

Daniel A. Cushing · Nancy R. Forsthoefel
Daniel R. Gestaut · Daniel M. Vernon

***Arabidopsis emb175* and other *ppr* knockout mutants reveal essential roles for pentatricopeptide repeat (PPR) proteins in plant embryogenesis**

Received: 7 October 2004 / Accepted: 11 November 2004 / Published online: 13 January 2005
© Springer-Verlag 2005

Abstract Pentatricopeptide repeat proteins (PPRPs) constitute one of the largest superfamilies in plants, with more than 440 identified in the *Arabidopsis thaliana* (L.) Heynh genome. While some PPRPs are known to take part in organelle gene expression, little is known about the broader biological contexts of PPRP gene function. Here, using developmental- and reverse-genetic approaches, we demonstrate that a number of PPRPs are essential early in plant development. We have characterized the *Arabidopsis embryo-defective175* mutant and identified the *EMB175* gene. *Emb175* consistently displays aberrant cell organization and undergoes morphological arrest before the globular-heart transition. The *emb175* mutation disrupts an intronless open reading frame encoding a predicted chloroplast-localized PPR protein—the first to be rigorously associated with an early embryo-lethal phenotype. To determine if other PPRP genes act in embryogenesis, we searched *Arabidopsis* insertion mutant collections for *pprp* knockout alleles, and identified 29 mutants representing 11 loci potentially associated with embryo-defective phenotypes. We assessed gene structures, T-DNA insertion position, and allelism for these loci and were able to firmly establish essential functions for six PPRP genes in addition to *EMB175*. Interestingly, Nomarski DIC microscopy revealed diverse embryonic defects in these lines, ranging from early lethality to dramatic late-stage morphological defects such as enlarged shoot apices and stunted cotyledons. Together, *emb175* and these *pprp* knockout mutants establish essential roles for PPRPs in embryogenesis, thus broadening the known organismal context for PPRP gene function. The diversity of *emb-pprp* knockout phenotypes indicates that mutation of

different PPRPs can, directly or indirectly, have distinct impacts on embryo morphogenesis.

Keywords *Arabidopsis* · Embryogenesis · Knockout mutants · Morphogenesis · Plant development · PPR motif

Abbreviations *EMB*: Embryo-defective · *DIC*: Differential interference contrast · *I-PCR*: Inverse polymerase chain reaction · *ORF*: Open reading frame · *PPR*: Pentatricopeptide repeat · *PPRP*: Pentatricopeptide repeat protein · *RT-PCR*: Reverse-transcription-PCR · *UTR*: Untranslated region

Introduction

Completion of the *Arabidopsis thaliana* genome provided the first full genome sequence of a plant, and comparisons with other genomes have allowed the identification of large gene classes greatly expanded in plants versus other eukaryotes (*Arabidopsis* Genome Initiative 2000). One such class is the pentatricopeptide repeat (PPR) family, encoding proteins containing tandem repeats of a 35-amino acid signature motif that may form a nucleic acid binding groove (Small and Peeters 2000). PPRPs have been identified in all eukaryotes examined, but the family has undergone dramatic expansion in plants. In the *Arabidopsis* genome, more than 440 PPR-encoding genes have been annotated, representing almost 2% of predicted protein encoding genes in this model system (Aubourg et al. 2000; Small and Peeters 2000; Lurin et al. 2004). Numerous PPR-encoding genes are also present in rice, suggesting expansion throughout the angiosperms.

Only a small number of PPRPs from various eukaryotes have been functionally defined. Most are localized to organelles where they have specialized roles in organelle gene expression. Some from fungi, animals, and plants act in mitochondria (Bentolila et al. 2002;

D. A. Cushing · N. R. Forsthoefel
D. R. Gestaut · D. M. Vernon (✉)
Department of Biology, and Program in Biochemistry,
Biophysics and Molecular Biology, Whitman College,
Walla Walla, WA 99362, USA
E-mail: vernondm@whitman.edu
Tel.: +1-509-5275326
Fax: +1-509-5275904

Kazama and Toriyama 2003; Coffin et al. 1997; Manthey and McEwen 1995; Koc and Spremulli 2003). For example, PET309 from yeast and its *Neurospora* counterpart CYA5 are required for post-transcriptional steps of mitochondrial *cox1* expression (Coffin et al. 1997; Manthey and McEwen 1995). In plants, most of the PPRPs that have been functionally characterized were identified through analysis of mutants with high chlorophyll fluorescence or altered seedling pigmentation, and they have highly specific roles in processing or translation of photosynthesis-related plastid transcripts. The first example was Maize CRP1, required for processing and efficient translation of *pet* RNAs encoding photosynthetic electron transport components (Barkan et al. 1994; Fisk et al. 1999). Similarly, HCF152, CRR2 and PGR3 also take part in photosynthesis-related RNA processing in *Arabidopsis* (Hashimoto et al. 2003; Yamazaki et al. 2004; Meierhoff et al. 2003). Maize PPR2 has a broader impact, being essential for plastid ribosome accumulation (Williams and Barkan 2003). Recently, a genomic study of *Arabidopsis* PPRPs has provided further support for the view that many of these proteins function in organelles (Lurin et al. 2004).

There is evidence that some PPRPs may have roles beyond organelle gene expression. A notable example is *Drosophila* BSF, which contributes directly to early embryonic patterning, binding to the *Bicoid* mRNA and stabilizing this key anterior determinant (Mancebo et al. 2001). Also, PPRPs with DNA-binding activity have been identified in both animals and plants. Wheat p63, a DNA-binding protein involved in mitochondrial transcription (Ikeda and Gray 1999), is a PPR. In mammals, LRP130/LRPPRC protein localizes to both the nucleus and cytoplasm, binds mini-satellite sequences, and may be involved in processes as diverse as vesicular trafficking, chromosome remodeling, and cytokinesis (Tsuchiya et al. 2002; Liu and McKeehan 2002). Thus, PPR proteins may have diverse cellular and developmental functions in a wide range of eukaryotes, including plants.

Plants contain far more PPRP genes than other eukaryotes, but the functional significance of the explosive expansion of the PPRP family in plant genomes is not known, and the developmental contexts in which these genes act have not been defined. Here, we provide developmental- and reverse-genetic evidence that a number of PPR proteins have essential, non-redundant roles in plant embryogenesis. We describe molecular and phenotypic characterization of the *Arabidopsis embryo-defective175* mutant, identifying the *EMB175* product as a PPR protein, the first rigorously shown to be essential in early development. We then extend these findings through the analysis of T-DNA knockout mutants, identifying six more essential PPRP loci. These mutants exhibit diverse defects in embryo morphology and timing of developmental arrest, revealing that disruption of PPRP genes can have diverse and dramatic morphological consequences in the *Arabidopsis* embryo.

Materials and methods

Plant material

Emb175 seeds were obtained from self-fertilized *emb175* heterozygotes originally isolated by the laboratory of D. Meinke (Oklahoma State University; Errampali et al. 1991). Seeds for *emb175* allelic lines (*emb1899-1* and *1899-2*) and other *emb*-PPRP knockout lines were obtained from the *Arabidopsis* Biological Resource Center (Ohio State University). *Emb175* plants used for inverse-PCR were grown from seed surface-sterilized and germinated in culture as previously described (Vernon and Meinke 1995), in the presence of 50 µg/ml kanamycin. Plants used for complementation crosses, embryo microscopy, and reverse transcription-polymerase chain reaction (RT-PCR) were seeded in soil, chilled at 4°C for 3 days, and grown in growth chambers at 22°C(day)/16°C(night) under cycles of 16 h light/8 h dark. Wild types and *emb* heterozygotes were distinguished among soil-grown plants by screening for abnormal seeds in siliques of self-fertilized plants (Meinke 1995).

Genetic and phenotypic characterization

For genetic and microscopy experiments, embryo-defective phenotypes were analyzed in seeds produced by selfed *emb* heterozygotes, or among F1 seeds produced in complementation crosses. Siliques were examined by dissection microscope to score or obtain abnormal seeds. Complementation tests were carried out by reciprocal crosses between plants heterozygous for putatively allelic mutations. Embryo phenotypes were visualized as described by Vernon and Meinke (1994), using an Olympus BX60 microscope equipped with Nomarski DIC optics. Images were captured using a Coolsnap digital camera and software (RS Photometrics), or by an Olympus SC35 with Kodak TMAX 100 film.

Inverse PCR

For identification of the T-DNA flanking region, genomic DNA was isolated from *emb175* heterozygotes using the procedures of Castle et al. (1993), followed by further purification by an additional 1:1 phenol:chloroform extraction and ethanol precipitation. Ten micrograms of genomic DNA was digested with 50 U of *HindIII* for 4 h at 37°C, with 30 units of additional enzyme added after 1 h to complete digestion. Enzyme was heat-inactivated (20 min, 70°C) and digestion verified by 1% agarose gel. Digested DNA was ligated in a dilute reaction to favor intramolecular circularization: 3 µg of digested genomic DNA, 50 µl of T4 ligation buffer, 10 µl T4 Ligase, and dH₂O to a final volume of 500 µl, incubated for 16 h at 16°C. Circularized ligation products

were subjected to inverse PCR with primers specific for sites within the T-DNA insert between the left-border (LB) terminus and the left-most *Hind*III site approximately 3 kbp inside the LB of the T-DNA (Castle et al. 1993). Primers were designed using Primer3 [http://www.genome.wi.mit.edu/genome_software/other/primer3.html]: 5'-TCTGGGAATGGCGTAACAAAGGC-3' (outward from T-DNA LB); 5'-ACGTTTTTCGCTGTCGGCAGATG-3' (inward toward T-DNA *Hind*III site). Six PCR reactions were performed using from 4 µl to 0.04 µl of circularized DNA mixture, with ExTaq DNA Polymerase (TaKaRa Bio Inc., Otsu, Japan) with reaction components supplied by the manufacturer. Reactions were carried out for 36 cycles (94°C for 45 s, 58°C for 60 s, 72°C for 120 s). Products were isolated with QIAquick Gel Extraction kits (Qiagen, Valencia, CA, USA) and sequenced.

Gene identification and molecular analysis

Junction sequences between T-DNA LB and right borders (RBs) and adjacent genomic DNA were amplified from individual *emb175* heterozygotes, using combinations of gene- and T-DNA-specific primers in the PCR conditions outlined above, but with an annealing temperature of 61°C. Primers for LB junction: *EMB175* reverse primer (5'-CGTTCTTCATAACCCGAACCGG-3') and T-DNA LB primer (5'-TCTGGGAA-TGGCGTAACAAAGGC-3'). For the RB junction: *EMB175* forward primer (5'-CCTTCAATTCCTCCGAACATCG-3') and the T-DNA RB primer (5'-GGACACCTACGGTCAAGGGAG-3'). The *EMB175* forward and reverse primers mentioned above were used to amplify the wild-type gene in control reactions. PCR products were visualized on 1% agarose gels, purified by Qiaquick extraction, and sequenced to confirm product identity.

All sequencing was done by the University of Arizona DNA sequencing core facility (Tucson, AZ, USA). DNA sequences comparisons were done by BLAST (Altschul et al. 1997) at Genbank [http://www.ncbi.nlm.nih.gov] and TAIR [http://www.Arabidopsis.org].

For protein predictions, DNA sequences were translated using MacVector 6.0 software (Oxford Molecular). Domain and motif searches were performed using PFAM [http://www.sanger.ac.uk/Software/Pfam/ (Bateman et al. 2004)] and InterProScan [http://www.ebi.ac.uk/InterProScan/ (Mulder et al. 2003)].

EMB175 and *MED24-9* transcription units were defined by RT-PCR using primers specific to exon regions of each hypothesized gene. The following primer combinations were used on cDNA populations isolated from both leaf and root tissues: *MED24-10* forward primer (5'-CATTTCGTCGGTGTCTTGGTTCAT-3') and *MED24-10* reverse primer (5'-CGTTCTTCATAACCCGAACCGG-3'); *MED24-9* forward primer (5'-CGTCTCT-TCTACGGCATCATTGG-3') and *MED24-9*

reverse primer (5'-CCACCGACCAATC-CAGTTAA-GGA-3').

To define termini of other prospective *EMB*-PPR genes, EST and/or full-length cDNA sequences were obtained by BLAST searches using query sequences derived from genomic DNA representing prospective *EMB* genes and putative neighboring loci. Sequences were downloaded from Genbank for ORF analyses, alignments, and identification of untranslated regions using MacVector 6.0 software (Oxford Molecular).

Reverse-transcription PCR

For RNA expression surveys, flower, leaf, and root issues of adult wild-type plants were ground with a mortar and pestle in liquid N₂ and RNA was isolated using a Clontech Nucleobond kit (Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions. Yield was estimated by spectrophotometry (Sambrook et al. 1989). Purification of polyA RNA was performed with Clontech Nucleotrap mRNA Mini kits. From each tissue, 0.1 µg of polyA RNA was reverse-transcribed using Clontech Advantage RT-for-PCR kits, following the manufacturer's instructions. Absence of genomic DNA in RNA preparations was established by control PCR without prior reverse transcription. RT-PCR was carried out with 1 µl of cDNA products using the PCR program described above for inverse PCR, but with an annealing temperature of 61°C, using the following primers: forward primer (5'-CATT-CGTCCGTGTTCTTGGTTCAT-3'); reverse primer (5'-CGTTCTTCATAACCC-GAACCGG-3').

Results

Emb175: a T-DNA tagged mutant that arrests early in embryogenesis

Embryo-defective175 (*emb175*) was originally isolated from the Feldmann T-DNA insertion mutant population (Forsthoefel et al. 1992), as part of a large collection of embryo-defectives identified on the basis of segregation of defective seeds in siliques of self-fertilized heterozygotes (Errampalli et al. 1991). The *EMB175* locus had been mapped to upper chromosome V (Franzmann et al. 1995). Previously published genetic and Southern blot analyses had established that the *emb175* line contains a single, recessive mutation associated with a T-DNA insert (Errampalli et al. 1991; Castle et al. 1993).

We used Nomarski DIC microscopy of cleared whole-mount seeds to investigate the *emb175* phenotype at different developmental stages. Wild type and *emb175* embryogenesis are compared in Fig. 1. Homozygous mutant embryos morphologically arrested at the globular-heart transition, consistently failing to initiate cotyledons and make the switch from radial to bilateral symmetry (West and Harada 1993; Goldberg et al.

1994). Despite morphological arrest, mutant embryos continued cell division and embryo enlargement to some extent, continuing to increase in complexity until siliques reached the linear or early cotyledon stages of development (Fig. 1 panels h, i, j). Mutant embryos from late-stage siliques contained bloated cells resulting in a bumpy protoderm similar to that observed in *Arabidopsis raspberry* (*rsy*) mutants (Yadegari et al. 1994; Apuya et al. 2002). In contrast to many *Arabidopsis* embryo-defective mutants that arrest morphologically at the globular stage (e.g., Schwartz et al. 1994; Yadegari et al. 1994; Apuya et al. 2002), mutant suspensors in *emb175* typically retained their wild-type morphology until siliques neared maturity, suggesting that embryo-suspensor interactions required for suspensor maintenance remain intact until late in seed development (Schwartz et al. 1997; Vernon and Meinke 1994). Suspensor proliferation and enlargement were observed at a low frequency in some mutant seeds from siliques nearing maturity (Fig. 1j).

Failure of morphogenesis in *emb175* coincided with irregularities in cell division and enlargement patterns. These were often evident as localized asymmetry within

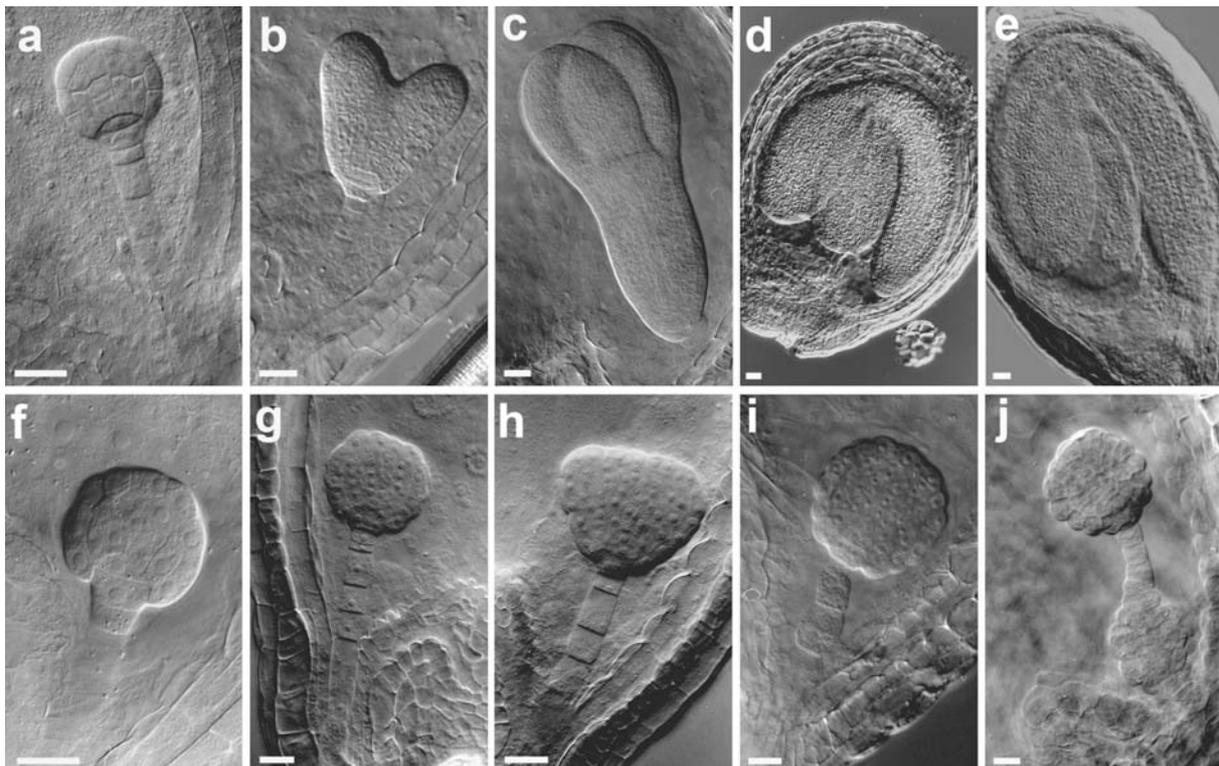
the focal plane, most easily observed in heart stage siliques following morphological arrest. Representative mutant embryos with asymmetric cell organization are shown in Fig. 2. The variable sites and timing of such local irregularities suggested stochastic, cell autonomous defects in cell division rate and orientation in *emb175* homozygotes.

Identification of the *EMB175* gene

To identify the *EMB175* gene, we isolated sequences adjacent to the *emb175* T-DNA tag, using inverse PCR (I-PCR) on genomic DNA from *emb175* heterozygotes. I-PCR with primers specific to the T-DNA LB generated a single product containing T-DNA LB sequence and 170 bp of plant DNA, linked by 5 bp of filler DNA, a common feature of T-DNA:plant junctions (Tax and Vernon 2001). I-PCR results and junction sequence are illustrated in Fig. 3a. BLAST alignments with the *Arabidopsis* genome indicated that the tagged plant sequence was derived from upper chromosome V, consistent with the previously determined *emb175* map position. Flanking sequence suggested the T-DNA was situated within a large open reading frame, MED24-10 (Fig. 3b).

The position of the *emb175* T-DNA within the MED24-10 ORF was verified by PCR amplification of both LB and RB T-DNA:plant junction sequences from *emb175* heterozygotes. Primers specific to predicted flanking regions of MED24-10, used in conjunction with T-DNA LB- or RB-specific primers (see Fig. 3b), successfully amplified both predicted LB and RB plant:T-DNA junction sequences (Fig. 3c). PCR product

Fig. 1 Developmental profiles of wild-type and *emb175* mutant embryos. Seeds were harvested at various stages of development, cleared, and viewed by Nomarski DIC microscopy. **a–c** Wild type embryos at the globular, heart, torpedo, early cotyledon, and mature cotyledon stages of embryogenesis, respectively. **f–i** Homozygous *emb175* embryos at ages corresponding to the wild-type morphological stages in the top row. **j** An *emb175* homozygote from mature siliques, illustrating suspensor cell proliferation observed at a low frequency in embryos at this stage. Scale bars = 20 μ m



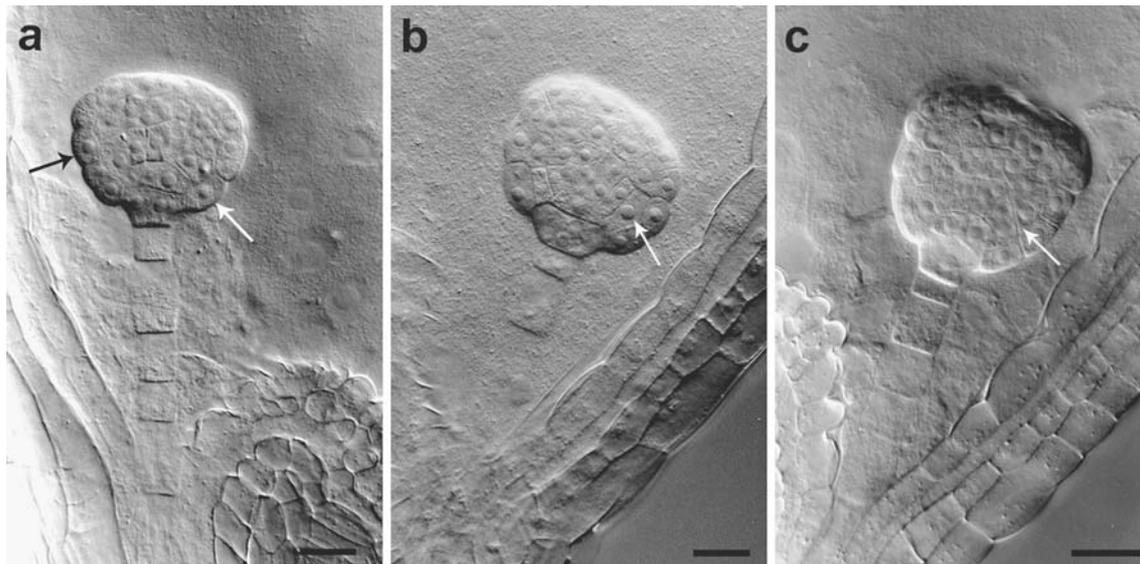


Fig. 2 Examples of localized asymmetry in morphologically-arrested *emb175* embryos. Seeds containing *emb175* embryos were harvested from late heart stage siliques and viewed in a longitudinal focal plane as whole mounts by Nomarski DIC microscopy. **a, b** Arrows indicate enlarged cells that are not dividing in synchrony with cells in corresponding positions on the other side of the embryo. **c** Arrow indicates large distorted internal cell with aberrant division plane. Scale bars = 20 μ m

identities were confirmed by sequencing. The *emb175* T-DNA had inserted 1,088 bp into the MED24-10 ORF, and that insertion was accompanied by a 23 bp deletion in the MED24-10 ORF and insertion of 5 bp of “filler” sequence. The presence of these plant/T-DNA junction fragments co-segregated with the production of *embryo-defective* seeds following self-fertilization, as expected based on prior Southern blot analyses of *emb175* heterozygotes with T-DNA probes (Castle et al. 1993).

To further confirm that the T-DNA-tagged MED24-10 ORF corresponds to the *EMB175* gene, we characterized two independent mutant alleles of this locus, *emb1899-1* and *1899-2*, which became available through the SeedGenes project (McElver et al. 2001; <http://www.seedgenes.org>). Locations of the T-DNA in these alleles are shown in Fig. 3d. Nomarski microscopy revealed that they had embryo phenotypes indistinguishable from that of *emb175* (data not shown). Complementation crosses between both *emb1899* mutants and *emb175* demonstrated that these mutations were indeed all allelic, confirming that disruption of the locus identified in *emb175* is responsible for the embryo-defective phenotype. All of these are likely null alleles, based on T-DNA location and small out-of-frame deletions adjacent to inserts (Fig. 3d).

Determining the correct structure of the *EMB175* locus

To make meaningful interpretations of *EMB175* sequence, we first needed to determine the correct gene

structure. The *Arabidopsis* genome annotation offered two very different competing annotations of the tagged *EMB175* locus, illustrated in Fig. 4. One predicted the gene to consist solely of the MED24-10 ORF, with a separate neighboring locus, MED24-9 approximately 1.3 kbp downstream (Fig. 4a). A more recent annotation by TIGR, At5g03800, predicted a single, large multi-exon locus consisting of both MED24-10 and MED24-9, plus an additional exon from within the MED24-9/10 intergenic region. We used RT-PCR and available EST sequences to distinguish between these models. RT-PCR with the primer pairs shown in Fig. 4a successfully detected RNA species corresponding both to MED24-9 and MED24-10 regions (Fig. 4b). However, no RNA bridging both of these regions was detected. Our inability to amplify the hypothetical chimeric MED24-9/10 transcript was not due to inadequate reverse transcription, because each of these regions was successfully detected separately using MED24-9- and MED24-10-specific primers (Fig. 4b).

Later in the course of this work EST sequences became available for the At5g03800 region: AU229134/RAFL16-75-K17 from MED24-10, and BE524816 and AV555272 from the neighboring MED24-9 locus. cDNA AU229134/RAFL16-75-K17 confirmed the boundaries of the MED24-10 gene prediction. Both this cDNA and BE524816 contained sequences from downstream of the MED24-10 ORF predicted by the At5g03800 annotation to be intronic, indicating that the TIGR At5g03800 annotation (second map, Fig. 4a) was incorrect. Furthermore, end regions of ESTs contained non-sense codons in-frame with adjacent coding sequences, confirming that these ESTs defined the terminal UTRs of what must be two separate mRNAs. Taken together, these RNA-derived sequences clearly established that *EMB175* consists of a single exon (MED24-10), containing an ORF of 2,688 bp, located only ~400 bp upstream of a transcription unit consisting of

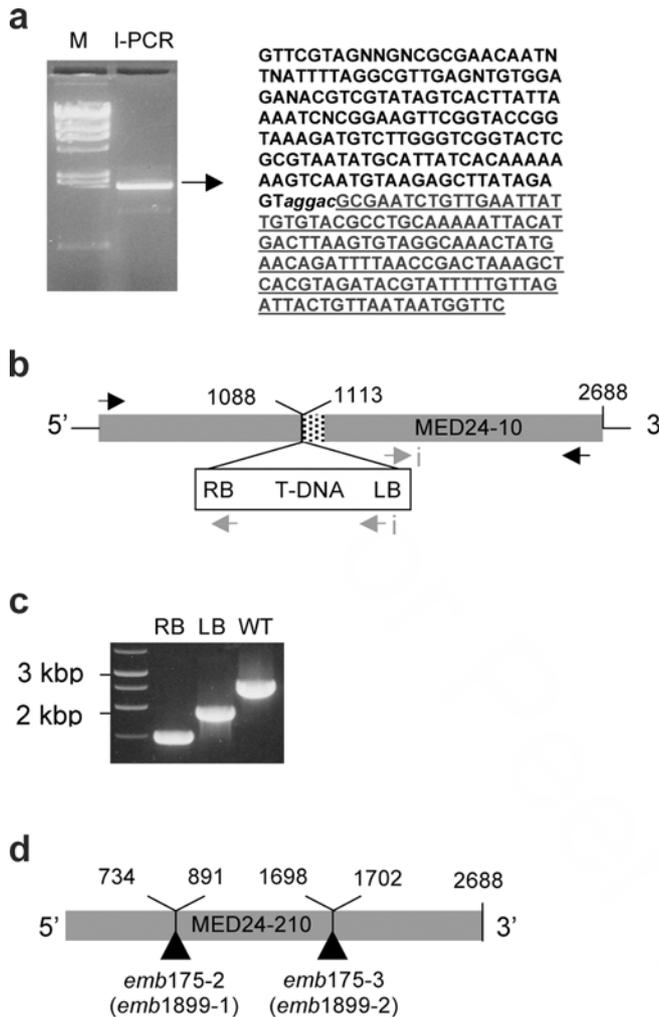


Fig. 3 Inverse PCR and identification of the tagged *emb175* locus. I-PCR was carried out on *Hind*III-digested, recircularized genomic DNA from *emb175* heterozygotes. **a** Agarose gel of I-PCR product, with marker (*M*) in left lane, and T-DNA:plant junction sequence obtained from it. *Underlined* Sequence corresponding to the T-DNA left border. *Italics* 5 bp filler sequence at insertion site. *Plain type* Flanking sequence from upper chromosome 5. **b** Diagram of T-DNA insertion in the MED24-10 ORF showing T-DNA orientation and position. Nucleotide positions adjacent to the T-DNA are indicated (*RB* and *LB*, respectively). *Stippled region* LB flanking region rescued by I-PCR (**a**). *Arrows* indicate positions of plant-specific (*black*) and T-DNA-specific (*gray*) PCR primers used in I-PCR (*i*) and subsequent experiments. **c** Confirmation of T-DNA location by PCR of predicted LB and RB T-DNA junction fragments. PCR with T-DNA- and gene-specific primer combinations was carried out on genomic DNA isolated from *emb175* heterozygotes. Product sizes correspond to those predicted from the map in (**b**); product identities were confirmed by sequencing. *RB* Right border junction product obtained with 5' gene-specific and RB-specific primers. *LB* Left border junction fragment obtained with 3' gene-specific and LB primers. *WT* Wild-type gene fragment amplified by 5' and 3' GSPs. Marker sizes are provided at left for reference. **d** Two other *emb175* T-DNA insertion alleles containing insertions in the MED24-10 ORF. T-DNA positions are indicated by *black triangles*; numbers indicate nucleotide positions at T-DNA borders

the predicted MED24-9 exons plus an additional 5' exon (Fig. 4a, bottom map).

EMB175 encodes a predicted plastid-localized PPR protein

EMB175 predicted a soluble primary protein product consisting of 896 amino acids, containing no trans-membrane domains. The amino acid sequence and key structural features are shown in Fig. 5. The most prominent feature was a large central domain of 14 internal pentatricopeptide motifs (some degenerate), arranged in tandem. Several other features of *EMB175* resemble features of many other plant PPR proteins. TargetP and Predotar programs suggest the protein is targeted to plastids and contains an N-terminal transit sequence, consistent with the organellar location of most previously identified PPR proteins from plants and fungi (Small and Peters 2000). Also, *EMB175* contains a potential heme-binding site near the C-terminus. The C-terminal amino acids are DLW, an apparent derivation of the DYW C-terminal motif found on many plant PPRPs (Aubourg et al. 2000).

EMB175 is transcribed during post-embryonic development

We carried out a qualitative RT-PCR survey of *EMB175* expression in leaves, roots and flowers. Results are shown in Fig. 6. Transcripts were detected in all of these organs, suggesting that *EMB175*'s function in wild-type plants is not restricted to embryogenesis.

Identification of other PPRP genes essential for embryogenesis

The *EMB175* gene provided the first example of a plant PPRP gene essential for embryogenesis. To extend these findings and determine if other plant PPRPs influence embryogenesis, we searched available mutant collections for knockout alleles of PPRP genes putatively associated with an embryo-defective phenotype. The SeedGenes collection, a set of prospective T-DNA-tagged *embryo-defectives* cataloged in a searchable database, provided a starting point (McElver et al. 2001; Tzafrir et al. 2004). Prospective allelic lines were then identified from the sequence-indexed Salk T-DNA mutant collection (Alonso et al. 2003). Table 1 lists putative *pprp* knockout mutants tentatively associated with seed defects by the SeedGenes database, and candidate second alleles of these loci. In total, in addition to *emb175*, we identified 29 insertion mutants representing 11 candidate

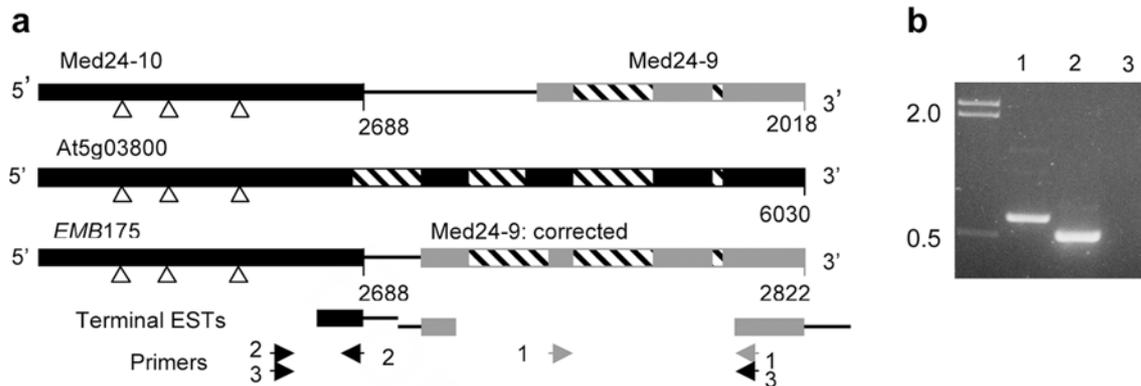


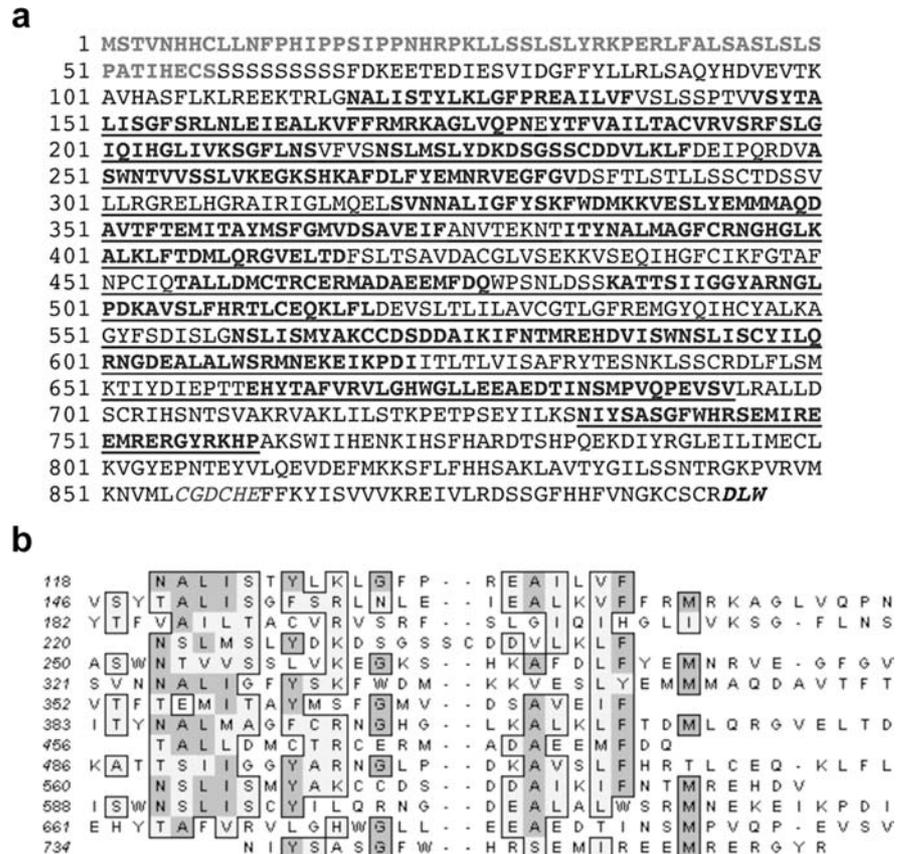
Fig. 4 Determination of the correct structure of the *EMB175* locus. **a** Diagrams of various annotations of the *EMB175* region. *Hollow triangles* mark the positions of T-DNA inserts in various *emb175* alleles; all are in the 5'-most ORF. *Top* Kazusa annotation consisting of MED24-10 and predicted downstream locus Med24-9. *Middle* Large multi-exon gene predicted by TIGR. *Bottom* correct gene maps derived from RT-PCR results (**b**), and 5' and 3' EST data. Positions of ESTs defining 5' and 3' ends of separate transcripts from this region are shown below (see Results). *Arrows* Positions of primers used for RT-PCR shown in **b**; numbers correspond to gel lanes in **b**. **b** RT-PCR of separate *EMB175* and MED24-9 transcripts, but not hypothetical chimera predicted by TIGR annotation. cDNA fragments were amplified from total RNA isolated from wild-type seedlings. *Lane 1* MED24-9 RT-PCR product; *lane 2* *EMB175*/MED24-10 product; *lane 3* lack of product from primer combination 3 (see **a**)

PPRP genes with potential roles in embryogenesis. As described below, we further characterized these loci by investigating gene structures, T-DNA insert positions, phenotype segregation, and allelism to confirm (or rule-out) the association of *PPRP* loci with embryo-defective phenotypes. For convenience in the following sections, we refer to these mutants and the corresponding loci as *emb-PPRPs*.

Confirmation of embryo-defective phenotypes in six *pprp* knockout lines

Due to the large scope of the SeedGenes project, it has been acknowledged that developmental phenotypes have

Fig. 5 Features of the predicted *EMB175* protein. **a** Predicted amino acid sequence. *Bold* Predicted N-terminal chloroplast transit sequence. *Underlined* PPR domain, with PPR unit motifs in *bold*. *Italics* Predicted heme-binding site, and C-terminal DLW motif, both found in other plant PPR proteins. **b** ClustalW alignment of PPR unit motifs, with conserved residues *shaded*



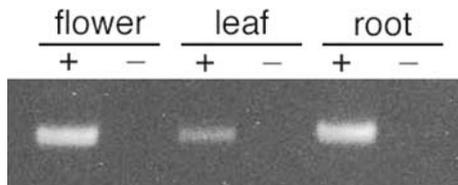


Fig. 6 Detection of *EMB175* transcript in adult tissues. RT-PCR with *EMB175*-specific primers was carried out with polyA RNA isolated from the indicated organs from mature plants. Mock reactions without reverse transcriptase were carried out to demonstrate lack of DNA contamination, because there is no intron that would allow for distinction between and mRNA-derived and contaminating DNA-derived products. Lanes: +, RT-PCR products; –, mock reaction products

not been rigorously assigned to prospective T-DNA tagged genes (Tzafrir et al. 2004). Furthermore, the identification of tagged loci in the SeedGene and Salk mutant collections is based on hypothetical gene annotations rather than empirically defined gene structures. Therefore, we viewed the SeedGenes collection only as a starting point for identification of *PPRP* genes for more detailed study, including genetic confirmation of embryo-defective phenotypes, definition of *PPRP* gene boundaries, and verification of T-DNA insertion status. Using this strategy on the eleven prospective *emb-PPRP* loci listed in Table 1, we were able to confirm the association between embryo defects and *emb-pprp* knockout alleles for six *PPRP* genes, which are listed in Table 2.

We assessed *emb* phenotype segregation in each prospective *pprp* knockout line and carried out complementation crosses between putative allelic lines. Segregation analysis confirmed that all T-DNA insertion alleles for these six loci produced *emb* phenotypes: progeny produced by selfed heterozygotes for all mutant alleles in Table 2 segregated for ~25% abnormal seeds

Table 1 Additional candidate Arabidopsis *PPRP* genes essential for embryogenesis

| AGI gene ID ^a | SeedGenes mutant ^b | Putative Salk T-DNA alleles ^c |
|--------------------------|-------------------------------|--|
| At3g18110 | <i>emb1270</i> | SALK 027171SALK 027183 |
| At1g06150 | <i>emb1444</i> | SALK 093892 |
| At3g49240 | <i>emb1796</i> | SALK 069042 |
| At3g49170 | <i>emb2261</i> | SALK 025014SALK 024975SALK 024582 |
| At4g39620 | <i>emb2453</i> | SALK 023575SALK 071838 |
| At4g20090 | <i>emb1025</i> | SALK 142675 |
| At5g50280 | <i>emb1006</i> | SALK 073046 |
| At1g30610 | <i>emb2279</i> | SALK 088420SALK 086107 |
| At5g39680 | <i>emb2744</i> | SALK 006056SALK 119171 |
| At5g39710 | <i>emb2745</i> | None available |
| At1g12770 | <i>emb1586</i> | SALK 134436 |

^a Loci as designated by AGI annotation [http://www.Arabidopsis.org]

^b Mutants from SeedGenes database [http://www.seedgenes.org] with T-DNA inserts tentatively assigned to *PPRP* loci (McElver et al. 2001; Tzafrir et al. 2004)

^c Putative allelic mutant lines from the Salk sequence-indexed insertion mutant collection (Alonso et al. 2003)

Table 2 Complementation crosses confirming the association between *PPRP* knockout mutations and embryo-defective phenotypes

| <i>PPRP</i> gene | Cross | Predicted <i>emb</i> /no. scored ^a | Observed <i>emb</i> (%) |
|------------------|------------------------------|---|-------------------------|
| At3g18110 | <i>emb1270</i> × SALK 027183 | 41/164 | 39/164 (24%) |
| At3g49240 | <i>emb1796</i> × SALK 069042 | 64/253 | 65/253 (26%) |
| At3g49170 | <i>emb2261</i> × SALK 024975 | 39/155 | 43/155 (28%) |
| At4g20090 | <i>emb1025</i> × SALK 142675 | 75/298 | 68/298 (23%) |
| At4g39620 | <i>emb2453</i> × SALK 071838 | 68/275 | 68/275 (25%) |
| At1g30610 | SALK 088420 × 086107 | 16/62 | 15/62 (24%) |

^a25% of F1 progeny were predicted to be homozygous recessive and exhibit an embryo-defective seed phenotype, if defects were caused by allelic *pprp* knockouts

following self-fertilization (data not shown). As shown in Table 2, complementation crosses between allelic heterozygotes produced approximately 25% abnormal F1 progeny, confirming that the *pprp* knockout mutation in each of these lines was responsible for the observed embryo-defective phenotype.

To confirm that T-DNA inserts in the candidate *pprp* insertion lines were actually situated within *PPRP*-encoding transcription units, *PPRP* gene boundaries were defined using cDNA and available EST sequence data (as was done with *EMB175*). For the six *PPRP* loci listed in Table 2, T-DNA inserts were situated within exons of *PPRP* transcription units. Furthermore, T-DNAs interrupted *PPR*-encoding or 5' transcribed regions of each gene in all but the *emb1796* alleles (in which both T-DNAs resided near the 3' end of the coding region). Thus, these mutants likely harbored true *pprp* knockout alleles.

For the remaining five candidate *emb-PPRP* loci listed in Table 1, we ruled out, or were unable to confirm, the association between *PPRP* gene disruption and an embryo-defective phenotype. These loci are listed in Table 3. For *emb1444* and *emb1586*, terminal EST sequences revealed that tagged loci had been misannotated and T-DNA inserts did not actually reside within *PPRP* transcription units, thus suggesting that phenotypes in these lines were not likely due to *PPRP* disruption. There was also apparent complementation between different *emb1444* alleles, further suggesting that the *PPRP* gene at that locus may not actually be responsible for the seed defect in the SeedGenes line. For three putative *emb-PPRPs*—*emb1006*, *emb2744*, and *emb2745*—we were unable to confirm embryo-defective phenotypes, due to phenotype complementation between alleles, a lack of phenotype in the candidate second allele, or the lack of an available allelic line (Table 3). These results underscore the view that functional assignments generated by large-scale gene-tagging projects are preliminary—perhaps more so than is generally acknowledged. Nevertheless, the successful identifica-

Table 3 *PPR* loci implicated in embryogenesis for which developmental roles could not be confirmed

| AGI gene ID | SeedGenes mutant ^a | Putative additional alleles ^b | Reason for database misdesigniation |
|-------------|-------------------------------|--|---|
| At1g06150 | <i>emb1444</i> | SALK 093892 SALK 143087 | Annotation error in databases; T-DNAs outside coding region; apparent complementation between alleles |
| At1g12770 | <i>emb1586</i> | SALK 134436 | Annotation error in databases; T-DNA inserts in unrelated adjacent gene |
| At5g50280 | <i>emb1006</i> | SALK 073046 | No <i>emb</i> phenotype associated with SALK insertion allele |
| At5g39680 | <i>emb2744</i> | SALK 006056 SALK 119171 | T-DNA insertion status of SALK alleles not clear; apparent T-DNA-associated translocation in one line |
| At5g39710 | <i>emb2745</i> | none available | No second allele yet identified |

^aSeedGenes database at <http://www.seedgenes.org> (McElever et al. 2001; Tzafrir et al. 2004)

^bSalk sequence indexed insertion lines (Alonso et al. 2003)

tion of six additional *emb-pprp* mutants indicated that in addition to *EMB175*, other members of the *PPRP* family are essential in plant development.

The *emb-pprp* knockout mutants display dramatic and distinct phenotypes

To assess the impact of these different *pprp* knockout mutations on embryogenesis, we carried out a preliminary survey of mutant phenotypes for five *PPRP* genes unambiguously associated with embryo-defective phenotypes: *embs1025*, 1270, 1796, 2261, and 2453. These genes all encode predicted organelle-localized proteins, but otherwise represent different subfamilies from within the PPR superfamily (Lurin et al. 2004). Table 4 lists locus accession numbers and summarizes the predicted cellular locations and PPR domain features of each *emb-PPRP* subjected to phenotype characterization. The diagrams in Fig. 7 further illustrate the size and PPR domain diversity among these proteins.

Figure 8 shows examples of mutant embryos from cotyledon stage siliques, illustrating the terminal morphological phenotypes resulting from *PPRP* gene disruption in each of these loci. In assessing these

mutants as a group and comparing their phenotypes, we noted two important features: (1) each *emb-pprp* mutation had consistent effects on embryo morphology; and (2) each mutation impacted embryo morphology differently.

While the *emb175* mutation had caused morphological arrest at the globular/heart transition, resulting in a “globular” phenotype common among *Arabidopsis* embryo-defectives (Fig. 1), such was not the case for most of the other *emb-pprp* mutants. Of these, only *emb1796* resulted in failure to develop beyond globular stage. *Emb1796* exhibited consistent and severe developmental arrest at the mid-globular stage, resulting in a slightly irregular embryo proper consisting of 32–64 cells (Fig. 8a). In contrast to *emb175*, arrested embryos persisted in this early arrested state into late-stage seeds, without any apparent further growth.

In contrast, *embs1025*, 1270, 2261, and 2453 consistently initiated cotyledons and continued growth and cell division beyond heart stage (Fig 8b–e). However, while these mutants were able to initiate cotyledons, they were often asymmetrical and severely stunted, and mutant embryos failed to progress to a proper curled cotyledon or “walking stick” morphology. These mutants displayed distinct morphological differences. *Emb2453*

Table 4 Predicted location and PPR domain features of *emb-PPRP* gene products subjected to knockout phenotype characterization

| Emb-PPRP (locus ID) | Predicted location ^a | Length, PPR motif number and PPR distribution ^b |
|---------------------|--|--|
| EMB175 (At5g03800) | Chloroplast | 896 aa, 14 PPR motifs (four degenerate) |
| EMB1025 (At4g20090) | Chloroplast (T) or mitochondria (P) | 660 aa, 14 PPR motifs |
| EMB1270 (At3g18110) | Chloroplast | 1429 aa, 24 PPR motifs |
| EMB1796 (At3g49240) | Mitochondria | 629 aa, 11 PPR motifs |
| EMB2261 (At3g49170) | Chloroplast (T) or mitochondria (P) | 849 aa, 17 PPR motifs (four degenerate) |
| EMB2453 (AT4g39620) | Chloroplast | 563 aa, 9 PPR motifs |

^aLocalization predicted with predotar (P) and targetP (T) programs: [<http://genoplante-info.infobiogen.fr/predotar/predotar.html>] and [<http://www.cbs.dtu.dk/services/TargetP/>] (Small et al. 2004; Emanuelsson et al. 2000). Predictions of each program are indicated by parentheses for proteins that yielded ambiguous results

^bNumber of amino acids in each protein is indicated (aa), followed by the number of PPR unit motifs. PPR motifs were identified using PFAM [<http://www.sanger.ac.uk/Software/Pfam/>] (Bateman et al. 2004)] and InterProScan [<http://www.ebi.ac.uk/InterProScan/>] (Mulder et al. 2003)]

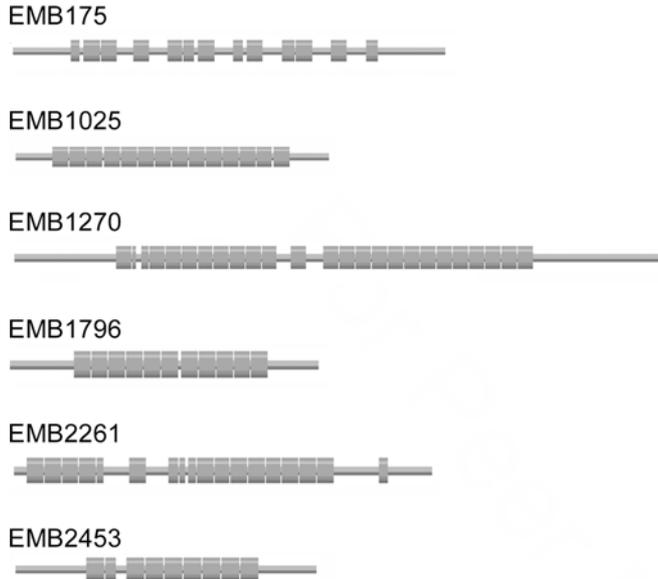


Fig. 7 Diversity of protein sizes and PPR motif distribution in emb-PPR proteins subjected to knockout phenotype analysis. Schematic protein diagrams are to scale, with PPR motifs designated by boxes. Protein lengths and motif numbers are summarized in Table 4. PPR motifs were identified using PFAM [<http://www.sanger.ac.uk/Software/Pfam/> (Bateman et al. 2004)] and InterProScan [<http://www.ebi.ac.uk/InterProScan/> (Mulder et al. 2003)].

appeared to arrest earliest, as a squat, triangular embryo with rigid asymmetric cotyledons, no defined hypocotyl, and suspensors that persisted even in mutants from maturation phase siliques. Like *emb2453*, *emb2261* consistently failed to elongate, developing instead as v-shaped embryos with wide, stunted cotyledons and no hypocotyl. In contrast, *emb1270* elongated and developed a clear hypocotyl, but featured dramatically stunted cotyledons. Interestingly, *emb2261* and *emb1270* both often exhibited a domed shoot apex along with severely stunted cotyledons, suggesting that apical cell division continued but cells were not allocated properly toward organogenesis. *Emb1025* was unique, developing slowly to a morphology resembling younger, narrow, somewhat pointed torpedo stage embryos with slightly stunted cotyledons (Fig. 8e). Together, the differential impacts on cotyledon, shoot apex, and hypocotyl development among these *emb-ppr* mutants indicated that disruption of different *PPRPs* can have surprisingly varied consequences for embryo morphogenesis.

Discussion

PPRPs comprise one of the largest plant gene groups, representing almost 2% of the predicted genes in *Arabidopsis* (Small and Peeters 2000). *PPRPs* are not so numerous in fungal and animal genomes, and thus have undergone a large evolutionary expansion in plants (Lurin et al. 2004; T. Anderson, D. Vernon, unpublished). Here, taking a developmental- and reverse-ge-

netic approach, we have demonstrated that despite the large, plant-specific expansion of this superfamily, a number of *PPRPs* are of fundamental importance, being indispensable in the earliest phases of plant development. In addition, the diverse phenotypes of *pprp* knockout mutants reveal that loss of *PPRP* gene functions can impact morphology in different ways.

emb175 reveals a broader developmental context for plant *PPRP* activity

Emb175 provides the first unambiguous example of a developmental mutant with a *PPRP* gene defect. Previous genetic investigations of plant *PPRP* function have focused on a small number of mutants identified by abnormal seedling pigmentation or high chlorophyll fluorescence (Barkan et al. 1994; Fisk et al. 1999; Hashimoto et al. 2003; Meierhoff et al. 2003; Williams and Barkan 2003; Yamazaki et al. 2004), or on mutations that restore fertility in male-sterile lines of various species (Bentolila et al. 2002; Kazama and Toriyama 2003; Komori et al. 2004). The characterization of *emb175* and identification of the *EMB175* gene now establish a requirement for a *PPRP* in early, pre-photosynthetic development: defects as early as the 32–64 cell globular stage in *emb175* embryos indicate that the gene is active and essential well before the *Arabidopsis* embryo achieves photosynthetic competence at the torpedo stage.

Morphological arrest in *emb175* mutants is likely due to a housekeeping defect, rather than direct disruption of morphogenetic regulation. Arrest appears to be a net consequence of cell-autonomous defects in cell division rate and orientation, rather than a breakdown in overall morphogenetic regulation (Fig. 2). Like many *PPRPs*, *EMB175* is predicted to be a chloroplast protein. Acting in the chloroplast, *EMB175* could potentially affect organelle biogenesis or any number of essential biosynthetic or metabolic processes based there, disruption of which could lead to the observed early developmental arrest. Consistent with this, *emb175*'s phenotype resembles those of other *Arabidopsis emb* mutants defective in plastid proteins, such as *rsy3* and *emb506*, which also arrest at the globular stage (Albert et al. 1999; Despres et al. 2001; Apuya et al. 2002). Thus, *emb175* provides support for the view that plastids are essential for organismal viability well before plants gain photosynthetic competence. Full elucidation of the biochemical function of *EMB175* will require protein localization and identification and analysis of non-lethal mutant alleles.

Using sequence-indexed T-DNA knockout collections to identify additional *emb-PPRPs*

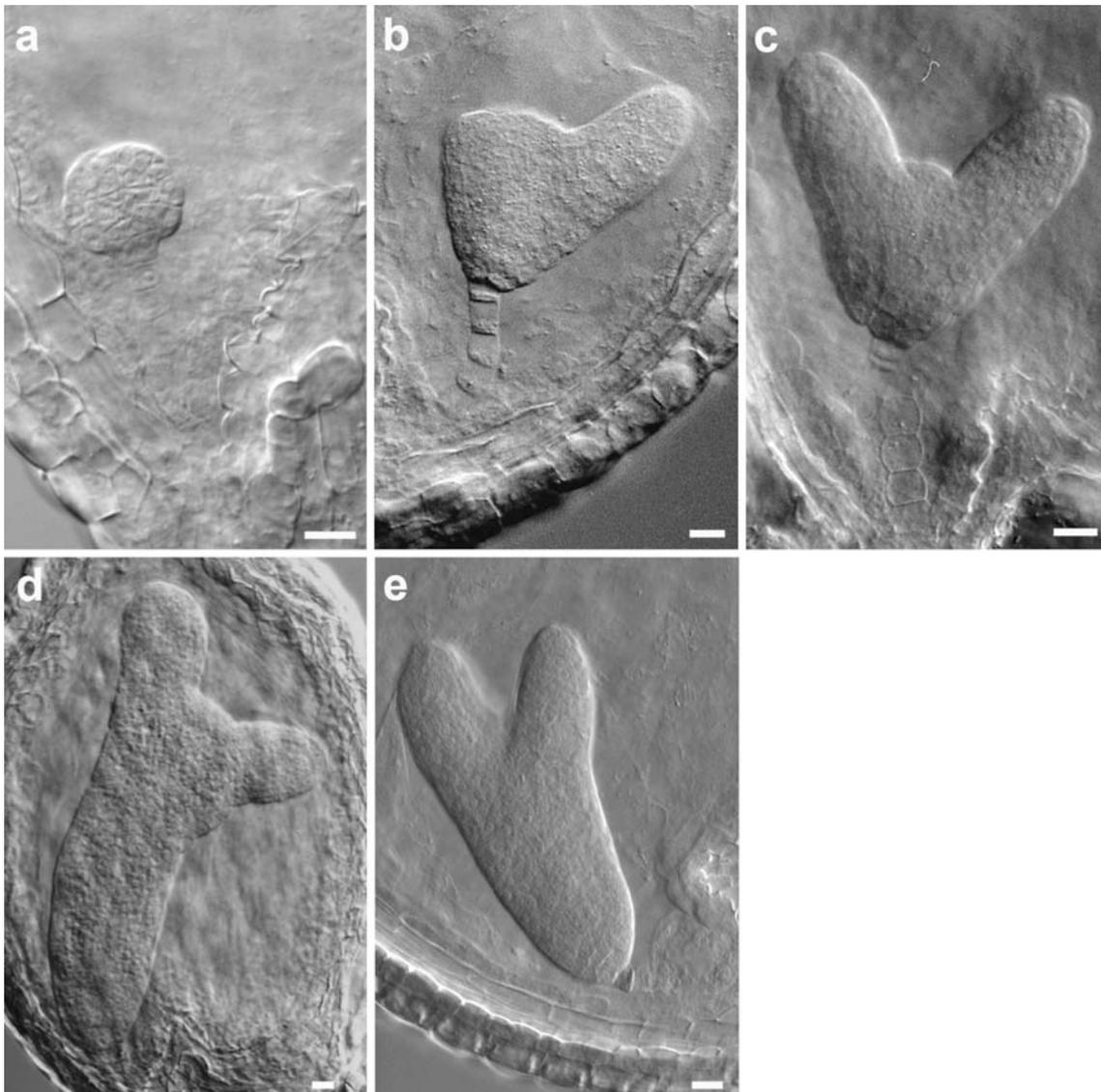
Using available genomic resources, we confirmed essential embryonic functions for six additional *Arabidopsis* *PPRPs*, in addition to *EMB175*. Publicly-available, sequence-searchable collections of T-DNA

knockout mutants allowed rapid identification of additional prospective PPRPs with developmental roles. However, our molecular and genetic analyses of these lines illustrates that these resources should be viewed chiefly as starting points for more in-depth molecular and genetic studies: of 11 prospective *emb-PPRPs* we identified in the SeedGenes and Salk collections, embryonic phenotypes could be confirmed for only six, due to mis-annotation, actual location of T-DNA inserts in neighboring genes (rather than in *PPRPs*), lack of phenotypes in prospective allelic lines, and/or complementation between presumptive allelic *pprp* knockout lines. Due to the broad scope of gene-tagging projects such as the SeedGenes database, such discrepancies were not unanticipated and the likelihood of such mistakes has been acknowledged (Tzafrir et al. 2004). Even so, our results suggest a high error rate in such databases and underscore the need for caution in interpreting data from large-scale gene tagging projects.

pprp knockout mutations affect diverse aspects of embryo morphology

The most notable feature of the *emb-pprp* knockout mutants was their phenotypic diversity—most notably with respect to the dramatic morphological defects in some lines. Other than *emb175*, only one *pprp* mutant—*emb1796*—arrested early in embryogenesis, and it did so consistently earlier than *emb175*. The other *emb-pprps* all arrested later in development, and each of those exhibited unique and dramatic morphological defects. This phenotype diversity indicates that the *emb-pprp* mutations must, somehow, ultimately be impacting distinct morphogenetic programs.

How can mutation of different *PPRP* genes affect embryogenesis in different ways? The simplest explanation, consistent with plant *PPRP* functions that have been documented to date, is that *emb-PPRPs* may affect morphogenesis indirectly: like previously characterized plant *PPRPs*, they may function strictly in organelle



gene expression, perhaps affecting the expression of one or a few chloroplast or mitochondrial genes. In this model, each knockout mutant would suffer from impairment of different organelle functions, resulting in a variety of effects on embryo morphology. Alternatively, different *pprp* knockout mutations may impair organelle function to different extents, some severely disrupting organelle gene expression, others leading only to partial reductions in organelle gene function. In either case, this model implies that different defects in organelle gene expression can have very different morphological consequences in the developing plant embryo.

Within the scenario described above, one hypothetical explanation for phenotype diversity is that these *emb-pprp* mutations could differ in severity, with some lines harboring true knockout alleles, while others may suffer from a partial loss of gene function that leads to aberrant late-stage morphology rather than globular arrest. Studies of other embryo-defective mutants have suggested that weak alleles of essential genes can result in abnormal late-stage embryo morphology (Vernon and Meinke 1995). However, the mutants characterized in this study likely harbor full knockout alleles, as judged by the positions of T-DNA inserts within 5' or PPR-encoding exons of *PPRP* transcribed regions in all but the *emb1796* mutants (which nevertheless exhibited a severe early-lethal phenotype). Mutants such as *emb1444*, for which knockout status seemed dubious, were not subjected to further phenotypic characterization in this study (see Table 3). Thus, the range of *pprp* phenotypes reported here likely results from null *pprp* alleles. Another possible explanation for the unusual late-stage phenotypes seen in some lines could be partial functional redundancy between these *emb-PPRPs* and related genes. Given the presumed specificity of PPR protein functions, and the embryo-lethality of all of these mutations, this seems unlikely. However, it remains possible that partial redundancy could result in the less-severe, late-embryo phenotypes observed for some *emb-pprp* mutants. This possibility could be experimentally addressed through identification of related *PPRP* genes and isolation of corresponding knockout mutants, followed by double mutant studies. Whatever the case, the knockout phenotypes reported here demonstrate that loss of individual *PPRP* functions can have dramatic and distinct effects on embryogenesis.

An alternative, more speculative, model for the phenotypic diversity reported here is that some *emb-pprp* mutations impact morphology directly, perhaps by affecting expression of developmentally significant RNAs, or by disrupting yet-undefined developmental mechanisms involving *PPRPs*. Diverse cellular functions have been identified for animal *PPRPs*, such as *Drosophila* Bicoid Stabilization Factor (Mancebo et al. 2001) and mammalian LRP130 (Liu and McKeegan 2002), as well as for at least one plant *PPRP*, DNA-binding p63 from wheat (Ikeda and Gray 1999). Therefore, it seems possible that some of the diverse morphological phenotypes reported here could result

more directly from disruption of novel *PPRP* functions. It is conceivable, for example, that some *PPRPs* could influence organelle–nucleus communications, or other processes that could have major impacts on embryo morphology and other developmental programs. More detailed molecular analyses, and characterization of individual *pprp* mutants with non-lethal alleles, will be needed to fully elucidate *PPRP* functions and their mechanistic relationship to embryo morphology.

The identification of *PPRPs* essential for embryogenesis broadens the developmental context in which *PPRP* functions must be considered and provides a foundation for more in-depth functional studies of individual *emb-PPRP* genes. As members of a large superfamily of proteins that has undergone dramatic, plant-specific expansion, essential *PPRPs* may provide insights into aspects of multicellular development specific to higher plants.

Acknowledgements We are grateful for support from the USDA (02-35304-12304 to DMV) and the Murdock Charitable Trust. *Emb175* characterization was initiated with support from NSF-9604344 to DMV. The W.M. Keck Foundation provided support for digital imaging equipment. We thank Dr. David Meinke (Oklahoma State University) for *emb175* complementation crosses with *emb1899-1* and *-2* when those mutants were identified in the course of this work. cDNAs and mutant seed stocks were provided by the ABRC. We thank Tovi M. Anderson and John E. Houghland for assistance.

References

- Albert S, Despres B, Guillemot J, Bechtold N, Pelletier G, Delseny M, Devic M (1999) The EMB506 gene encodes a novel ankyrin repeat containing protein that is essential for the normal development of *Arabidopsis* embryos. *Plant J* 17:169–179
- Alonso JM et al (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301:653–657
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Apuya NR, Yadegari R, Fischer RL, Harada JJ, Goldberg RB (2002) Raspberry3 gene encodes a novel protein important for embryo development. *Plant Physiol* 129:1–15
- Aubourg S, Boudet N, Kreis M, Lecharny A (2000) In *Arabidopsis thaliana*, 1% of the genome codes for a novel protein family unique to plants. *Plant Mol Biol* 42:603–613
- Barkan A, Walker M, Nolasco M, Johnson D (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. *EMBO J* 13:3170–3181
- Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR (2004) The Pfam protein families database. *Nucleic Acids Res* 32:D138–D141
- Bentolila S, Alfonso AA, Hanson MR (2002) A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. *Proc Natl Acad Sci USA* 99:10887–10892
- Castle LA, Errampalli D, Atherton T, Franzmann, Yoon E, Meinke DW (1993) Genetic and molecular characterization of embryonic mutants identified following seed transformation in *Arabidopsis*. *Mol Gen Genet* 241:504–514
- Coffin JW, Dhillon R, Ritzel RG, Nargang FE (1997) The *Neurospora crassa* *cya-5* nuclear gene encodes a protein with a

- region of homology to the *Saccharomyces cerevisiae* PET309 protein and is required in a post-transcriptional step for the expression of the mitochondrially encoded COXI protein. *Curr Genet* 32:273–280
- Despres B, Delseny M, Devic M (2001) Partial complementation of embryo defective mutations: a general strategy to elucidate gene function. *Plant J* 27:149–159
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300:1005–1016
- Errampalli D, Patton D, Castle L, Mickelson L, Hansen K, Schnell J, Feldmann K, Meinke DW (1991) Embryo lethals and T-DNA insertion mutagenesis in *Arabidopsis*. *Plant Cell* 3:149–157
- Fisk DG, Walker MB, Barkan A (1999) Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO J* 18:2621–2630
- Forsthoefel NR, Wu Y, Shulz B, Bennett, MJ, Feldmann KA (1992) T-DNA insertion mutagenesis in *Arabidopsis*: prospects and perspectives. *Aust J Plant Physiol* 19:353–366
- Franzmann L, Yoon E, Meinke DW (1995) Saturating the genetic map of *Arabidopsis thaliana* with embryonic mutations. *Plant J* 7:341–350
- Goldberg RB, de Paiva G, Yadegari R (1994) Plant embryogenesis: zygote to seed. *Science* 266:605–614
- Hashimoto M, Endo T, Peltier G, Tasaka M, Shikanai T (2003) A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in *Arabidopsis*. *Plant J* 36:541–549
- Ikeda TM, Gray MW (1999) Characterization of a DNA-binding protein implicated in transcription in wheat mitochondria. *Mol Cell Biol* 19:8113–8122
- Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796–826
- Kazama T, Toriyama K (2003) A pentatricopeptide repeat-containing gene that promotes the processing of aberrant *atp6* RNA of cytoplasmic male-sterile rice. *FEBS Lett* 544:99–102
- Koc EC, Spremulli LL (2003) RNA-binding proteins of mammalian mitochondria. *Mitochondrion* 2:277–291
- Komori T, Ohta S, Murai N, Takakura Y, Kuraya Y, Suzuki S, Hiei Y, Imaseki H, Nitta N (2004) Map-based cloning of a fertility restorer gene, *Rf-1*, in rice. *Plant J* 37:315–325
- Liu L, McKeegan WL (2002) Sequence analysis of LRPPRC and its SEC1 domain interaction partners suggests roles in cytoskeletal organization, vesicular trafficking, nucleocytosolic shuttling and chromosome activity. *Genomics* 79:124–136
- Lurin C et al (2004) Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16:2089–2103
- Mancebo R, Zhou X, Shillinglaw W, Henzel W, Macdonald PM (2001) BSF binds specifically to the bicoid mRNA 3' untranslated region and contributes to stabilization of bicoid mRNA. *Mol Cell Biol* 21:3462–3471
- Manthey GM, McEwen JE (1995) The product of the nuclear gene PET309 is required for translation of mature mRNA and stability or production of intron-containing RNAs derived from the mitochondrial COXI locus of *Saccharomyces cerevisiae*. *EMBO J* 14:4031–4043
- McElver J, Tzafrir I, Aux G, Rogers R, Ashby C, Smith K, Thomas C, Schetter A, Zhou Q, Cushman MA, Tossberg J, Nickle T, Leven JZ, Law M, Meinke D, Patton D (2001) Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics* 159:1751–1763
- Meierhoff K, Felder S, Nakamura T, Bechtold N, Schuster G (2003) HCF152, and *Arabidopsis* RNA binding pentatricopeptide repeat protein involved in the processing of Chloroplast *psbB-psbT-H-petB-D* RNAs. *Plant Cell* 14:1480–1495
- Meinke DW (1995) Molecular genetics of plant embryogenesis. *Annu Rev Plant Physiol Plant Mol Biol* 46:369–394
- Mulder NJ et al (2003) The InterPro Database, 2003 brings increased coverage and new features. *Nucleic Acids Res* 31:315–318
- Sambrook J, Fritsch EF, Maniatis, T (1989) *Molecular Cloning: a laboratory manual*. Cold Spring Harbor, Cold Spring Harbor
- Schwartz BW, Yeung EC, Meinke DW (1994) Disruption of morphogenesis and transformation of the suspensor in abnormal suspensor mutants of *Arabidopsis*. *Development* 120:3235–3245
- Schwartz BW, Vernon DM, Meinke DW (1997) Development of the suspensor: differentiation, communication, and programmed cell death during plant embryogenesis. In: Larkins BA, Vasil IK (eds) *Cellular and molecular biology of plant seed development*. Kluwer, Dordrecht, pp 73–115
- Small ID, Peeters N (2000) The PPR motif: a TRP-related motif prevalent in plant organellar proteins. *Trends Biochem Sci* 2:46–47
- Small ID, Peeters N, Legeai F, Lurin C (2004) Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4:1581–1590
- Tax FE, Vernon DM (2001) T-DNA associated duplication/translocations in *Arabidopsis*: implications for mutant analysis and functional genomics. *Plant Physiol* 126:1527–1538
- Tsuchiya N, Fukuda H, Sugimura T, Nagao M, Nakagama H (2002) LRP130, a protein containing nine pentatricopeptide repeat motifs interacts with a single-stranded cytosine-rich sequence of mouse minisatellite Pc-1. *Eur J Biochem* 269:2927–2933
- Tzafrir I, Pena-Muralla R, Dicherman 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
- Vernon DM, Meinke DW (1994) Embryogenic transformation of the suspensor in *twin*, a polyembryonic mutant of *Arabidopsis thaliana*. *Dev Biol* 165:566–573
- Vernon DM, Meinke DW (1995) Late *embryo-defective* mutants of *Arabidopsis*. *Dev Genet* 16:311–320
- West MA, Harada JJ (1993) Embryogenesis in higher plants: an overview. *Plant Cell* 5:1361–1369
- Williams PM, Barkan A (2003) A chloroplast-localized PPR protein required for plastid ribosome accumulation. *Plant J* 36:675–686
- Yadegari R, De Paiva G, Laux T, Kalunow AM, Apuya N, Zimmerman JL, Fischer B, Harada JJ, Goldberg RB (1994) Cell differentiation and morphogenesis are uncoupled in *Arabidopsis* raspberry embryos. *Plant Cell* 6:1713–1729
- Yamazaki H, Tasaka M, Shikanai T (2004) PPR motifs of the nucleus-encoded factor PGR3 function in selective and distinct steps of chloroplast gene expression in *Arabidopsis*. *Plant J* 38:152–163