Together, these data indicated that the phenotype conferred by *fac19* mutation is attributable to a single recessive nuclear mutation with an equal parental contribution.

Map-based cloning of *FAC19*

To clone the *FAC19* gene, *FAC19*/*fac19* was crossed to the wild type (L*er*) to generate the mapping population. Using 38 wild

Table 1. Genetic analyses of the *fac19* mutant

Cross combination	Normal seeds	Aborted seeds	Unfertilized ovules	Percentage of aborted seeds[a]
WT self-pollination	169	5	1	2.9
FAC19/fac19 self-pollination	247	79	0	24.2[b]
♀WT×*FAC19/fac19*♂	134	5	1	3.6
♀*FAC19/fac19*×WT♂	118	3	2	2.5
SALK_012298 self-pollination	241	79	1	24.7[b]
SALK_143867 self-pollination	200	83	1	29.3[b]
♀*FAC19/fac19*×*SALK_012298*♂	264	107	5	28.8[b]
♀*SALK_012298*×*FAC19/fac19*♂	161	65	4	28.8[b]
♀*FAC19/fac19*×*SALK_143867*♂	429	169	5	28.3[b]
♀*SALK_143867*×*FAC19/fac19*♂	88	32	1	26.7[b]

[a]Aborted seeds/(aborted seeds + normal seeds).

[b]Percentage of aborted seeds was not significantly different from the theoretical proportion of 25% for a single recessive mutation ($P > 0.05$).

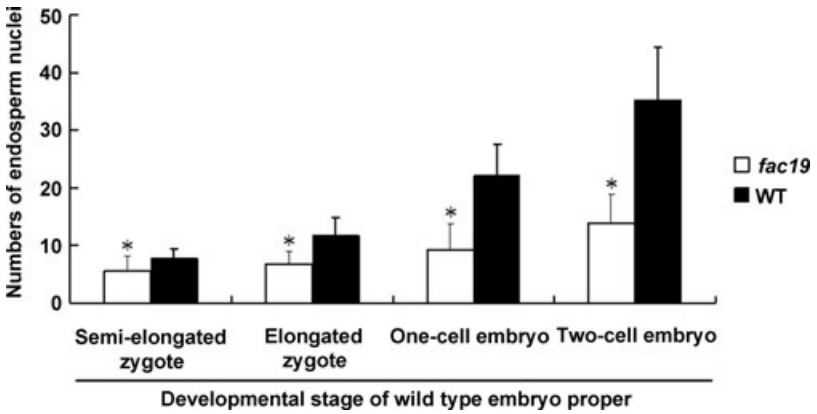

Figure 2. Numbers of endosperm nuclei in *fac19* and wild type (WT) seeds.

$n \geq 11$, *t*-test, *, $P < 0.05$.

Table 2. Transmission of the *fac19* mutant allele

Cross combinations	Phenotypes of progeny plants		
	WT	*FAC19/fac19*	Ratio
WT♀×*FAC19/fac19*♂	66	53	1:0.8[a]
FAC19/fac19♀×WT♂	29	31	1:1.1[a]

[a]The segregation ratio was not significantly different from 1:1.

type plants obtained from the F2 population, *FAC19* was first roughly positioned in a 789 kb region on chromosome 1, between the molecular markers F3F19-1 and F1L1-1 (**Figure 3A**). By utilizing these two markers, 950 F2 and 792 F3 plants were screened for recombinants in the vicinity of the mutation, which allowed us to narrow the mutation down to a 73 kb region between markers F21F23-3 and F7A19-5 (**Figure 3A**). Based on The *Arabidopsis* Information Resource database (TAIR, http://www.arabidopsis.org), this region contains 28 genes. Based on the annotated gene structure, function and expression pattern, the open reading frames of candidate genes were amplified by polymerase chain reaction (PCR) from a *FAC19/fac19* plant and sequenced. One single base change (G to A) was identified (as a double peak in the sequencing output) in the only exon of *At1G13800* (**Figure 3B**), which led to an amino acid (AA) change in the 674th amino acid from hydrophilic serine (S) to hydrophobic leucine (L) (**Figure 3C**). Sequencing of its parental lines (Col and L*er*) confirmed that it was a true mutation. Amplified fragments from four F3 plants, two wild type and two heterozygous, were also sequenced and showed that both phenotypically heterozygous plants carried the mutation, while none of the two wild type ones did. G to A transition is the most common mutation in ethyl methanesulfonate (EMS)-mutagenized plants.

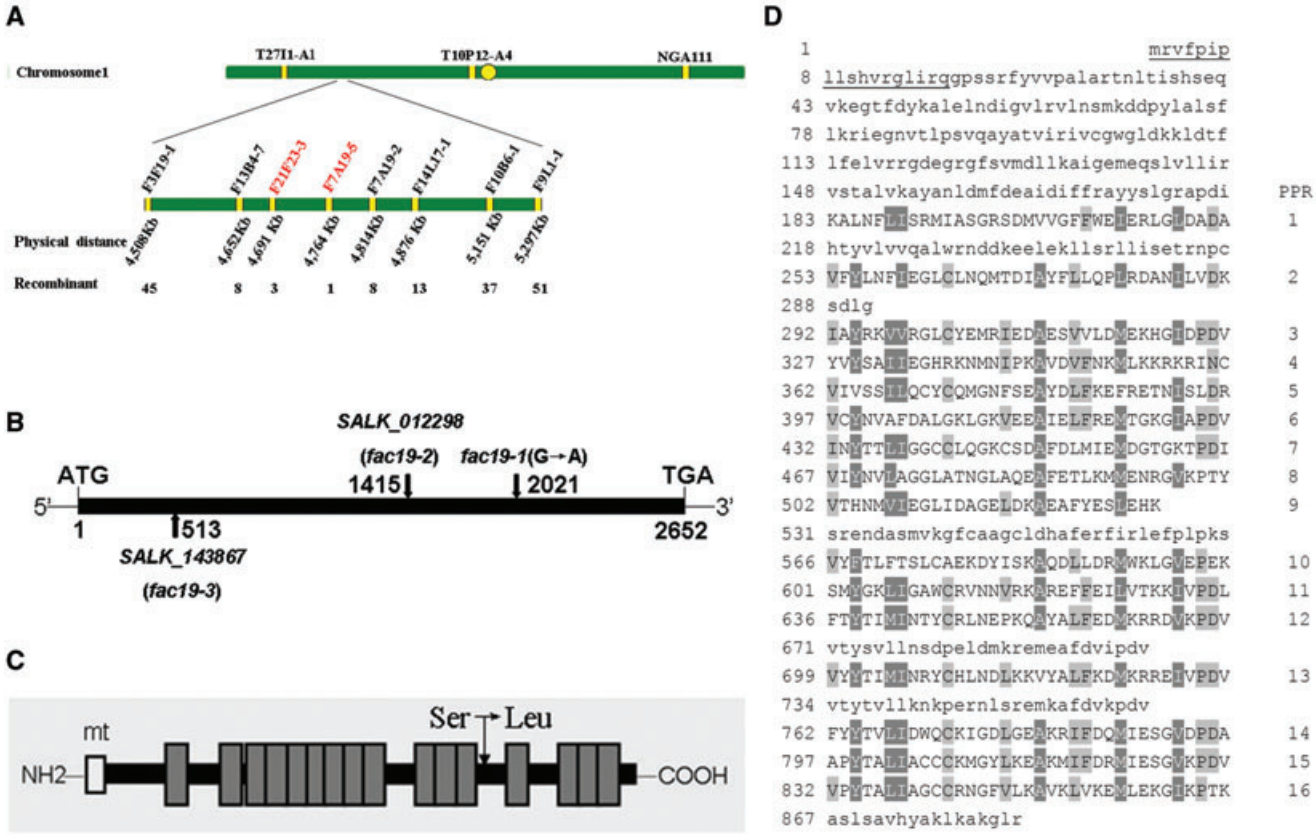

Figure 3. Molecular identification of the *FAC19* gene.

(A) Schematic representation of map-based cloning of *FAC19*, which is located to a 73 kb region between markers F21F23-3 and F7A19-5. **(B)** The structure of *FAC19* gene, showing the point mutation in *fac19-1* and the locations of two T-DNA insertions *(fac19-2* and *fac19-3)*. The gene is predicted to have a single open reading frame.

(C) Scheme of the FAC19 protein and the AA change in *fac19-1*. The point mutation converts the first serine located between 12th and 13th pentatricopeptide repeat (PPR) domains to leucine. The predicted mitochondrial targeting signal (mt) is indicated by a white box, while PPR motifs are represented by grey boxes.

(D) Predicted AA sequence of the FAC19 protein. Putative mitochondrial targeting signal is underlined. PPR motifs are in upper case, aligned and numbered from 1 to 16. Highly and moderately conserved AAs are highlighted by black and grey boxes, respectively.

Characterization of the knockout allele and complementation analysis

By searching the T-DNA insertion database (TAIR, http://www.arabidopsis.org), we found two independent T-DNA insertion lines (*SALK_012298* and *SALK_143867*), with T-DNAs inserted in different positions of *At1G13800* (**Figure 3B**). Seeds for these lines were obtained from the *Arabidopsis* Biological Resource Center (ABRC). The insertion sites were verified in these lines by PCR analysis using a T-DNA left border primer and two gene-specific primers, respectively (Table S1), in which no homozygous insertion plants were identified. Similar to *FAC19*/*fac19*, approximately 25% of seeds in silliques of *SALK_012298* and *SALK_143867* plants were aborted (**Table 1**). Similar to *fac19,* microscopic analyses revealed zygote arrest phenotypes in both lines (**Figure 1 L–U**). Allelic tests were performed by reciprocal crosses among two insertion lines and the *FAC19*/*fac19.* All F1 plants showed the zygote-lethal phenotype, confirming the deficiencies of *At1G13800* gene were responsible for the *fac19* phenotype (**Table 1**). As such, *fac19*, *SALK_012298* and *SALK_143867* were named *fac19-1, fac19-2* and *fac19-3,* respectively.

FAC19 encodes a PPR protein

FAC19 encodes a putative pentatricopeptide repeat (PPR) protein composed of 883 AAs, with a calculated molecular weight 100.52 kD and a pI of 6.96. Sequence analysis revealed that the FAC19 protein carried a predicted mitochondrial targeting signal at its N terminus (**Figure 3C, D**), 16 degenerate PPR

motifs from residues 183 to 866. Among these PPR motifs, only the 3rd to the 9th, the 10th to the 12th, and the 14th to the 16th repeats were contiguous (**Figure 3C, D**). From the perspective of C-terminus, it belongs to the P subclass of the PPR protein family (Small and Peeters 2000; Lurin et al. 2004). A BLASTP search revealed a number of proteins sharing high similarities with FAC19. Most of these proteins shared primarily their PPR domains with FAC19. The closest homolog is AlFAC19 (XP_002892774) from *Arabidopsis lyrata,* which carried the same number of PPR motifs and shared 89% sequence identities with FAC19. The closest paralog in the genome of *A. thaliana* is At2g26790, which shares 49.87% sequence identity with FAC19 at the protein level. Alignment of FAC19 with its homologs showed that the mutated Ser^{674} in *fac19* was located in a highly conserved previously unknown motif of DVVTY(S/T)VLLN (**Figure 4**). We named this conserved domain the "TV motif", since it is rich in threonine and valine. As the *fac19-1* mutant with the point mutation showed very similar phenotype to the T-DNA insertion mutants, we believe the TV motif is critical for the function of this particular class of PPR proteins.

Expressions of *FAC19*

Pentatricopeptide repeat genes are usually constitutively expressed at low levels in all tissues throughout plant development (Lurin et al. 2004). Microarray data indicated that it is the case for *FAC19* (**Figure 5A** and Figure S1). To elucidate the expression in detail, real-time PCR analysis was performed to examine the expression of *FAC19,* with RNA samples extracted from different tissues of wild type plants. The result showed that *FAC19* transcripts were most abundant in open flowers, followed by seedlings and inflorescences (**Figure 5B**).

Discussion

In this paper, we present data on molecular characterizations of the *FAC19* gene that is essential for the activation of zygotic embryogenesis. A G-to-A transition of *FAC19* leads to a zygote-lethal phenotype without parental effects. Genetic analysis showed that neither the male nor the female gametophyte development was affected. *FAC19* encodes a predicted mitochondrial protein with 16 conserved PPR motifs. The *fac19* mutant caused by a single AA change in a conserved TV motif failed to complement two mutants with T-DNA insertions in the same gene, confirming the cloning result.

Pentatricopeptide repeat motif consists of 35 degenerate AAs and often occurs as tandem repeats in proteins (Small and Peeters 2000). In fact, in addition to the canonical PPR motif (P motif), there are two variants closely related to it, namely the PPR-like S (S motif) and PPR-like L (L motif). P motif is common to all eukaryotes, but S motif and L motif are specific to plants (Andrés et al. 2007). PPR proteins form a huge family widely spread across all eukaryotes. In many fully sequenced and annotated genomes and almost completely annotated genomes, only one PPR-like gene has been identified in prokaryotes of *Ralstonia solanacearum*, which appears to be a fairly recent capture by horizontal gene transfer (Lurin et al. 2004). Although all sequenced eukaryotic genomes contain genes encoding PPR proteins, the numbers of PPR genes in animals and fungi are relatively small. In plants, however, this gene family has undergone dramatic expansion (Geddy and Brown 2007). There are approximately 450 PPR proteins in *Arabidopsis*, 242 of them with only P motifs, whereas the others are predominantly comprised of repeats of P-L-S triplets (PLS subfamily). Almost half of the PPR proteins are predicted to be targeted to mitochondria and one-quarter to plastids (Lurin et al. 2004; Andrés et al. 2007). Structural predictions suggest that PPR motifs consist of two anti-parallel α-helices and presumably tandem arrays would form a superhelix enclosing a groove. The predicted residues projecting into the central groove are almost exclusively hydrophilic and the bottom of the groove is positively charged. This means PPR proteins could interact with RNA and be involved in RNA processing (Small and Peeters 2000; Delannoy et al. 2007). Now more available data support this hypothesis (Schmitz-Linneweber and Small 2008). It is anticipated that PPR proteins are sequence-specific RNA-binding proteins and this is consistent with the idea that they cannot complement each other (Lurin et al. 2004; Delannoy et al. 2007). Though the DYW domain of the PPR protein has been reported to possess endoribonuclease activity, most PPR proteins lack any known catalytic sites (Nakamura and Sugita 2008). So other proteins must be recruited to carry out the RNA processing events in association with the PPR-RNA complex.

Despite the huge number of PPR genes in the *Arabidopsis* genome, genetic analyses reveal a very small functional redundancy among different members (Lurin et al. 2004; Schmitz-Linneweber and Small 2008). Several reports have shown that loss-of-function mutants of PPR genes often exhibit embryo-lethal phenotypes, although at relatively late stages (Tzafrir et al. 2004; Cushing et al. 2005; Kocabek et al. 2006; Schmitz-Linneweber et al. 2006; Chi et al. 2008; Hammani et al. 2011; Lu et al. 2011). PPR proteins are involved in RNA editing, especially in mitochondria and plastids (Meierhoff et al. 2003; Wang et al. 2006; Pfalz et al. 2006; de Longevialle et al. 2007; Tang et al. 2010). As a nucleus-localized PPR protein, mutation of GRP23 causes an arrest of embryo development before dermatogen stage (Ding et al. 2006). GRP23 can physically interact with RNA polymerase II subunit III in both yeast and plant cells (Ding et al. 2006). SIG6, which is a chloroplast sigma factor responsible for the transcription of plastid-encoded RNA polymerase (PEP)-dependent chloroplast genes in cotyledons, is identified through yeast two-hybrid screening to act as a protein partner with the PPR protein DELAYED GREEN1 (DG1)

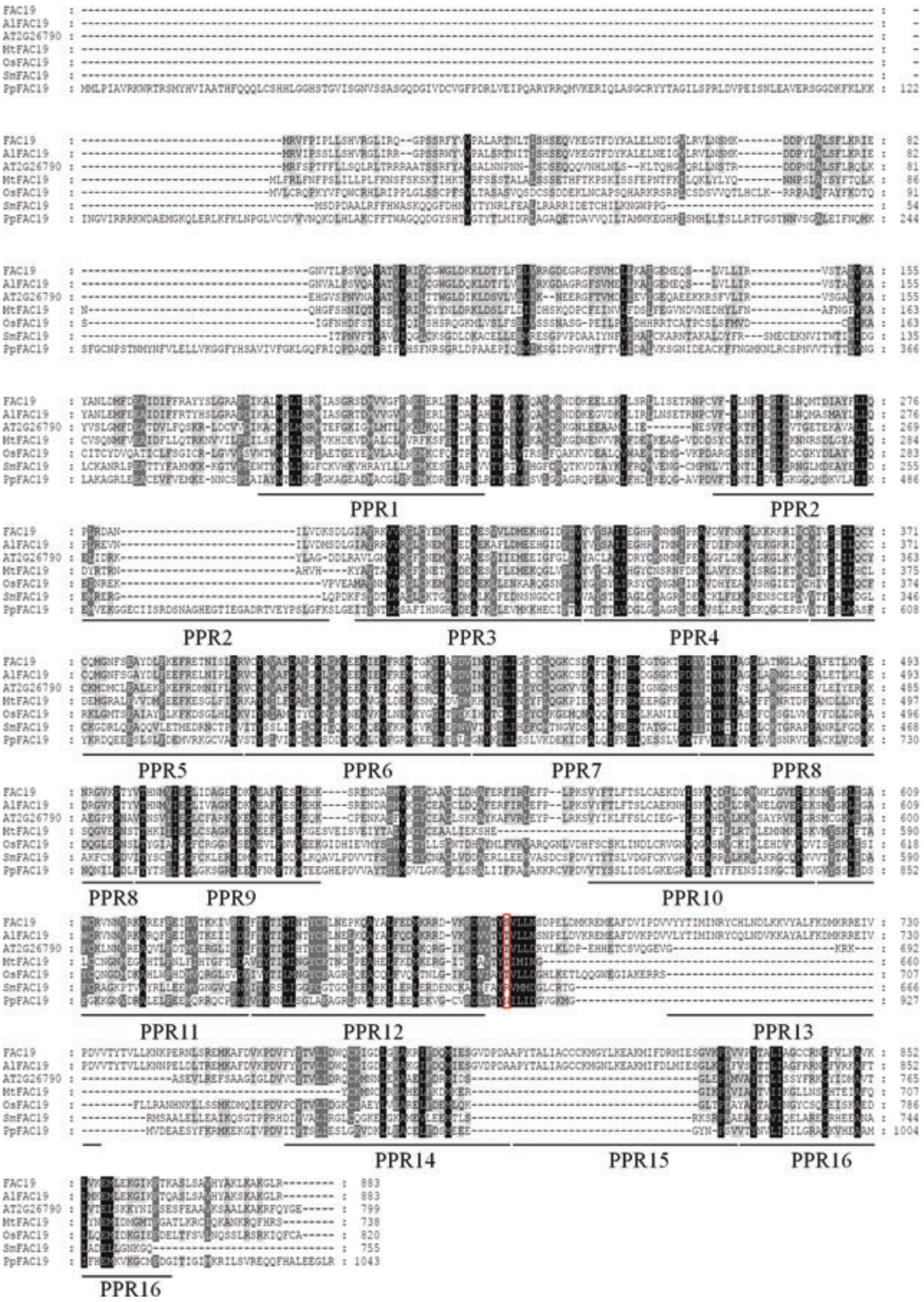

Figure 4. Alignment of FAC19 protein with its closest homologs from *Arabidopsis lyrata* (AlFAC19), *Arabidopsis thaliana* (At2g26790), *Medicago truncatula* (MtFAC19), *Oryza sativa* (OsFAC19), *Selaginella moellendorffi* (SmFAC19) and *Physcomitrella patens* (PpFAC19).

Sixteen pentatricopeptide repeat (PPR) motifs of FAC19 are underlined. The point mutation in *fac19* is marked by a red box. Highly and moderately conserved AAs are highlighted by black and grey boxes, respectively. Note PPR13 is only present in FAC19 in *A. thaliana* and in *A. lyrata*.

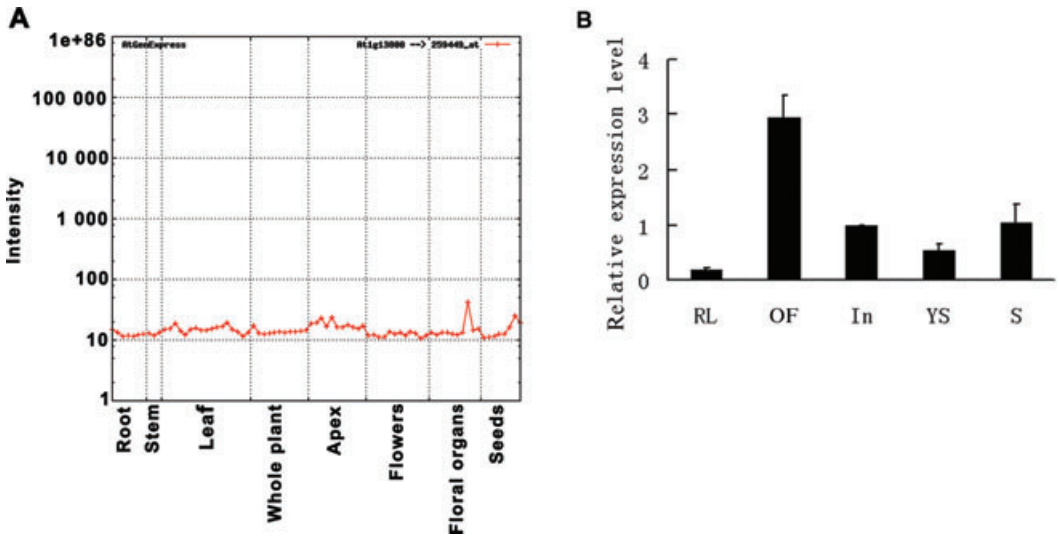

Figure 5. Expressions of *FAC19* in *Arabidopsis*.

(A) Expression profile of *FAC19,* analyzed using AtGenExpress Visualization Tool (http://jsp.weigelworld.org/expviz/expviz.jsp), showing its constitutively low expression during development.

(B) Expressions of *FAC19* in different tissues, revealed by using real-time polymerase chain reaction (PCR) analysis. In, inflorescences; OF, open flowers; RL, rosette leaves; S, seedlings; YS, young silliques.

(Chi et al. 2010). PNM1 in mitochondria is associated with ribosomes in an RNA-dependent manner, while in the nucleus PNM1 interacts with TCP8 and NUCLEOSOME ASSEMBLY PROTEIN1 (NAP1) to control gene expressions (Hammani et al. 2011). All these data seem to support the functions of PPR proteins in transcriptional regulation. FAC19 and three other PPR proteins mentioned above all belong to the P subclass. FAC19 is predicted to be a mitochondria targeted protein. Mutations of *FAC19* led to a zygote-lethal phenotype, which represents the earliest abortion among all PPR mutants known, suggesting its critical role in activation of zygotic embryogenesis. *FAC19* is expressed at a low level in all tissues, with a slightly higher level in open flowers, which may represent its expression in fertilized ovules. Most likely FAC19 is involved in RNA editing for genes essential for executing mitochondrial functions in the zygote, although the exact targets remain to be elucidated.

Materials and Methods

Plant materials and growth conditions

The *Arabidopsis thaliana* L. accessions Col-0 and L*er* were used in all experiments. The *fac19* mutant was identified from an EMS mutagenized population in the Col-0. The T-DNA insertion lines *SALK_012298* and *SALK_143867* in the Col-0 were obtained from the ABRC at Ohio State University. All plants were grown in soil at 21 ± 2 °C in the greenhouse with a 16:8 h light : dark cycle.

Microscopic examinations

Immature silliques from *FAC19*/*fac19* were opened with hypodermic needles and all seeds were excised intact and mounted with the HCG solution. After 15 to 60 min clearing, samples were examined under a Leica florescence microscope equipped with Nomarski optics. For histological analysis, seeds were excised and fixed in modified FAA solution (5 mL of formaldehyde, 5 mL of acetic acid, 5 mL of glycerol and 90 mL of 50% ethanol). Following fixation, materials were dehydrated in a graded ethanol series, infiltrated and embedded in LR White resin, and polymerized at 60 °C for 24 h. Sections were cut at 1 to 2 μm with a glass knife on a Leica microtome, stained with periodic acid-Schiff's reagent (Sigma), and counter-stained with 1% Amido black (Sigma).

Map-based cloning of *FAC19*

Rough mapping was carried out by establishing linkage with markers from different chromosomes in an F2 population generated between wild type (L*er*) and *FAC19*/*fac19* (Col-0) heterozygous plant. Thirty-eight wild type F2 plants were used to localize *FAC19* to chromosome 1 by 15 simple sequence length polymorphism (SSLP) markers distributed evenly across the whole genome (Table S1). For fine mapping, SSLP and single nucleotide polymorphism (SNP) markers were designed based on the sequence difference between Col-0 and L*er* deposited in the Cereon database (http://www.arabidopsis.org, Table S1). By using 950 F2 and 792 F3 segregated wild type

plants, *FAC19* was positioned to a 73 kb region between markers F21F23-3 and F7A19-5.

Genomic DNAs corresponding to the candidate genes were amplified and directly sequenced from both the wild type and *FAC19*/*fac19* heterozygous plants. Allelic tests were carried out by reciprocal crosses between *FAC19*/*fac19* heterozygous plants and SALK T-DNA insertion lines and then F1 seeds were examined under a dissection microscope.

Bioinformatic analysis

Protein targeting was analyzed using the TargetP (http://www.cbs.dtu.dk/services/TargetP/), Predotar (http://urgi.versailles.inra.fr/predotar.html) and Mitoprot (http://ihg.gsf.de/ihg/mitoprot.html) algorithms. The homologous sequences from different organisms were obtained from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). Multiple sequence alignments were generated using the CLUSTALX 1.83. Protein domain analyses were performed with the SMART service (http://smart.embl-heidelberg.de/).

RNA extraction and real-time PCR analysis

Total RNA was extracted from different organs of wild type (Col-0) plants using the TRIzol reagent (Invitrogen), and reverse transcribed to cDNA with a cDNA synthesizing kit (Toyobo) after digestion with RNase-free DNase I (TaKaRa). For real-time PCR analysis, triplicate quantitative assays were performed using the SYBR *Premix Ex Taq* kit (TaKaRa) with a Rotor-Gene 3000 (Corbett Research) detection system and software according to the manufacturer's instructions. The relative expression level was analyzed using the comparative CT method. *ACTIN2* was amplified as an internal control for normalization (Remans et al. 2008, Table S1).

Acknowledgements

We thank Fengqin Dong for technical support in histological analysis, Xiufeng Song, Shipeng Li, Min Chen, Xiuli Lu and Shuqi Qian for their technical help. This work was supported by the National Natural Science Foundation of China (30625018 and 30821007), CAS/SAFEA International Partnership Program for Creative Research Teams (20090491019) and CAS 100-Talent Projects of China.

Received 21 Apr. 2011 Accepted 22 May 2011

References

Andrés C, Lurin C, Small ID (2007) The multifarious roles of PPR proteins in plant mitochondrial gene expression. *Physiol. Plant.* **129,** 14–22.

Chi W, Ma J, Zhang D, Guo J, Chen F, Lu C, Zhang L (2008) The pentratricopeptide repeat protein DELAYED GREENING1 is involved in the regulation of early chloroplast development and chloroplast gene expression in *Arabidopsis*. *Plant Physiol*. **147,** 573–584.

Chi W, Mao J, Li Q, Ji D, Zou M, Lu C, Zhang L (2010) Interaction of the pentatricopeptide-repeat protein DELAYED GREENING 1 with sigma factor SIG6 in the regulation of chloroplast gene expression in *Arabidopsis* cotyledons. *Plant J.* **64,** 14–25.

Cushing DA, Forsthoefel NR, Gestaut DR, Vernon DM (2005) *Arabidopsis emb175* and other *ppr* knockout mutants reveal essential roles for pentatricopeptide repeat (PPR) proteins in plant embryogenesis. *Planta* **221,** 424–436.

de Longevialle AF, Meyer EH, Andres C, Taylor NL, Lurin C, Millar AH, Small ID (2007) The pentatricopeptide repeat gene *OTP43* is required for trans-splicing of the mitochondrial nad1 Intron 1 in *Arabidopsis thaliana*. *Plant Cell* **19,** 3256–3265.

Delannoy E, Stanley WA, Bond CS, Small ID (2007) Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. *Biochem. Soc. Trans.* **35,** 1643–1647.

Ding YH, Liu NY, Tang ZS, Liu J, Yang WC (2006) *Arabidopsis GLUTAMINE-RICH PROTEIN23* is essential for early embryogenesis and encodes a novel nuclear PPR motif protein that interacts with RNA polymerase II subunit III. *Plant Cell* **18,** 815–830.

Fan Y, Jiang L, Gong H, Liu CM (2008) Sexual reproduction in higher plants I: Fertilization and the initiation of zygotic program. *J. Integr. Plant Biol.* **50,** 860–867.

Geddy R, Brown GG (2007) Genes encoding pentatricopeptide repeat (PPR) proteins are not conserved in location in plant genomes and may be subject to diversifying selection. *BMC Genomics* **8,** 130.

Jürgens G (1998) Embryogenesis. In: Westhoff P, Jeske H, Jürgens G, Kloppstech K, Link G, eds. *Molecular Plant Development: From Gene to Plant*. Oxford University Press, New York, USA. pp. 163–176.

Grimanelli D, Perotti E, Ramirez J, Leblanc O (2005) Timing of the maternal-to-zygotic transition during early seed development in maize. *Plant Cell* **17,** 1061–1072.

Hammani K, Gobert A, Hleibieh K, Choulier L, Small I, Giege P (2011) An Arabidopsis dual-localized pentatricopeptide repeat protein interacts with nuclear proteins involved in gene expression regulation. *Plant Cell* **23,** 730–740.

Kocabek T, Repkova J, Dudova M, Hoyerova K, Vrba L (2006) Isolation and characterization of a novel semi-lethal *Arabidopsis thaliana* mutant of gene for pentatricopeptide (PPR) repeat-containing protein. *Genetica* **128,** 395–407.

Li HJ, Liu NY, Shi DQ, Liu J, Yang WC (2010) YAO is a nucleolar WD40-repeat protein critical for embryogenesis and gametogenesis in *Arabidopsis*. *BMC Plant Biol*. **10,** 169.

Lin Z, Yin K, Zhu D, Chen Z, Gu H, Qu LJ (2007) *AtCDC5* regulates the G2 to M transition of the cell cycle and is critical for the function of *Arabidopsis* shoot apical meristem. *Cell Res*. **17,** 815–828.

Liu M, Yuan L, Liu NY, Shi DQ, Liu J, Yang WC (2009) *GAMETOPHYTIC FACTOR **1***, involved in pre-mRNA splicing, is essential for megagametogenesis and embryogenesis in *Arabidopsis*. *J. Integr. Plant Biol*. **51,** 261–271.

Lu Y, Li C, Wang H, Chen H, Berg H, Xia Y (2011) AtPPR2, an *Arabidopsis* pentatricopeptide repeat protein, binds to plastid 23S rRNA and plays an important role in the first mitotic division during gametogenesis and in cell proliferation during embryogenesis. *Plant J*. **67,** 13–25.

Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A, Le Ret M, Martin-Magniette, ML, Mireau H, Peeters N, Renou JP, Szurek B, Taconnat L, Small I (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* **16,** 2089–2103.

Mansfield S, Briarty L (1991) Early embryogenesis in *Arabidopsis thaliana*: II. The developing embryo. *Can. J. Bot.* **69,** 461.

Meierhoff K, Felder S, Nakamura T, Bechtold N, Schuster G (2003) HCF152, an *Arabidopsis* RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbB-psbT-psbH-petB-petD RNAs. *Plant Cell* **15,** 1480–1495.

Meyer S, Scholten S (2007) Equivalent parental contribution to early plant zygotic development. *Curr. Biol.* **17,** 1686–1691.

Nakamura T, Sugita M (2008) A conserved DYW domain of the pentatricopeptide repeat protein possesses a novel endoribonuclease activity. *FEBS Lett*. **582,** 4163–4168.

Ning J, Peng XB, Qu LH, Xin HP, Yan TT, Sun M (2006) Differential gene expression in egg cells and zygotes suggests that the transcriptome is restructed before the first zygotic division in tobacco. *FEBS Lett*. **580,** 1747–1752.

O'Toole N, Hattori M, Andres C, Iida K, Lurin C, Schmitz-Linneweber C, Sugita M, Small I (2008) On the expansion of the pentatricopeptide repeat gene family in plants. *Mol. Biol. Evol.* **25,** 1120–1128.

Pfalz J, Liere K, Kandlbinder A, Dietz KJ, Oelmuller R (2006) pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression. *Plant Cell* **18,** 176–197.

Pillot M, Baroux C, Vazquez MA, Autran D, Leblanc O, Vielle-Calzada JP, Grossniklaus U, Grimanelli D (2010) Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in *Arabidopsis*. *Plant Cell* **22,** 307–320.

Remans T, Smeets K, Opdenakker K, Mathijsen D, Vangronsveld J, Cuypers A (2008) Normalisation of real-time RT-PCR gene expression measurements in *Arabidopsis thaliana* exposed to increased metal concentrations. *Planta* **227,** 1343–1349.

Ronceret A, Gadea-Vacas J, Guilleminot J, Lincker F, Delorme V, Lahmy S, Pelletier G, Chaboute ME, Devic M (2008) The first zygotic division in *Arabidopsis* requires *de novo* transcription of thymidylate kinase. *Plant J*. **53,** 776–789.

Schmitz-Linneweber C, Williams-Carrier RE, Williams-Voelker PM, Kroeger TS, Vichas A, Barkan A (2006) A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast *rps12* pre-mRNA. *Plant Cell* **18,** 2650–2663.

Schmitz-Linneweber C, Small I (2008) Pentatricopeptide repeat proteins: A socket set for organelle gene expression. *Trends Plant Sci*. **13,** 663–670.

Shen WH, Parmentier Y, Hellmann H, Lechner E, Dong A, Masson J, Granier F, Lepiniec L, Estelle M, Genschik P (2002) Null mutation of *AtCUL1* causes arrest in early embryogenesis in *Arabidopsis*. *Mol. Biol. Cell* **13,** 1916–1928.

Small ID, Peeters N (2000) The PPR motif – a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* **25,** 46–47.

Tang J, Kobayashi K, Suzuki M, Matsumoto S, Muranaka T (2010) The mitochondrial PPR protein LOVASTATIN INSENSITIVE 1 plays regulatory roles in cytosolic and plastidial isoprenoid biosynthesis through RNA editing. *Plant J*. **61,** 456–466.

Tzafrir I, Pena-Muralla R, Dickerman A, Berg M, Rogers R, Hutchens S, Sweeney TC, McElver J, Aux G, Patton D, Meinke D (2004) Identification of genes required for embryo development in *Arabidopsis*. *Plant Physiol*. **135,** 1206–1220.

Vielle-Calzada JP, Baskar R, Grossniklaus U (2000) Delayed activation of the paternal genome during seed development. *Nature* **404,** 91–94.

Wang Z, Zou Y, Li X, Zhang Q, Chen L, Wu H, Su D, Chen Y, Guo J, Luo D, Long Y, Zhong Y, Liu YG (2006) Cytoplasmic male sterility of rice with boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. *Plant Cell* **18,** 676–687.

Weijers D, Geldner N, Offringa R, Jürgens G (2001) Seed development: Early paternal gene activity in *Arabidopsis*. *Nature* **414,** 709–710.

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.

(Co-Editor: Wei-Cai Yang)

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Expression profile of *FAC19* from *Arabidopsis* eFP Brower (http://www.bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi).
Table S1. Primer information.