

# Mitochondrial Pentatricopeptide Repeat Protein, EMB2794, Plays a Pivotal Role in NADH Dehydrogenase Subunit *nad2* mRNA Maturation in *Arabidopsis thaliana*

Fernanda Marchetti<sup>1</sup>, Maximiliano Cainzos<sup>1</sup>, Sofia Shevtsov<sup>2</sup>, Juan Pablo Córdoba<sup>1</sup>, Laure Dora Sultan<sup>2</sup>, Axel Brennicke<sup>3,†</sup>, Mizuki Takenaka<sup>4</sup>, Gabriela Pagnussat<sup>1</sup>, Oren Osterseztzer-Biran<sup>2</sup> and Eduardo Zabaleta<sup>1,\*</sup>

<sup>1</sup>Instituto de Investigaciones Biológicas (IIB)-Universidad Nacional de Mar del Plata (UNMDP)-CONICET, Funes 3250 3er nivel, 7600 Mar del Plata, Argentina

<sup>2</sup>Department of Plant and Environmental Sciences, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, 919040 Jerusalem, Israel

<sup>3</sup>Institut für Molekulare Botanik, Universität Ulm, Ulm 89069, Germany

<sup>4</sup>Department of Botany, Graduate School of Science, Kyoto University, Oiwake-cho, Sakyo-ku, Kyoto, 606-8502 Japan

\*Corresponding author: E-mail, ezabalet@mdp.edu.ar; Fax, +54 223 475 30 30.

†Passed away on February 2017.

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession number: Arabidopsis EMB2794 (At2g02150).

(Received December 27, 2019; Accepted March 8, 2020)

The Arabidopsis genome encodes >450 proteins containing the pentatricopeptide repeat (PPR) motif. The PPR proteins are classified into two groups, termed as P and P Long-Short (PLS) classes. Typically, the PLS subclass proteins are mainly involved in the RNA editing of mitochondrial and chloroplast transcripts, whereas most of the analyzed P subclass proteins have been mainly implicated in RNA metabolism, such as 5' or 3' transcript stabilization and processing, splicing and translation. Mutations of PPR genes often result in embryogenesis and altered seedling developmental defect phenotypes, but only a limited number of *ppr* mutants have been characterized in detail. In this report, we show that null mutations in the EMB2794 gene result in embryo arrest, due to altered splicing of *nad2* transcripts in the Arabidopsis mitochondria. In angiosperms, *nad2* has five exons that are transcribed individually from two mitochondrial DNA regions. Biochemical and in vivo analyses further indicate that recombinant or transgenic EMB2794 proteins bind to the *nad2* pre-mRNAs in vitro as well as in vivo, suggesting a role for this protein in trans-splicing of *nad2* intron 2 and possibly in the stability of the second pre-mRNA of *nad2*. Homozygous *emb2794* lines, showing embryo-defective phenotypes, can be partially rescued by the addition of sucrose to the growth medium. Mitochondria of rescued homozygous mutant plants contain only traces of respiratory complex I, which lack the NADH-dehydrogenase activity.

**Keywords:** Complex I • *nad2* transcript • Pentatricopeptide repeat protein • Plant mitochondria.

## Introduction

The pentatricopeptide repeat (PPR) protein family is defined by the presence of a degenerative 35-amino-acid motif repeated in

tandem up to 30 times (Small and Peeters 2000, Cheng et al. 2016). In plants, the gene family coding for this kind of proteins has greatly expanded (Schmitz-Linneweber and Small 2008, Cheng et al. 2016). In Arabidopsis, this family is composed of 458 members divided into two subfamilies based on the structure of the repeated motif, named P and PLS subfamilies. While members of the P-type contain the canonical P motif, which is common to all eukaryotes, members of the PLS subfamily contain two additional derived PPR motifs, including a short (S) and a long (L) motifs, further divided into the PLS, E, E+ and DYW subtypes (Lurin et al. 2004). PPR proteins are predominantly located within organelles; about 54% of the proteins reside in mitochondria and 19% in plastids (Lurin et al. 2004). So far, only a single nuclear PPR protein has been shown to function in the nucleus in Arabidopsis plants (i.e. GRP23 encoded by the At1g10270 gene) where it seems to act as a regulator of gene expression (Ding et al. 2006). Another member of the family is dually targeted to both the mitochondrion and the nucleus (i.e. PNM1 encoded by the At5g60960 gene). Yet, genetic complementation assays indicated a function for PNM1 only in the mitochondria, where the protein seems to play a critical role in organellar translation (Hammani et al. 2011). Given the mechanism of recognition and slight variation in sequence repeats, all PPR proteins are proposed to function as gene-specific regulators of RNA metabolism (Lurin et al. 2004). In plant organelles, PPR proteins were shown to play important roles in RNA stability, cleavage, splicing, editing and translational regulation (Schmitz-Linneweber and Small 2008, Hammani et al. 2011). The PLS-type PPR proteins are almost exclusively implicated in C-to-U RNA editing in plant organelles (Takenaka et al. 2013, Sun et al. 2016), while the P-class proteins are involved in a wide range of posttranscriptional processes within the organelles, ranging from RNA maturation to translation

(Barkan and Small 2014). Several PPR proteins have been reported to be essential for plant growth and development, as indicated by embryo lethality or severe growth and developmental defects associated with loss-of-function mutants in different PPRs (Falcon de Longevialle et al. 2007, Liu et al. 2010, Sung et al. 2010, Lee et al. 2017). In addition, only few P-class PPR proteins contain a C-terminal small MutS-related (SMR) domain. This subclass plays essential roles in embryo development, chloroplast biogenesis and gene expression and plastid-to-nucleus retrograde signaling (summarized in Zhang and Lu 2019). The SMR domain also confers RNA endonucleolytic activity in plants (Zhou et al. 2017), which was shown essential for chloroplastic 23S and 4.5S rRNA maturation.

Mitochondrial complex I (CI, NADH:ubiquinone oxidoreductase, EC 1.6.5.3) is a major entry site for electrons into the respiratory chain, by mediating the transfer of electrons from NADH to ubiquinone. In *Arabidopsis thaliana*, CI is comprised of 49 subunits (Meyer 2012, Braun et al. 2014, Senkler et al. 2017). The majority of the subunits are nuclear-encoded, and only nine genes have been retained in the mitochondrial DNA in *Arabidopsis*, namely, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7* and *nad9* (Unsel et al. 1997). Many *Arabidopsis* mutants affecting the expression of different CI subunits have been previously characterized (e.g. Lee et al. 2002, Perales et al. 2005, Nakagawa and Sakurai 2006, Meyer et al. 2011, Wydro et al. 2013, Colas Des Francs-Small and Small 2014), but only three mutants in mitochondrial CI loci have been reported so far. These include two *Nicotiana sylvestris* mutants with large deletions in the *nad7* locus (Pla et al. 1995) and a deletion in the 3' end of *nad4* in *Zea mays* (Karpova and Newton 1999). A nuclear mutant of *N. sylvestris* exhibits extremely low CI levels due to a splicing defect in *nad4* intron 1 (Brangeon et al. 2000), while several nuclear mutants affected in CI have been identified in *Arabidopsis* (Lee et al. 2002, Perales et al. 2005, Nakagawa and Sakurai 2006, Meyer et al. 2011, Wydro et al. 2013). Some mutants lacking CI activity (Kühn et al. 2015, Lee et al. 2017) or completely lacking CI (Córdoba et al. 2016, Fromm et al. 2016, Córdoba et al. 2019) have also been recently reported. The phenotypes of CI mutants vary from weak growth phenotypes (*ndufs8*, Pétriacq et al. 2017) to reduced male fertility (Pla et al. 1995, Karpova and Newton 1999), strong developmental defects or even embryo lethality, i.e. *ndufv1* (Kühn et al. 2015), and double mutants of gamma CAs, *ca1ca2*, *ca1ca3* and *cal1cal2* (Córdoba et al. 2016, Fromm et al. 2016, Córdoba et al. 2019).

In the mitochondria of different angiosperms, the coding sequences of *nad1*, *nad2*, *nad4*, *nad5*, *nad7*, *ccmF<sub>c</sub>*, *cox2*, *rpl2* and *rps3* genes are interrupted by group II intron sequences (see e.g. Bonen 2008, Brown et al. 2014). Most of these introns are found in a *cis* configuration, where the intron and the flanking exons are transcribed in a single RNA molecule. Yet, in different angiosperms, the splicing of *nad1*, *nad2* and *nad5* genes occurs between separate RNA molecules, to generate a single mature transcript (reviewed by Brown et al. 2014, Schmitz-Linneweber et al. 2015, Zmudjak and Osterseizer-Biran 2017). While the splicing of bacterial group II introns requires a single, or a few, splicing cofactor, the maturation of the plant organellar introns requires many protein cofactors (Li-Pook-Than and

Bonen 2006, Bonen 2008, Brown et al. 2014, Schmitz-Linneweber et al. 2015, Sultan et al. 2016, Zmudjak and Osterseizer-Biran 2017). Although multiple splicing cofactors were identified in *Arabidopsis* mitochondria, the molecular basis of group II-type introns splicing in plant mitochondria is still under investigation. PPR proteins most likely determine cleavage sites required for the maturation of group II introns by specific RNA recognition (Schmitz-Linneweber and Small 2008). Then, maturases and RNA helicases may stabilize catalytically active structures required in the splicing reactions (Matsuura et al. 2001, Mohr et al. 2006, Köhler et al. 2010, Schmitz-Linneweber et al. 2015, Zmudjak and Osterseizer-Biran 2017).

Although a large number of PPR-related proteins are encoded in plants, the physiological roles and molecular functions of many of them are less clear. One of these proteins, EMB2794, encodes an essential factor that harbors 17 PPR. Mutations in the *EMB2794* (*At2g02150*) gene lead to strong embryogenesis defects, which are tightly associated with altered *trans*-splicing of *nad2* intron 2 and the stability of the second pre-RNA. Analyses of *emb2794* mitochondria indicate strong defects in CI biogenesis and reduced NADH dehydrogenase activity (i.e. below detectable levels). Furthermore, we analyzed the RNA-binding characteristics of a recombinant EMB2794 protein *in vitro* as well as *in vivo* and discuss its putative molecular functions in mitochondria biogenesis and organellar RNA metabolism.

## Results

### *emb2794* insertional mutant plants show delayed embryo development

Pentatricopeptide-repeat (PPR) proteins play important roles in the gene expression of eukaryotic cells, by affecting cellular and organellar RNA processing and translation (Small and Peeters 2000). Several PPRs were found in an initial survey of embryo-defective (EMB) mutants (Tzafrir et al. 2004) and recently summarized by Meinke (2020). These include 34 PPRs out of 510 EMB genes (6.7%) compared with total PPR containing genes at the genome-wide scale (1.6%). The *EMB2794* gene contains no introns and encodes a PPR protein (761aa) with a predicted mRNA of 2,286 nucleotides in length (Fig. 1A). The terminal phenotype was studied in the context of the Seed Genes Project (<http://www.seedgenes.org/>, Meinke et al. 2008). Using two independent alleles (*emb2794-1* and *emb2794-2*, Fig. 1A), we observed that about 25% of the seeds in both lines show an abnormal shrunken phenotype, easily recognizable by their pale appearance in the middle stages of embryogenesis and by the dark brown coloration of mature seeds (Fig. 1B and Supplementary Fig. S2). We confirmed that the most common terminal phenotype associated with this gene is Early Curled Cotyledon (Fig. 1C). The abnormal seeds were isolated and genotyped and were found to be homozygous mutants for the *emb2794* locus (Fig. 1D). In addition, although homozygous knockout *emb2794* seeds are unable to germinate on Murashige & Skoog (MS) media, in rare cases, a few of the seeds were found to germinate in the presence of

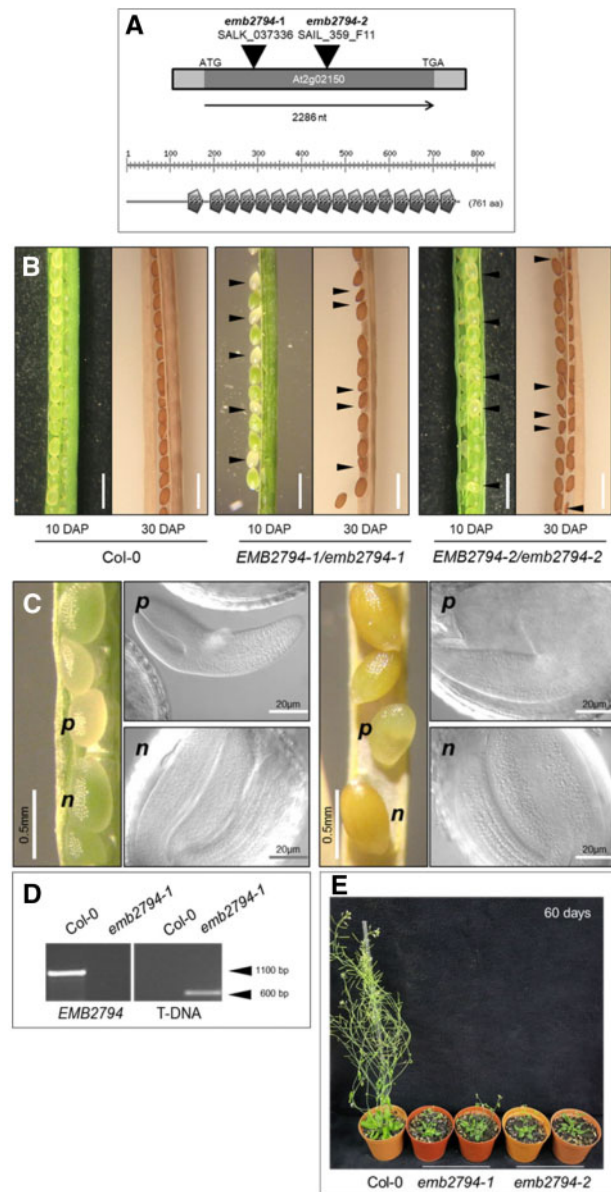
high sucrose concentration in the media and once established were transferred to soil. The rescued homozygous plants show a strong degenerate growth phenotype curled leaves (Fig. 1E and Supplementary Fig. S3). Seeds collected from homozygous plants are all shrunken and could not be rescued, even when grown on MS media supplemented with sucrose. A gene-complementation strategy was conducted introducing a construct containing the full-length coding sequence (2,286 bp) of At2g02150 gene driven by *UBQ10* or its own promoter in the mutant *emb2794* background. Analyses of the T1 plants showed several heterozygous plants for *emb2794* containing the transgene. These plants show siliques containing reduced amounts of shrunken seeds (~12%, Supplementary Fig. S4A and B). These data further provide support that the transgene complements the embryo lethal phenotype associated with *emb2794* mutant plants.

### EMB2794 is a mitochondrial protein in Arabidopsis plants

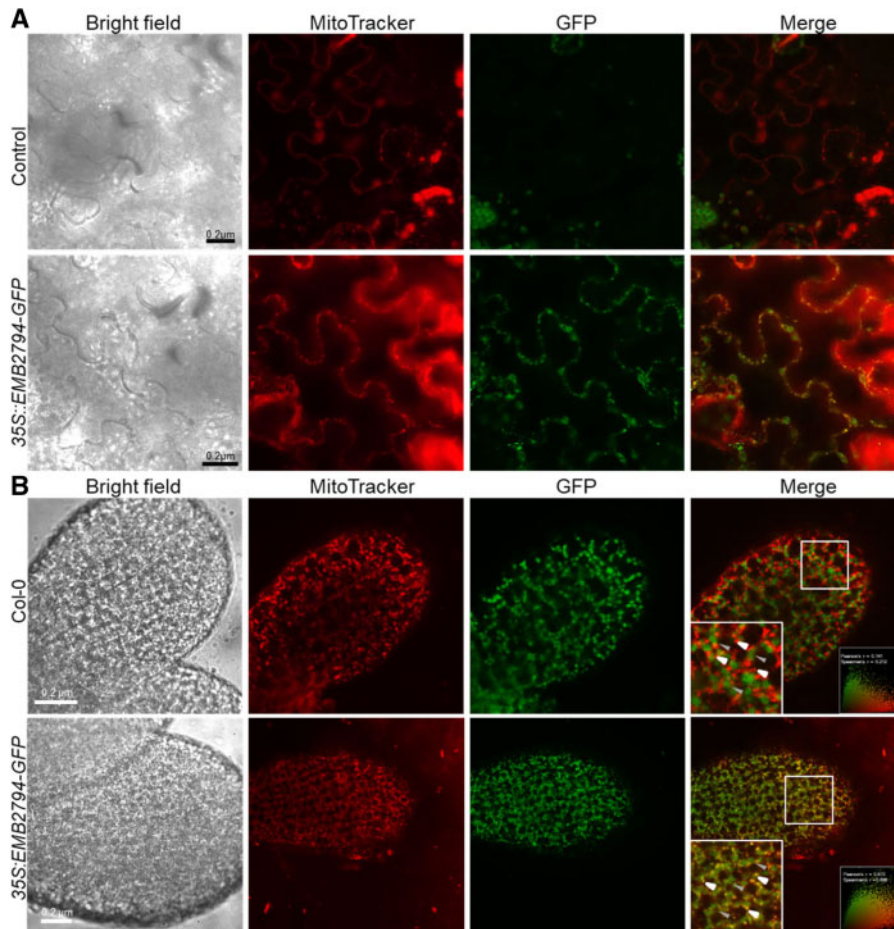
Based on several in silico protein-localization prediction servers (i.e. Predotar, TargetP, iPSORT, MultiLoc, LocTree and AtsubP), EMB2794 was predicted to be located in chloroplast (Colcombet et al. 2013), however, with strong differences among predictors. Predotar, LocTree and AtSubP showed a higher probability of mitochondrial localization, while TargetP and iPSORT provided ambiguous predictions (i.e. possible dual-localization to both organelles). To experimentally determine the subcellular localization of this protein in vivo, we carried out translational fusions of the coding region of EMB2794 with green fluorescent protein (GFP) under the control of the strong promoter 35S and delivered into *Nicotiana benthamiana* epidermal cells via leaf infiltration for transient expression. Forty-eight hours after infiltration, leaves were analyzed under confocal microscope. EMB2794-GFP clearly co-localizes with MitoTracker, staining specifically mitochondria (Fig. 2A). Chloroplasts present in guard cells also show green fluorescence, but this signal is most likely due to auto-fluorescence of chloroplasts, since a similar signal was also present in the negative control (Supplementary Fig. S5). In addition, the construct containing EMB2794-GFP was stably introduced in Arabidopsis plants. After the verification of GFP expression in embryos, experiments with MitoTracker were performed, confirming the co-localization with mitochondria and not with chloroplasts (Fig. 2B). From these experiments, we concluded that EMB2794 is located exclusively in mitochondria.

### Mitochondrial transcriptome analyses reveal that EMB2794 activities are required for the maturation of *nad2* in Arabidopsis mitochondria

It is commonly accepted that PPR proteins are RNA-binding proteins, which are important for posttranscriptional regulation of gene expression in plant organelles, such as *cis*- and *trans*-splicing, C-to-U RNA-editing, trimming and/or RNA stability (Barkan and Small 2014). Therefore, we examined the putative effects of EMB2794 protein on mitochondrial RNA



**Fig. 1** The embryo lethal mutant *emb2794* is rescued with sucrose. (A) Schematic representation of the *EMB2794* gene showing 17 pentatricopeptide repeats. In addition, the site of insertion of T-DNA of two mutants is indicated. (B) *emb2794* siliques show abnormal seeds. In immature siliques (10 days after pollination; DAP), ~25% of developing seeds of heterozygous *emb2794/EMB2794* selfed plants are still pale (arrowheads) in both *emb2794-1* and *emb2794-2* while WT (Col-0) are green. In mature siliques (30 DAP), the ~25% are shrunken (arrowheads). Bars: 1 mm. (C) Embryos within pale seeds are delayed in their development. Terminal phenotype of delayed embryos, n, normal, p, pale seeds. Left panel: 10 d after pollination and right panel: 25 d after pollination. (D) Genotype of pale and green embryos. Genomic PCR detecting the WT *EMB2794* gene or the T-DNA insertion reveals that the pale ones (p) are the homozygous *emb2794* mutants while the normal seeds (n) are heterozygous or WT segregants. (E) Homozygous *emb2794* shrunken seeds were sown on medium supplemented with sucrose as indicated in the Materials and Methods. Rescued plants are smaller than wild-type plants showing curled leaves. Flowering is delayed producing no viable seeds.



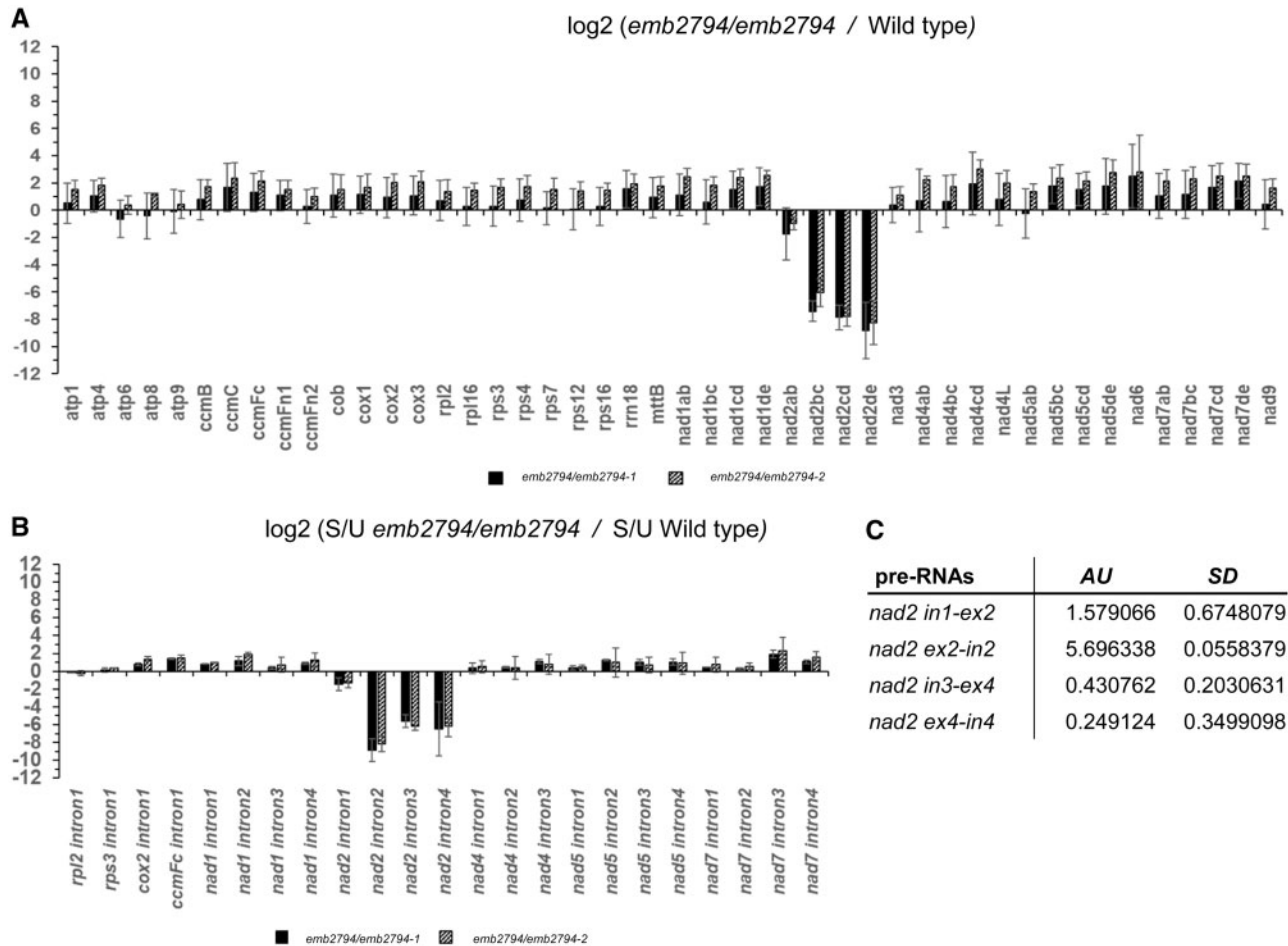
**Fig. 2** EMB2794 is localized in mitochondria. Confocal microscope images showing the subcellular localization of EMB2794-GFP translational fusion in epidermal cells of *Nicotiana benthamiana*-infiltrated leaves (A) and embryos of stable transformed Arabidopsis plants (B). Samples were treated with MitoTracker Red to label mitochondria. The left panel shows the bright field, the second shows mitochondria labeled with the MitoTracker Red dye, the third shows the GFP signal and the right panels present the merged images. (A) Individual epidermal cells are shown. Large vacuoles without fluorescent signals occupy the center of each epidermal cell. (B) Images show one cotyledon of Col-0 or EMB2794-GFP embryos. A close up is shown (box). Chloroplasts are shown (grey arrows). Yellow spots indicate EMB2794-GFP protein co-localized in mitochondria (white arrows). Scatter plot is shown (inset); Pearson's and Spearman's correlation are indicated.

processing and metabolism. As knockout *emb2794* mutants show embryogenesis-defective phenotypes, we first used artificial microRNAs (amiRs) to reduce the *EMB2794* expression. We designed two independent allelic amiRs (amiR-A and amiR-B) targeting the *EMB2794* transcript. These constructs were introduced into a binary plasmid to direct their expression by the constitutive 35S promoter. To minimize the risk of silencing the 35S promoter driving the amiR transgenes, the constructs were introduced into the Arabidopsis *rdl6* mutant background (Carbonell et al. 2014). Total RNA was then extracted from several transgenic lines for each silencing construct, and transcript level of *EMB2794* was assessed by reverse transcriptase (RT)-quantitative PCRs (qPCRs). Two lines showing the lowest expression, named *emb2794-2A4* and *2B17* [29% and 10% of the wild-type (WT) *EMB2794* expression, respectively, **Supplementary Fig. S6**] were selected for further analyses.

Using these amiR lines, we undertook a global transcriptome analysis, considering changes in the steady-state levels of different mitochondrial genes, excluding tRNAs, rRNAs and mtORFs, as well

as alterations in the splicing efficiencies of the 23 group II intron sequences in Arabidopsis mitochondria (Unseld et al. 1997). In both amiR lines, we could not detect any consistent variation in transcript abundance, considering a minimum log<sub>2</sub> ratio cutoff of  $\pm 2$  (**Supplementary Fig. S7**). For this reason, we performed the same global analysis using RNA extracted from rescued homozygous *emb2794* plants, as a partial reduction in *EMB2794* levels in the amiR lines may not produce strong phenotypes. The transcriptome data clearly revealed that, in both *emb2794* mutant lines, the accumulation of *nad2* mRNA transcripts and, particularly, those corresponding to *nad2* exons 2–3 (*ex-bc*) exons 3–4 (*ex-cd*) and exons 4–5 (*ex-de*) is strongly affected (**Fig. 3A**). As these transcripts are all interrupted by group II intron sequences, a large reduction in *nad2* exons 2–3, 3–4 and 4–5 may indicate that their splicing was affected in the mutant lines.

To further establish the roles of *EMB2794* in the maturation of *nad2*, we examined the splicing efficiencies of the 23 group II introns in embryo-rescued plants. The maturation of *nad2* transcripts involves the removal of four different group II-type



**Fig. 3** *nad2* transcripts are strongly affected in homozygous mutant-rescued plants. RNA extracted from leaves of Col-0 plants (control) and *emb2794-1*- and *emb2794-2*-rescued plants was analyzed by quantitative RT-PCR using specific primers. (A) The histogram shows the  $\log_2$ -transformed ratios of mature RNA expression for each mitochondrial transcript between the two mutant alleles and Col-0 plants with three biological repeats. Error bars indicate standard deviation. (B) Splicing efficiency in *emb2794* rescued plants was analyzed using primers that specifically amplified pre-RNAs and mRNAs. The histogram shows the  $\log_2$ -transformed ratios of spliced/unspliced in *emb2794* lines vs. those of Col-0 plants. Values are mean  $\pm$  SD of three biological replicates. (C) Table showing the ratio of unspliced/spliced *nad2* transcripts in *emb2794-2* mutant line. AU, arbitrary units.

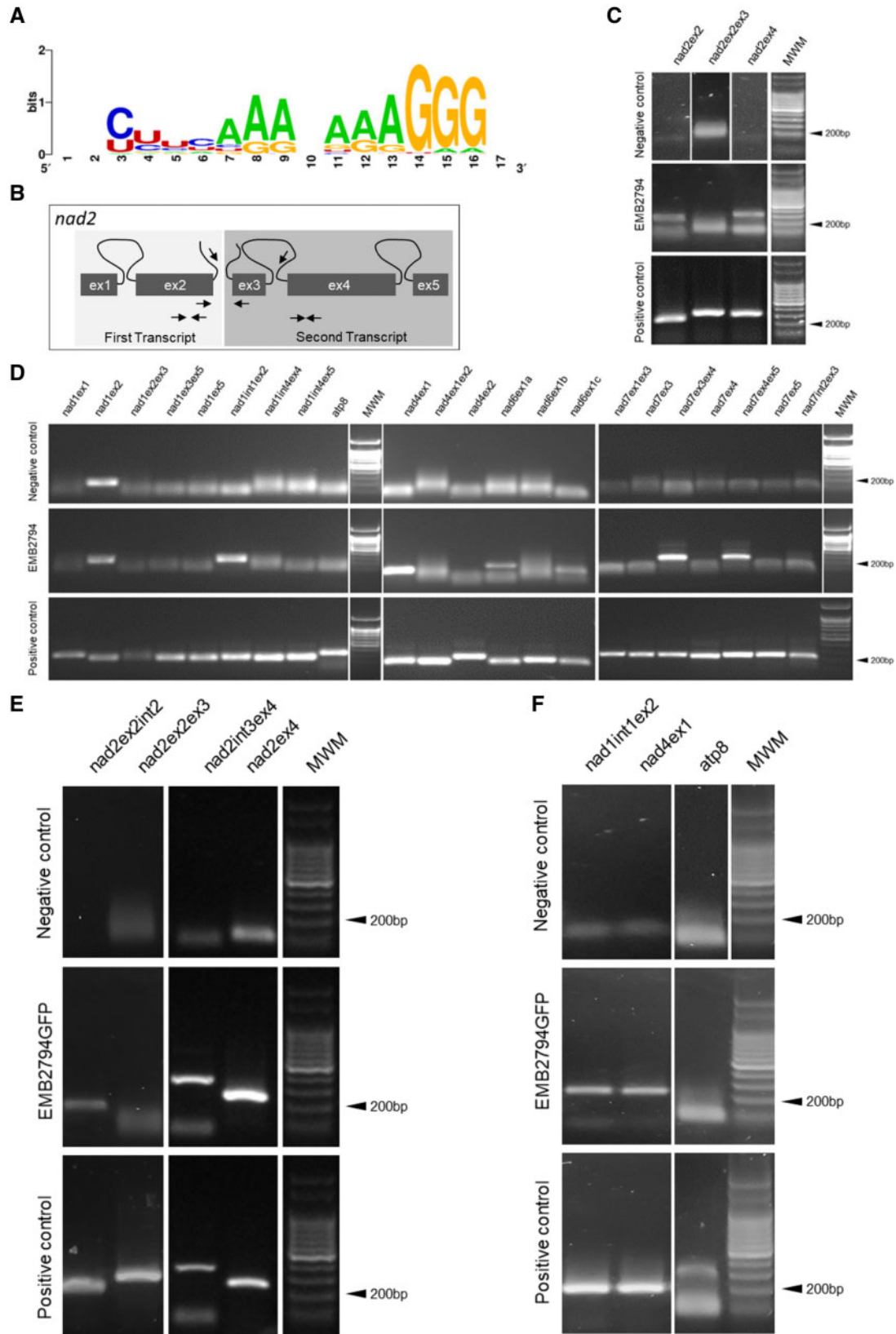
introns, including three *cis*-spliced intron (i.e. introns 1, 3 and 4) and the *trans*-spliced *nad2* intron 2. Analysis of the splicing efficiencies in *emb2794* mutants was estimated by comparing the ratios of pre-RNA to mRNA in the rescued lines vs. Col-0 plants. These analyses further indicated that the efficiency of the *trans*-splicing *nad2* intron 2 is strongly affected in *emb2794* mutants (Fig. 3B) and the corresponding unspliced transcript is accumulated (Fig. 3C). As similar effects were seen in two independent mutant alleles, we concluded that the functions of EMB2794 proteins are required for the splicing of the second (*trans*-spliced) intron in *nad2*. While the first intron is slightly affected, all transcripts corresponding to the second pre-RNA are downregulated in the mutants, including the unspliced versions. We thus suggest that EMB2794 may also be required for the stability of the second *nad2* pre-RNA.

### EMB2794 is associated with *nad2* pre-RNAs in vivo

PPR proteins are canonical RNA-binding proteins (Yin et al. 2013, Coquille et al. 2014), where two to three amino acids in

each PPR domain form hydrogen bonds with a specific ribonucleotide base (Shen et al. 2016). Since our experimental data suggested the reduction in mature *nad2* transcript, we predicted the binding consensus sequence of EMB2794 (Fig. 4A) and searched for possible binding sites within the coding region of *nad2*. We found several possible binding sites in intron 3 as well as in intron 2 near exon 2 (Supplementary Fig. S8).

To confirm whether EMB2794 protein is directly associated with their identified RNA targets, i.e. *nad2* pre-mRNA transcripts, we performed in vitro RNA-binding assays with a recombinant EMB2794 protein. For this purpose, we expressed a recombinant version of EMB2794 in *Escherichia coli* as a fusion protein to Maltose Binding Protein (MBP) with a TEV cleavage site inserted between the two chimeric proteins (MBP-EMB2794) as described in the Materials and Methods. MBP alone attached to the column was used as a negative control. We analyzed the presence of transcripts whose accumulation was either affected or unaffected in the mutant lines, considering the results of the mitochondrial transcriptome obtained



**Fig. 4** EMB2794 is able to bind both *nad2* pre-mRNAs. (A) Consensus of EMB2794 putative binding site shown as WebLogo created in <https://weblogo.berkeley.edu/logo.cgi> using all RNA target predicted sequences for EMB2794 PPR motif. (B) Scheme of *nad2* pre-transcripts (first and second transcripts). Primers used over exons and intron are shown (arrows). (C, D) Recombinant MBP–EMB2794 or MBP (negative control) bound

(continued)

from rescued homozygous *emb2794* plants (Fig. 3). In Fig. 4A, the consensus sequence of EMB2794 putative binding sites is shown. The recombinant EMB2794 protein was found to be stably associated with the two primary transcripts of *nad2* (Fig. 4C). These must undergo a *trans*-splicing activity to allow the joining of the two *nad2* fragments into a single mature *nad2* RNA. However, a transcript containing a predictive binding site within the coding sequence such as *atp8* was not found associated to the recombinant EMB2794 protein (Fig. 4D). Surprisingly, other non-coding regions of several mitochondrial transcripts, including the first transcript of *nad1*, *nad4* ex2-intron 2 and *nad7* ex3-intron 3-ex4, were also predicted as putative RNA targets of EMB2794 protein. For these particular *nad* transcripts (*nad1*, *nad4* and *nad7*), EMB2794 was also found associated in vitro (Fig. 4D), although this is not consistent with the transcript analyses of rescued plants. These results suggest that EMB2794 has the capacity to bind a portion of several *nad* transcripts, although these interactions may be precluded in vivo.

To determine these capabilities in vivo, we took advantage of the 35S::EMB2794-GFP plants. Using these plants, we performed an in vivo binding assay using a commercial GFP agarose matrix as outlined in the Materials and Methods. After incubation and washing steps, agarose beads containing the PPR-GFP fusion protein complexed with their cognate RNAs were used to isolate RNA. cDNA was synthesized and PCR was conducted for relevant positive and negative controls. As shown in Fig. 4E, the fusion protein was found to be stably bonded to the first and second pre-RNA of *nad2* as we detected in the in vitro assay. Again, our results show that *atp8* transcript was not associated with EMB2794 in vivo. However, other *nad* transcripts, such as *nad1* and *nad4*, were also detected but with less reproducibility (Fig. 4F). Taken together, we propose that EMB2794 is associated with both pre-RNAs of *nad2* in vivo.

### The *emb2794* mutant embryos show normal mitochondrial membrane potential

Several mutants displaying embryo lethal phenotypes show a significant reduction in their mitochondrial membrane potentials, in general at heart stage, just before the growth delay is noticeable (Kühn et al. 2015, Córdoba et al. 2016, Fromm et al. 2016). To test whether such phenotypes are also associated with *emb2794* mutants, siliques of hemizygous plants (for both alleles) were manually pollinated. The membrane potential of mutant embryos was analyzed at different times after fertilization using the tetramethylrhodamine, methyl ester (TMRM) fluorescent probe, which fluoresces red when the mitochondrial membrane potential is high. Embryos at the

heart stage or at later embryogenesis stages were examined, as we could easily distinguish Arabidopsis embryos with normal embryogenesis development from those that showed developmental delay. These analyses indicate that the mutants show a typical mitochondrial membrane potential, equivalent to the hemizygous or WT segregants of the same silique or even compared to WT embryos at the same developmental stage (Fig. 5). These data suggest that despite the altered expression of transcripts coding for the NAD2 subunit, the mitochondria of *emb2794* mutants are still able to maintain some oxidative phosphorylation activity.

### Homozygous *emb2794* plants contain only traces of holo-CI

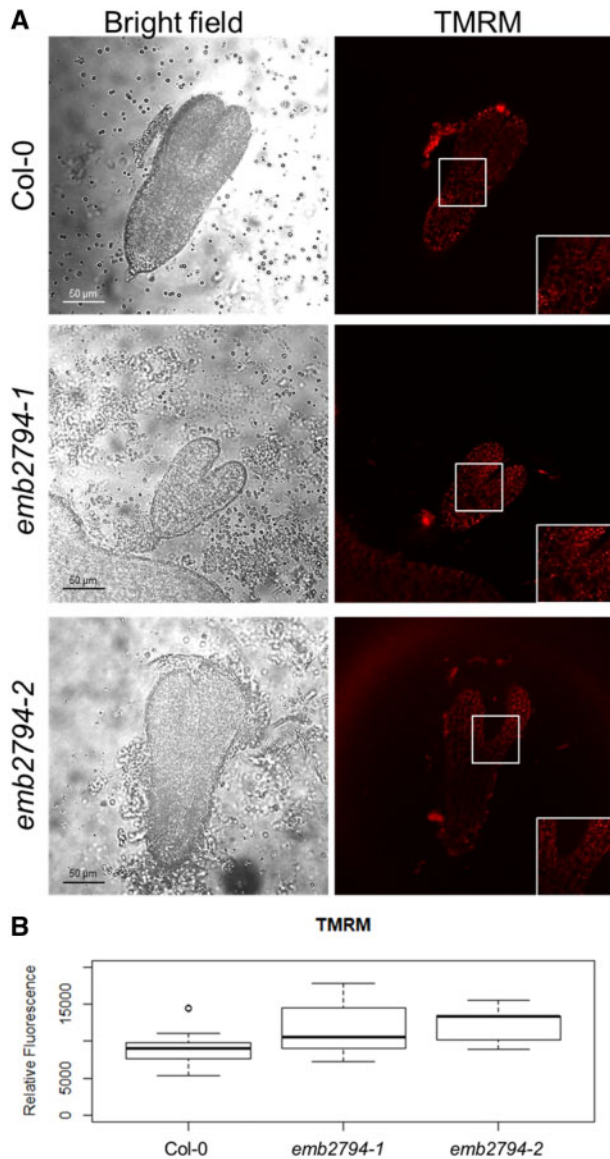
Based on the strong reduction in *nad2* mRNA, we assumed that NAD2 protein is strongly reduced in *emb2794* mutants, which might be affecting the biogenesis of the mitochondrial respiratory CI. To test this assumption, we prepared mitochondria enriched membranous fractions from leaves of WT and mutant Arabidopsis plants. The enriched mitochondria preparations were subjected to 1-D Blue-Native PAGE, in-gel NADH DH activity and immunoblotting assays. An antibody raised against the CI carbonic anhydrase-like subunit 2 (CA2), which is found among the earliest known intermediates of CI (i.e. an 85-kDa sub-particle of CI, Ligas et al. 2019), was used to examine the assembly status of CI in the native mitochondrial preparations. These experiments revealed that fully assembled CI exists only in trace amounts in *emb2794* mutants, although completely inactive. Lower-molecular mass intermediates were found accumulated to higher levels in the *emb2794* mutants (Fig. 6). These include a CI intermediate that contains CA2, with a calculated mass between 85 and 200 kDa (Supplementary Fig. S9), which likely corresponds to the second intermediate proposed by Ligas et al. (2019) that does not include the NAD2 subunit. As CI is inactive, the level of alternative oxidase (AOX) proteins is increased (Supplementary Fig. S10). These data further support the critical roles of NAD2 in CI biogenesis.

### Discussion

In this work, we present evidence showing that EMB2794 is required for the maturation of *nad2* pre-RNAs. In vitro binding assays suggest that EMB2794 directly associates with *nad2* transcripts. In the absence of a functional EMB2794 protein, strong reductions in the steady-state levels of fully mature *nad2* transcripts are evident. Under these circumstances, it is expected that the levels of NAD2 protein, an essential component of CI, would be strongly reduced (Liu et al. 2010, Kühn et al. 2011,

#### Fig. 4 Continued

to a maltose resin was incubated with mRNA extracted from Col-0 plants. The EMB2794-mRNA complexes were eluted by cutting the MBP moiety using TEV protease. The RNAs were purified from the complexes using TRIzol and used for cDNA synthesis and PCR using specific primers over different portions of *nad2* transcript (C) and other mitochondrial transcripts (D). Transgenic plants expressing 35S::EMB2794-GFP fusion proteins were used to perform an RNA-binding assay using GFP-trap agarose beads. After incubation and washing steps, total RNA was extracted from eluted EMB2794-RNA complexes and RT-PCR was performed using the indicated primers over *nad2* (E) and other mitochondrial transcripts (F). In all cases, positive control refers to amplification using mRNA from Col-0 plants as a template.



**Fig. 5** *emb2794* mutant embryos show nearly normal mitochondrial membrane potential. (A) Col-0 and *emb2794-1* and *emb2794-2* embryos were incubated with TMRM at the end of embryo development in the mutants and at an equivalent stage in the WT (torpedo). Embryos were then analyzed under confocal microscopy. Within mutant *EMB2794/emb2794* siliques, there are around 25% of delayed embryos after heart stage, which, however, show near-normal mitochondrial membrane potential. Pictures in the left show bright-field images, in the right, the red fluorescence and the inset. Bars: 50  $\mu$ m. (B) Plots showing red fluorescence of Col-0 and both *emb2794* alleles analyzed by *t*-test ( $P < 0.05$ ).

Braun et al. 2014, Hsu et al. 2014, Subrahmanian et al. 2016, Xiu et al. 2016, Cai et al. 2017, Weißenberger et al. 2017, Sun et al. 2019). Accordingly, respiratory CI occurs only in trace amounts in *emb2794* plants. We suggest that the absence, or a large reduction, of CI activity is directly linked to the embryo-development arrest phenotypes of *emb2794* mutants. These assumptions are supported by previous analyses of Arabidopsis and maize affected in CI biogenesis or activity, which also show notable defects in embryo development

(Kühn et al. 2015, Córdoba et al. 2016, Fromm et al. 2016, Ostersetzer-Biran 2016, Best et al. 2019, Ren et al. 2019, Sun et al. 2019), regardless of the fact that some plants, e.g. mistletoe, are able to grow in the absence of CI in their mitochondria, although in a semiparasitic way (Maclean et al. 2018, Senkler et al. 2018).

### EMB2794 is a factor involved in the *trans*-splicing of *nad2* transcripts

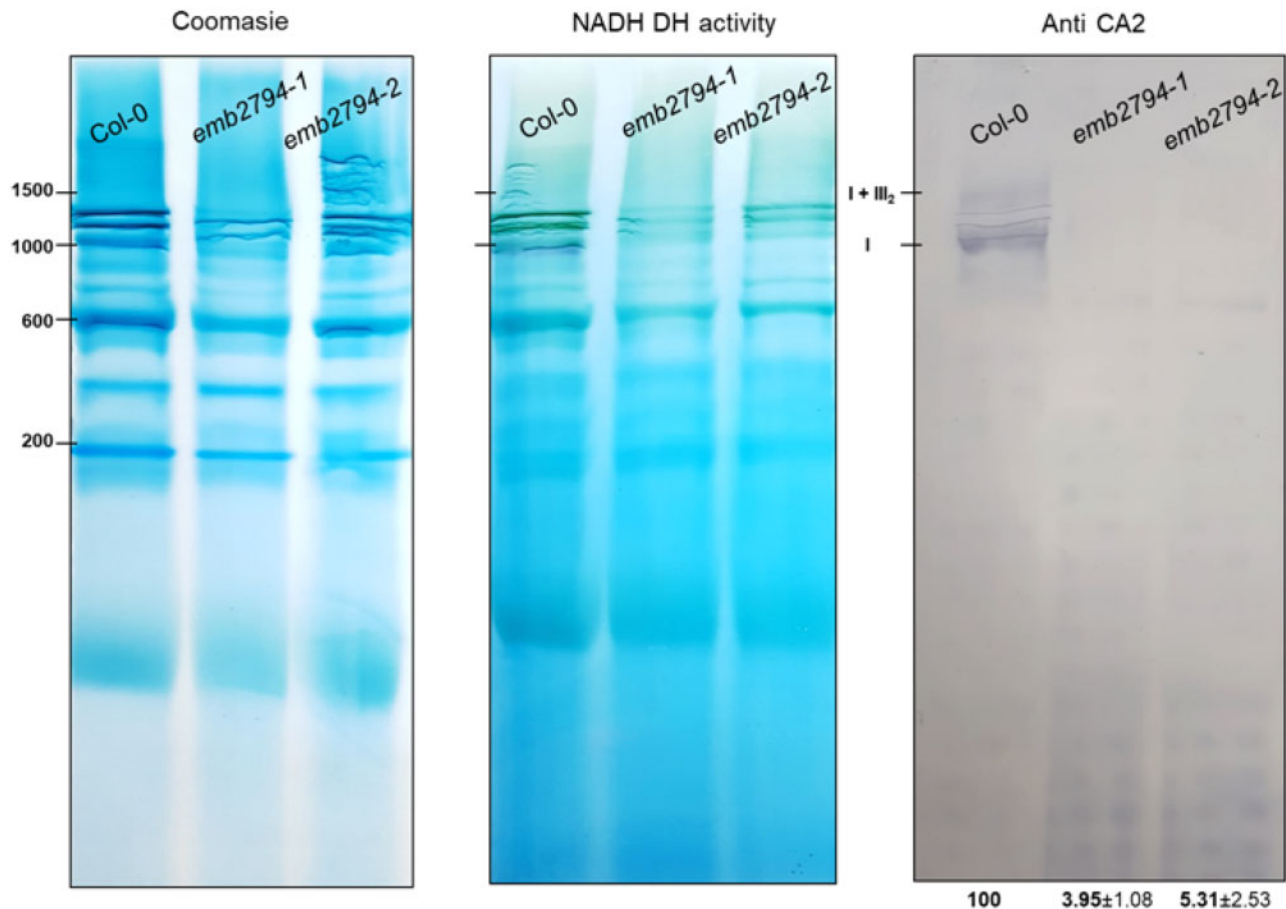
Using the amiR *emb2794* lines to perform global RNA analyses, we could not detect any consistent variation in transcript abundance, presumably because silencing was not complete and residual *EMB2794* transcripts may be efficiently translated at a level high enough to assure its function. On the contrary, analyses of the mitochondrial RNA profiles of *emb2794* homozygous mutants indicate defects in the *trans*-splicing of *nad2* intron 2, which was evident by the notable accumulation of unspliced *nad2* pre-RNA species and the strong reduction in the corresponding mature forms of *nad2* (Fig. 3). Furthermore, the observed decay in the steady state of the ex4 to ex5 seems to be a consequence of a failure in the *trans*-splicing between two pre-transcripts because accumulation of unspliced ex2–3 is only detected by this method. However, we could not rule out a possible function in RNA stability.

Based on the predicted protein sequences, predicted binding sites (Supplementary Fig. S8) and the in vitro and in vivo activity assays (Fig. 4), it is strongly anticipated that EMB2794 is directly associated with *nad2* pre-RNA transcripts and thereby facilitates the maturation of *nad2* intron 2. Because we found putative binding sites in both pre-transcripts, we speculate that the entire PPR protein is binding to each transcript; thus, one PPR protein is required per transcript using all 17 repeats. This function of EMB2794 may involve a direct role in the splicing reaction, or might be accomplished by affecting the structure as well as the stability of *nad2* pre-RNAs. By in silico predictions, we found putative binding sites in other *nad* genes (*nad1*, *nad4*, *nad5* and *nad7*) and also in other mitochondrial genes like *atp8* and *rps3*, 7 and 12; however, this is not consistent with the results obtained in the transcriptome. Low complexity of the predicted binding sites may explain the large number of predicted RNA targets. Some of the transcripts of these genes, particularly *nad1* and *nad4*, were found in in vitro and in vivo binding assays suggesting that the protein is able to bind these transcripts but, depending on transcription rates, levels of EMB2794 protein, RNA secondary structures and possible other factors, these interactions are precluded in the WT.

### Active assembled CI is absent in *emb2794* mutant plants

In plants, like in other organisms studied to date, the assembly of the membrane arm of CI involves the joining together of an intermediate containing NAD1 with an intermediate containing NAD2, followed by the addition of a module containing NAD4 and NAD5 subunits (Ligas et al. 2019). A major difference in CI biogenesis between the plant and mammalian pathways involves the CA domain, which is present in plants and proposed in other eukaryotes to be essential for the initial step of





**Fig. 6** Inactive and smaller complex I species are detected in homozygous *emb2794* plants. (A) Enriched mitochondrial fractions were isolated from leaves of Col-0 and rescued *emb2794* plants. After solubilization with digitonin, membrane protein complexes were separated on Blue-Native PAGE and stained with colloidal Coomassie blue G250. (B) In gel, NADH DH activity was determined in duplicate gels. Molecular weight (kDa) markers are indicated at the left. (C) Western blot analyses using antiCA2 antibodies to mark complex I species.

the membrane arm assembly (Perales et al. 2005, Meyer et al. 2011), while it does not exist in Opisthokonta (animals and fungi).

The embryo-lethal phenotype of *emb2794* is thought to be caused by the lack of an active CI. In fact, this mutant was named *emb* (Meinke et al. 2008), for embryo lethality, because the abnormal *emb2794* homozygous seeds are unable to germinate on the soil. In this study, we show that the embryogenesis process is not fully accomplished and that mutant plants show developmentally delayed embryos from heart stage onwards. These delayed embryos never undergo the cotyledon greening process, so they are pale in color and become dark brown reaching the Early Curved Torpedo final stage. Since *nad2* transcripts are compromised in the *emb2794* mutants, it is anticipated that NAD2 protein would be strongly reduced in the mutants. In this scenario, the biogenesis of the membrane arm of CI is expected to be arrested during the early assembly steps (Liu et al. 2010, Kühn et al. 2011, Braun et al. 2014, Hsu et al. 2014, Subrahmanian et al. 2016, Xiu et al. 2016, Cai et al. 2017, Weißenberger et al. 2017, Sun et al. 2019). Indeed, knock-out rescued plants did not show any NADH dehydrogenase activity, while a small intermediate (i.e. partial assembled) CI

particle of about 150 kDa was detected in leaf extracts in some experiments (Supplementary Fig. S9). Recently, Ligas et al. (2019) postulated that early CI membrane arm assembly involves the assembly of an 85-kDa intermediate composed of gamma-type carbonic anhydrases, i.e. the CA domain, most likely trimers of one CAL and two CAs (CA1, CA2 or CA3) (Córdoba et al. 2016, Fromm et al. 2016, Córdoba et al. 2019). The following step includes the formation of a 200-kDa intermediate composed of P2 (a plant-specific protein), NDUFC2, 20.9 kDa and NAD2 proteins. We speculate that the ~150-kDa intermediate we detected, which contains at least CA2 (Supplementary Fig. S9), is accumulated due to the impaired membrane arm assembly in the absence, or large reduction, of NAD2 caused by the strong splicing defects in *nad2* intron 2. Nevertheless, traces of nearly normal size of CI are also detected, but these complexes seem to be inactive. These may represent holo-CI particles that contain tiny amounts of NAD2 subunits, which were also evident in the RT-PCR analyses (Figs 3, 6). Unexpectedly, the arrested *emb2794* embryos show a nearly normal mitochondrial membrane potential. This might relate to alternative electron flow derived from alternative NADH dehydrogenases, which might eventually go through complexes

III and IV. Moreover, AOXs 1/2 are strongly induced in the *emb2794* mutants (**Supplementary Fig. S10**). Both alternative proteins are known to accumulate to high levels in Arabidopsis plants affected in CI biogenesis or activity.

A very similar phenotype was recently reported in maize (**Sun et al. 2019**), in which the *emp12* mutant affecting a PPR with 10 repeats and showing embryogenesis arrest displayed severely affected *trans*-splicing of intron 2, and *cis*-splicing of intron 1 and 4. As a result, CI assembly is perturbed. However, ZmEMP12 and AtEMB2794 are different PPR proteins with non-homologous binding sites. Since a similar binding site was found in maize *nad2* intron 2 (for a putative ortholog PRJNA10769), both factors are essential and might work together for the *nad2* *trans*-splicing event.

### Lack of CI leads to embryo lethality in the *emb2794* mutants

Many mutants with defects in the splicing or stability of *nad* transcripts exhibit severely impaired CI assembly and activity and show visual phenotypes such as reduced germination, retarded growth or curled leaves and in some cases embryo arrest phenotypes.

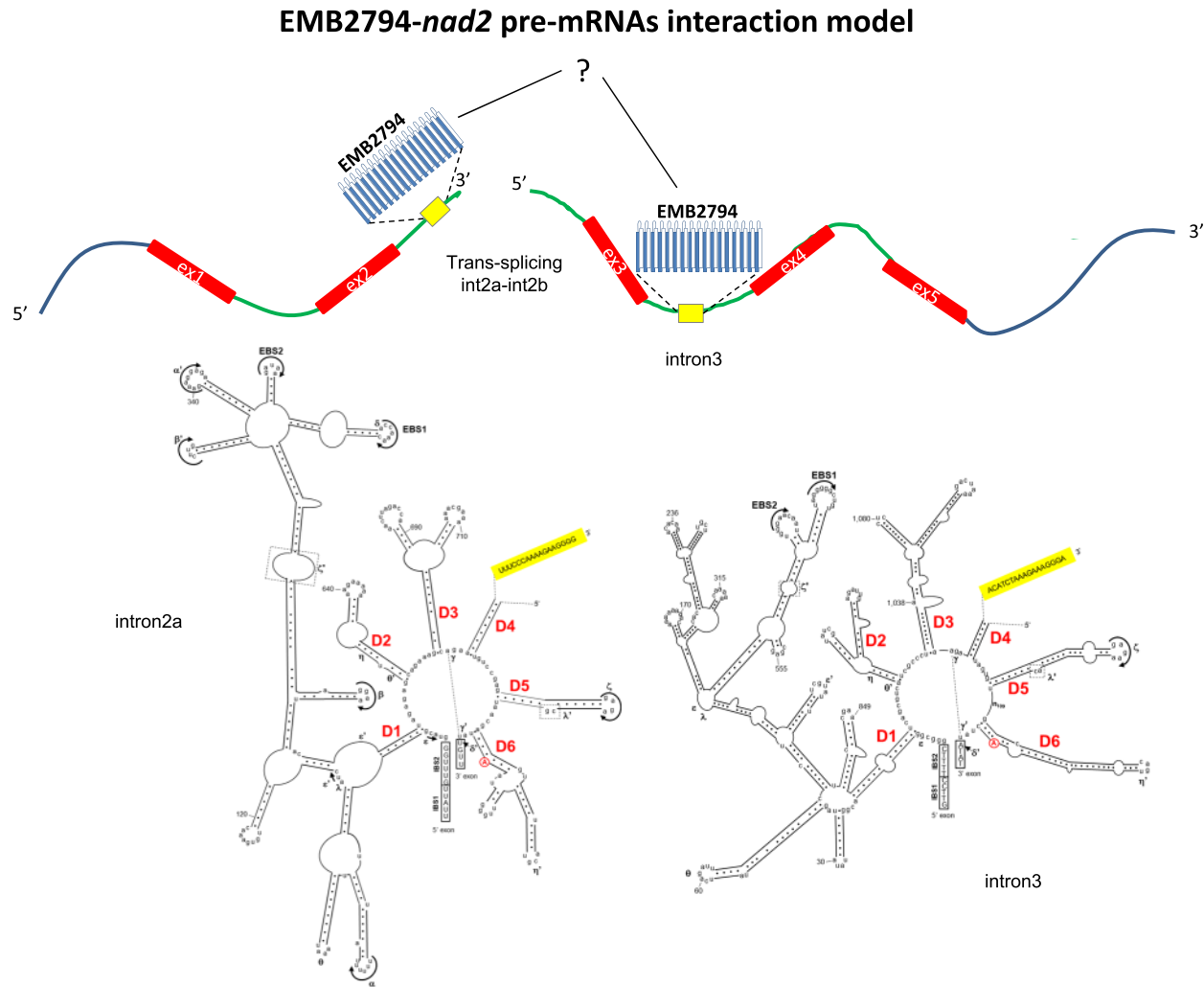
Mutants affecting *nad1* splicing (**Falcon de Longevialle et al. 2007**, **Keren et al. 2009**, **Köhler et al. 2010**, **Keren et al. 2012**, **Zmudjak et al. 2013, 2017**, **Cohen et al. 2014**, **Haïli et al. 2016**, **Lee et al. 2017**, **Ren et al. 2017**, **Wang et al. 2017, 2018**) have been extensively studied showing the complexity of this particular *trans*-splicing process. Interestingly, PMH2, mCSF1 and nMAT2 are involved in the maturation of numerous mitochondrial intron-containing pre-RNAs and thus may serve as key components of an early proto-spliceosomal system (**Schmitz-Linneweber et al. 2015**, **Zmudjak et al. 2017**). In addition to these factors, the mitochondria-encoded MatR protein (**Sultan et al. 2016**) and nMAT4 (**Cohen et al. 2014**) were shown to be essential for the *trans*-splicing of both *nad1* introns 3 and 4, where intron 3 is found in a *trans* configuration, while intron 4 is *cis*-spliced. The RADIATION SENSITIVE RAD52-like protein ORGANELLAR DNA-BINDING PROTEIN 1 (ODB1), originally identified as a mitochondrial DNA repair component, has been shown to be also required for the excision of two *cis*-spliced group II introns, *nad1* intron 2 and *nad2* intron 1 (**Gualberto et al. 2015**).

Other P-class PPR proteins together with other splicing factors have been shown to play key roles in CI biogenesis. ABA OVERLY SENSITIVE 5 (ABO5) (**Liu et al. 2010**) and MITOCHONDRIAL INTRON SPLICING FACTOR 26 (MISF26, **Wang et al. 2018**) are required for the *cis*-splicing of *nad2* intron 3 (together with the mitochondrial TRANSCRIPTION TERMINATION FACTOR 15 mTERF15), (**Hsu et al. 2014**). In maize, DEFECTIVE KERNEL 37 (DEK37) and PPR20 were proposed to be required for the *cis*-splicing of *nad2* intron 1 and 3, respectively, affecting seed development (**Dai et al. 2018**, **Yang et al. 2019**). RCC1/UVR8/GEF-Like 3 (RUG3), a protein related to the human guanine nucleotide exchange factor, was also shown to be an important but nonessential factor for the splicing of *nad2* intron 3 (**Kühn et al. 2011**). The MITOCHONDRIAL

STABILITY FACTOR 1 (MTSF1) is involved in stabilizing the *nad4* mRNA (**Haïli et al. 2013**). ABO8 functions in the splicing of *nad4* pre-mRNA (**Yang et al. 2014**), ORGANELLE TRANSCRIPT PROCESSING (OTP439) and TANG2 in the *trans*-splicing of *nad5* introns 2 and 3 (**Colas Des Francs-Small et al. 2014**). In maize, DEK35 and DEK43 (dual targeted to nucleus and mitochondria) have shown to be involved in the splicing of *nad4* intron 1 and introns 1 and 3, respectively (**Chen et al. 2017**, **R. Ren et al. 2020**, **Z. Ren et al. 2019**). BUTHIONINE SULFOMIXINE-INSENSITIVE ROOTS 6 (BIR6) (**Koprivova et al. 2010**) and SLOW GROWTH3 (**Hsieh et al. 2015**) were shown to be involved in the splicing of *nad7* introns 1 and 2, respectively. **Haïli et al. (2016)** showed that the PPR Mitochondrial Translation Factor 1 (MTL1) is involved in the translation of NAD7. In addition, the *misf68* mutant affects the processing of *nad2* intron 2 and *nad4* intron 1 and *nad5* intron 4 (**Wang et al. 2018**). Consequently, in many mutant plants affected in the processing of transcripts coding for NAD subunits, CI assembly is compromised and, consequently, seed germination and/or plant growth and development are strongly affected. The *emb2794* mutants show affected *nad2* intron 2 *trans*-splicing, no NADH DH activity and impaired CI assembly. However, these mutants show a more severe phenotype than other mutants affecting *nad 2* transcripts and it is probably due to the complete absence of NADH DH activity shown by *emb2794* mutants. Many mutants affecting CI assembly were shown to contain traces of NADH DH activity (**Keren et al. 2009**, **Keren et al. 2012**, **Cohen et al. 2014**). Moreover, mutants like *ndufs4* and *ndufv1* (**Meyer et al. 2009**, **Kühn et al. 2015**) contain an accumulation of assembly intermediates of CI without any detectable NADH dehydrogenase activity. Only in the case of *ndufs4*, traces of fully assembled CI are present while, in *ndufv1*, CI is not detected. Consequently, germination in the absence of sugar was not much affected in *ndufs4* mutants, while it is strongly affected and most seedlings died only in *ndufv1* mutants. In the case of *emb2794* mutants, trace amounts of CI were sometimes detected but completely inactive and germination without sugar is dramatically affected. With a combination of 1% and 3% sucrose, a small proportion of mutant plants is able to be rescued; however, adult mutant plants show a strong defective phenotype.

### EMB2794 is the important factor for maturation of *nad2* transcripts

According to our in vitro and in vivo experiments, recombinant or transgenic EMB2794 proteins are able to bind the two immature *nad2* transcripts (**Fig. 7**). A secondary structure was generated by the prediction algorithms showing that both putative binding sites are located at similar positions within both introns, near the D4 subdomain (**Fig. 7**). Putative binding sites are conserved in some dicotyledonous as well as monocotyledonous plants (**Supplementary Fig. S11**). However, the binding site located in intron 2a corresponding to the first transcript appears to be more conserved than the one located in intron 3 of the second transcript. Indeed, a similar binding site is in a different location within *nad2* intron 3 of many dicotyledonous species. In a recent study, **Chen et al. (2019)** have shown that



**Fig. 7** Model of *nad2* transcript-EMB2794 protein interaction. Upper panel: schematic representation of two *nad2* pre-transcripts. Solid lines indicate RNA, in blue 5' and 3' UTRs, in green introns. Red boxes indicate exons. Green boxes indicate the position of putative EMB2794-binding sites located in both pre-transcripts, one in intron 2a and the other in intron 3. All 17 PPRs are proposed to be involved in the binding of the putative binding sites. Lower panel: secondary structure model of *Arabidopsis nad2* introns 2 and 3. The conserved domains in the second (A) and third (B) introns of *Arabidopsis thaliana nad2* gene are illustrated in a secondary structure model. Each subdomain of D1–D6 is outlined within the structure. The conserved bulged-A residue in D6, the exon–intron-binding sites (i.e. EBS1/IBS1 and EBS2/IBS2) and tertiary interactions between different intron regions (indicated by roman letters) are shown in the model structure. The predicted binding sites of EMB2794 are illustrated by yellow boxes. The putative secondary structures of *nad2* i2 and i3 were generated by the prediction algorithms 'Alifold-Vienna RNA Webserver' (Hofacker et al. 2002) and 'the Mfold Web Server' (Zuker 2003), based on different *nad2* sequences from different angiosperms available on the National center for Biotechnology Information 'Organelle Genome Resources' server.

ZmPPR-SMR1 is required for the splicing of 75% mitochondrial group II introns and Zm-mCSF1 (a CRM domain containing protein) for at least six introns in maize. Both proteins interact and are implicated in *nad2* intron 2 *trans*-splicing, the same splicing event we suggest EMB2794 is involved in, in *Arabidopsis*. However, EMB2794 is suggested to act only on *nad2* transcripts and ZmPPR-SMR1 is a more general factor. A possible ortholog of EMB2794 in maize is the AQR38883.1 sequence, which has 19 repeats predicted and no SMR domain. Conversely, ZmPPR-SMR1 is the ortholog of the *Arabidopsis* At1g79490 gene, which contains an SMR domain and the mutant is annotated as EMB2217 protein (Meinke 2020).

In the homozygous rescued mutant plants, the lack of EMB2794 protein leads to an accumulation of unspliced *nad2* transcripts and a strong decrease in spliced *nad2* ex2–3 transcripts, suggesting that the *trans*-splicing involving pre-transcripts containing ex1–2 and, on the other side, ex3–5 is not able to be spliced together. Furthermore, the second pre-transcript is strongly diminished in the mutants suggesting a role in stability when EMB2794 is bonded presumably in the intron 3 binding site. Consequently, NAD2 protein is strongly diminished, assembly of CI is altered and traces are formed that are completely inactive leading to a strong failure in embryo development. Thus, we propose that EMB2794 is a splicing

factor required for the *trans*-splicing and stability of *nad2* transcripts.

## Materials and Methods

### Plant material

*Arabidopsis thaliana* Ecotype Columbia (Col-0) WT, *rdr6* lines and the mutants described were grown under long-day conditions (16/8 h L/D), 100  $\mu\text{mol}$  quanta ( $\mu\text{E}$ )  $\text{m}^{-2} \text{s}^{-1}$  light intensity and at 22°C constant temperature. After sowing seeds were maintained in darkness, at 4°C for 2 d. Single-mutant hemizygous lines for the *EMB2794* gene used in this work were *emb2794-1* (SALK\_037336) and *emb2794-2* (SAIL\_359\_F11). Growth on plate was conducted on Murashige and Skoog medium (Sigma, sigmaaldrich.com), supplemented with 30  $\mu\text{g}/\text{ml}$  kanamycin or 20  $\mu\text{g}/\text{ml}$  hygromycin or ammonium glufosinate 6  $\mu\text{g}/\text{ml}$  depending on the vector used. Seeds were previously sterilized on SDS-sodium hypochlorite solution and washed five times with sterile water. Homozygous *emb2794* plants could be rescued only if shrunken seeds were sown on MS media supplemented with 1% sucrose during 10 d and then changed to 3% for another 10 d. Some small seedlings germinate after 10 d, and we could maintain them after flowering, but no viable seeds develop.

### Complementation

Genomic DNA containing the entire coding sequence (2,286 bp) of the At2g02150 (*EMB2794*) gene was synthesized (Gene Universal Inc., China) together with *attL* sequences (5' and 3') for recombination in a pUC57 plasmid. It was used for a clonase reaction (Gateway Clonase LR II) together with a destination plasmid containing the *UBQ10* regulatory region and a terminator sequence derived from plasmid pUB-DEST (Grefen et al. 2010). The resulting vector was transformed into *Agrobacterium tumefaciens* GV3101 (pMP90) and used for floral dip transformation (Clough and Bent 1998). Resulting seeds were sown on MS plates containing 30  $\mu\text{g}/\text{ml}$  kanamycin and 6  $\mu\text{g}/\text{ml}$  BASTA to select both, the T-DNA insertion on *EMB2794* gene and the transgene *UBQ10*: *EMB2794* construct used to complement. A second vector containing the own promoter, 1,431 bp, *EMB2794::EMB2794*, was used for complementation experiments. Selected plants were genotyped, and siliques of the complemented homozygous *emb2794* lines were subjected to phenotypic analysis.

### Genotyping of insertional mutant lines

Detection of insertional T-DNA on mutant lines was performed by genomic PCR methods (5 min at 94°C, 35 cycles of 30 s at 94°C, 30 s at 50°C, 1.15 min at 72°C and 5 min at 72°C), using specific primers listed in Supplementary Fig. S1.

### Phenotype analysis

WT and mutant (*emb2794-1* and *emb2794-2*) siliques (corresponding to different stages of embryogenesis, Bowman 1994) were mounted on glass slides and cleared for 16 h on Hoyer's solution. After clearing, embryos were visualized under a Zeiss Axioplan Imager-A2 microscope under differential interference contrast optics. Images were captured with an Axiocam HRC charge-coupled device camera (Zeiss) using the AxioVision program (version 3.1).

### Subcellular localization

To determine the subcellular localization of *EMB2794*, the entire coding sequence was subcloned into the destination vector pH7FWG2 (Karimi et al. 2002) using a clonase reaction. The resulting construct containing 35S: *EMB2794*-GFP was used to transform *A. tumefaciens* and then infiltrate *N. benthamiana* leaves. In addition, the same construct was stably introduced in *Arabidopsis* using the floral dip methods (Clough and Bent 1998). Red MitoTracker was used as a mitochondria marker (Thermo Fischer Scientific). Samples were examined under confocal microscopy (Nikon Eclipse C1 Plus) and quantified using ImageJ software.

### Quantitative RT-PCR analysis

Total RNA was extracted with TRIzol reagent (Invitrogen, thermofisher.com). RNA was digested with RQ1 DNase (RQ1 RNase-Free DNase; Promega (worldwide.promega.com)) and used for cDNA synthesis with ImPromII RT (ImProm-

II<sup>TM</sup> Reverse Transcription System; Promega). cDNA synthesis (Fig. 3A, B) was carried out with the Superscript III reverse transcriptase (Invitrogen), using 1–2  $\mu\text{g}$  of total RNA and 250 ng of a mixture of random hexanucleotides (Promega) and incubated for 50 min at 50°C. Reactions were stopped by 15 min incubation at 70°C, and the RT samples served directly for real-time PCR. qPCRs were run on a LightCycler 480 (Roche), using 2.5  $\mu\text{l}$  of LightCycler 480 SYBR Green I Master mix and 2.5  $\mu\text{M}$  forward and reverse primers in a final volume of 5  $\mu\text{l}$ . Reactions were performed in triplicate in the following conditions: preheating at 95°C for 10 min, followed by 40 cycles of 10 s at 95°C, 10 s at 58°C and 10 s at 72°C. The nucleus-encoded 18S rRNA (At3g41768) and mitochondrial 26S rRNA (*rrn26*) genes were used as reference genes in the qPCR analyses. Primers used are listed in Supplementary Fig. S1.

### Silencing strategy and mitochondrial transcriptome

An amiR strategy was used. Two independent sequences targeting *EMB2794* gene were designed by Web MicroRNA designer (<http://wmd3.weigelworld.org/cgi-bin/webapp.cgi>). These sequences were assembled into the natural miR390, and the entire sequence was synthesized (Genscript, [www.genscript.com](http://www.genscript.com)). Then, they were cloned together into pCHF3 binary vector containing the 35S CaMV promoter and transformed into *rdr6* mutant plants by floral dip method (Clough and Bent 1998). Several independent lines were obtained with different degrees of silencing for *EMB2794* gene. Two lines, one for each amiR with the strongest silencing, were chosen for further analyses. Total RNA was extracted from leaves as explained and subjected to a mitochondrial transcriptome essentially as described in Zmudjak et al. (2013, 2018).

### Mitochondria isolation, endomembrane separation and gel electrophoresis procedures

Mitochondrial membrane protein complexes from plants (leaves) were isolated as outlined in Pineau et al. (2008). After that, proteins were separated as described above for 1D gels. Membrane protein complexes were solubilized in digitonin. Ten milligrams of mitochondria (~1 mg mitochondrial proteins) were mixed and incubated with 2.5 mg of digitonin in 100  $\mu\text{l}$  (Eubel et al. 2003) previous to electrophoresis separation. One-dimension blue native electrophoresis was performed as described in Wittig et al. (2006), in large, polyacrylamide gradient (4–16% p/v) gels. After 18 h running at 500 V, 15 mA, lanes were stained overnight (ON) in colloidal Coomassie blue G-250. Duplicate gels were transferred to PDF membranes and incubated with anti-CA2 antiserum as indicated (Perales et al. 2005).

Immune detection of AOX isoforms was performed on proteins obtained from Col-0, *emb2794-1* and *emb2794-2* rescued plants. After membrane protein complexes solubilization in digitonin, proteins were treated under denaturing conditions (5 min at 98°C heating and  $\beta$ -mercaptoethanol) and then separated in 12% polyacrylamide SDS-PAGE. Protein transfer and immunodetection were performed as described previously (Córdoba et al. 2019). Briefly, polyacrylamide gels were transferred to nitrocellulose membranes (GE Healthcare) and incubated with these primary antibodies and then with an anti-rabbit secondary antibody (Sigma) conjugated with alkaline phosphatase, following standard protocols. The final development was carried out with NBT/BCiP (Promega). We used anti-AOX primary polyclonal antibodies from Agriser (agrisera.com; Sweden, Catalog AS04 054).

### NADH dehydrogenase activity staining

Separation of protein complexes under native conditions allows in-gel activity staining. NADH dehydrogenase activity was assayed incubating the gels with 0.14 mM NADH (Sigma) and 1 mg/ml NBT in 0.1 M Tris-HCl (pH 7.4) until violet band is clearly viewed. After that, gels were fixed in 40% v/v methanol and 10% v/v acetic acid.

### Mitochondrial membrane potential

To analyze the mitochondrial membrane state, tetramethylrhodamine, methyl ester (TMRM) (Molecular Probes) was used. Stock solutions were prepared as 1000 $\times$  solutions in di methyl sulfoxide and stored at  $-20^\circ\text{C}$ . Immature pistils were emasculated and pollinated manually. Embryos at torpedo stage were isolated from seed coats and incubated with 200 nM TMRM on 0.5 $\times$  MS medium, 3% sucrose. After 30 min incubation, seeds were mounted with buffer and

images were taken under a confocal microscope (Nikon Eclipse C1 Plus). Red (for TMRM, excitation/emission wavelength = 485/590 nm) fluorescence was analyzed for embryos from WT and mutant plants.

### Target site prediction of EMB2794

Target site prediction of EMB2794 was performed as described by Takenaka et al. (2013). Briefly, the combination of the amino acids at positions 5 and 35 of the individual 17 PPR elements was compared to the nucleotide most often associated with this set in the PPR elements from previously identified RNA-editing factors. In addition to the major amino acid combinations at 5th and 35th, three minor combinations, GN, TT and CN, of which the nucleotide preferences were experimentally examined (Yan et al. 2019) were also employed for the target prediction. Using the profiles for the 17 PPR elements in the EMB2794, the probability to match the respective nucleotide bias is calculated for Arabidopsis Col-0 mitochondrial genome (GI: BK010421.1) by the FIMO program (Grant et al. 2011). Among the putative binding sequences of EMB2794 ( $P$ -value  $< 1.0 \times 10^{-4}$ ), only mapped on the transcribed regions were listed up (Supplementary Fig. S8).

### Protein–RNA-binding assay

EMB2794 was expressed in *E. coli* as a fusion with Recombinant MBP using 1 mM IPTG and 18°C ON. The MBP–EMB2794 protein was loaded into an amylose resin that binds MBP. Total RNA extracted from leaves of WT plants was applied to the column and incubated for 30 min with EMB2794 attached to the resin. Several washing steps were required to clean RNA in excess which was not bound. TEV protease was used to cleave a TEV site between MBP and EMB2794. The PPR–RNA complex was thus released from the resin. Aliquots obtained from input loaded in the resin (MBP–PPR), TEV cleavage (PPR–RNA complex) and elution of MBP from the resin were analyzed by Western blot with antibody against MBP to check the correct cleavage of TEV site. Once PPR–RNA complexes were obtained, RNA was isolated using TRIzol method and cDNA was synthesized. MBP was used as a negative control of the experiment. The cDNA thus synthesized was used as a template for RT-PCRs. To detect RNA targets, primers for specific mitochondrial transcripts over exons or introns were used. The cDNA synthesized from RNA of WT plants and the cDNA from RNA derived from MBP columns were used as positive and negative controls, respectively.

For in vivo analysis, we used GFP-trap agarose beads (Chromotek, Germany) with the capacity to bind GFP or proteins fused to GFP. We use transgenic plants containing 35S::EMB2794-GFP, which are able to express the fusion protein in leaves. One hundred milligrams of leaves were extracted with a buffer containing sodium deoxycholate to stabilize RNA possible bonded to the fusion protein. After incubation according to the manufacturer, RNA was isolated using TRIzol, cDNA synthesized and PCR performed using cognate primers for the indicated transcripts. At least three lines were used.

### Supplementary Data

Supplementary data are available at PCP online.

### Acknowledgments

F.M. and M.C. are doctoral fellows and J.P.C. was a postdoctoral fellow of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). G.P. and E.Z. are CONICET Researchers. M.T. and O.O.-B. are associate professors. We are grateful to Daniela Villamonte and Viviana Daniel for excellent technical assistance with confocal microscopy.

### Funding

Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, PICT16 0036); Consejo Nacional de Investigación

Científica y Tecnológica (CONICET); Howard Hughes Medical Institute (HHMI 55007430); ‘Israeli Science Foundation’ (ISF 741/15).

### Disclosure

The authors have no conflicts of interest to declare.

### References

- Barkan, A. and Small, I. (2014) Pentatricopeptide repeat proteins in plants. *Annu. Rev. Plant Biol.* 65: 415–442.
- Best, C., Zmudjak, M. and Ostersetzer, O. (2019) The PPR-related splicing cofactor MSP1/EMB1025 protein, encoded by At4g20090, encode an essential protein that is required for the splicing of *nad1* intron 1 and for the biogenesis of complex I in Arabidopsis mitochondria. *bioRxiv*, 615856.
- Bonen, L. (2008) Cis- and trans-splicing of group II introns in plant mitochondria. *Mitochondrion* 8: 26–34.
- Bowman, J. (1994) Arabidopsis: An Atlas of Morphology and Development. Springer.
- Brangeon, J., Sabar, M., Gutierrez, S., Combettes, B., Bove, J., Gendy, C., et al. (2000) Defective splicing of the first *nad4* intron is associated with lack of several complex I subunits in the *Nicotiana sylvestris* NMS1 nuclear mutant. *Plant J.* 21: 269–280.
- Braun, H.-P., Binder, S., Brennicke, A., Eubel, H., Fernie, A.R., Finkemeier, I., et al. (2014) The life of plant mitochondrial complex I. *Mitochondrion* 19 Pt B: 295–313.
- Brown, G.G., Colas Des Francs-Small, C. and Ostersetzer-Biran, O. (2014) Group II intron splicing factors in plant mitochondria. *Front. Plant Sci.* 5: 35.
- Cai, M., Li, S., Sun, F., Sun, Q., Zhao, H., Ren, X., et al. (2017) Emp10 encodes a mitochondrial PPR protein that affects the cis-splicing of *nad2* intron 1 and seed development in maize. *Plant J.* 91: 132–144.
- Carbonell, A., Takeda, A., Fahlgren, N., Johnson, S.C., Cuperus, J.T. and Carrington, J.C. (2014) New generation of artificial microRNA and synthetic trans-acting small interfering RNA vectors for efficient gene silencing in Arabidopsis. *Plant Physiol.* 165: 15–29.
- Chen, X., Feng, F., Qi, W., Xu, L., Yao, D., Wang, Q., et al. (2017) Dek35 encodes a PPR protein that affects cis-splicing of mitochondrial *nad4* intron 1 and seed development in maize. *Mol. Plant* 10: 427–441.
- Chen, Z., Wang, H.C., Shen, J., Sun, F., Wang, M., Xu, C., et al. (2019) PPR-SMR1 is required for the splicing of multiple mitochondrial introns, interacts with Zm-mCSF1, and is essential for seed development in maize. *J. Exp. Bot.* 70: 5245–5258.
- Cheng, S., Gutmann, B., Zhong, X., Ye, Y., Fisher, M.F., Bai, F., et al. (2016) Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. *Plant J.* 85: 532–547.
- Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735–743.
- Cohen, S., Zmudjak, M., Colas Des Francs-Small, C., Malik, S., Shaya, F., Keren, I., et al. (2014) nMAT4, a maturase factor required for *nad1* pre-mRNA processing and maturation, is essential for holocomplex I biogenesis in Arabidopsis mitochondria. *Plant J.* 78: 253–268.
- Colas Des Francs-Small, C., Falcon de Longevialle, A., Li, Y., Lowe, E., Tanz, S. K., Smith, C., et al. (2014) The pentatricopeptide repeat proteins TANG2 and ORGANELLE TRANSCRIPT PROCESSING439 are involved in the splicing of the multipartite *nad5* transcript encoding a subunit of mitochondrial complex I. *Plant Physiol.* 165: 1409–1416.
- Colas Des Francs-Small, C. and Small, I. (2014) Surrogate mutants for studying mitochondrially encoded functions. *Biochimie* 100: 234–242.

- Colcombet, J., Lopez-Obando, M., Heurtevin, L., Bernard, C., Martin, K., Berthomé, R., et al. (2013) Systematic study of subcellular localization of Arabidopsis PPR proteins confirms a massive targeting to organelles. *RNA Biol.* 10: 1557–1575.
- Coquille, S., Filipovska, A., Chia, T., Rajappa, L., Lingford, J.P., Razif, M.F., et al. (2014) An artificial PPR scaffold for programmable RNA recognition. *Nat. Commun.* 5: 5729.
- Córdoba, J.P., Fassolari, S., Marchetti, F., SotoD.B., Pagnussat, G.C., Zabaleta, E. (2019) Different types domains are present in complex I from immature seeds and of CA adult plants in *Arabidopsis thaliana*. *Plant Cell Physiol.* 60: 986–998.
- Córdoba, J.P., Marchetti, F., Soto, D., Martin, M.V., Pagnussat, G.C. and Zabaleta, E. (2016) The CA domain of the respiratory complex I is required for normal embryogenesis in *Arabidopsis thaliana*. *J. Exp. Bot.* 67: 1589–1603.
- Dai, D., Luan, S., Chen, X., Wang, Q., Feng, Y., Zhu, C., et al. (2018) Maize *Dek37* encodes a P-type PPR protein that affects cis-splicing of mitochondrial *nad2* intron 1 and seed development. *Genetics* 208: 1069–1082.
- Ding, Y.H., Liu, N.Y., Tang, Z.S., Liu, J. and Yang, W.C. (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.
- Eubel, H., Jansch, L. and Braun, H.-P. (2003) New insights into the respiratory chain of plant mitochondria. Supercomplexes and a unique composition of complex II. *Plant Physiol.* 133: 274–286.
- Falcon de Longevialle, A., Meyer, E.H., Andrés, C., Taylor, N.L., Lurin, C., Millar, A.H., et al. (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.
- Fromm, S., Braun, H.-P. and Peterhansel, C. (2016) Mitochondrial gamma carbonic anhydrases are required for complex I assembly and plant reproductive development. *New Phytol.* 211: 194–207.
- Grant, C.E., Bailey, T.L. and Noble, W.S. (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27: 1017–1018.
- Grefen, C., Donald, N., Hashimoto, K., Kudla, J., Schumacher, K. and Blatt, M. R. (2010) A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. *Plant J.* 64: 355–365.
- Gualberto, J.M., Le Ret, M., Beator, B. and Kühn, K. (2015) The RAD52-like protein ODB1 is required for the efficient excision of two mitochondrial introns spliced via first-step hydrolysis. *Nucleic Acids Res.* 43: 6500–6510.
- Haïli, N., Arnal, N., Quadrado, M., Amiar, S., Tcherkez, G., Dahan, J., et al. (2013) The pentatricopeptide repeat MTSF1 protein stabilizes the *nad4* mRNA in Arabidopsis mitochondria. *Nucleic Acids Research* 41: 6650–6663.
- Haïli, N., Planchard, N., Arnal, N., Quadrado, M., Vrielynck, N., Dahan, J., et al. (2016) The MTL1 pentatricopeptide repeat protein is required for both translation and splicing of the mitochondrial NADH DEHYDROGENASE SUBUNIT7 mRNA in Arabidopsis. *Plant Physiol.* 170: 354–366.
- Hammani, K., Gobert, A., Hleibieh, K., Choulier, L., Small, I. and Giegé, P. (2011) An Arabidopsis dual-localized pentatricopeptide repeat protein interacts with nuclear proteins involved in gene expression regulation. *Plant Cell* 23: 730–740.
- Hofacker, I.L., Fekete, M. and Stadler, P.F. (2002) Secondary structure prediction for aligned RNA sequences. *J. Mol. Biol.* 319: 1059–1066.
- Hsieh, W.Y., Liao, J.C., Chang, C.Y., Harrison, T., Boucher, C. and Hsieh, M.H. (2015) The SLOW GROWTH3 pentatricopeptide repeat protein is required for the splicing of mitochondrial NADH dehydrogenase subunit7 intron 2 in Arabidopsis. *Plant Physiol.* 168: 490–501.
- Hsu, Y.W., Wang, H.J., Hsieh, M.H., Hsieh, H.L. and Jauh, G.Y. (2014) Arabidopsis mTERF15 is required for mitochondrial *nad2* intron 3 splicing and functional complex I activity. *PLoS One* 9: e112360.
- Karimi, M., Inzé, D. and Depicker, A. (2002) GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.* 7: 193–195.
- Karpova, O.V. and Newton, K.J. (1999) A partially assembled complex I in NAD4-deficient mitochondria of maize. *Plant J.* 17: 511–521.
- Keren, I., Bezawork-Geleta, A., Kolton, M., Maayan, I., Belausov, E., Levy, M., et al. (2009) AtnMat2, a nuclear-encoded maturase required for splicing of group-II introns in Arabidopsis mitochondria. *RNA* 15: 2299–2311.
- Keren, I., Tal, L., Des Fracs-Small, C.C., Araújo, W.L., Shevtsov, S., Shaya, F., et al. (2012) nMAT1, a nuclear-encoded maturase involved in the trans-splicing of *nad1* intron 1, is essential for mitochondrial complex I assembly and function. *Plant J.* 71: 413–426.
- Köhler, D., Schmidt-Gattung, S. and Binder, S. (2010) The DEAD-box protein PMH2 is required for efficient group II intron splicing in mitochondria of *Arabidopsis thaliana*. *Plant Mol. Biol.* 72: 459–467.
- Koprivova, A., Colas Des Fracs-Small, C., Calder, G., Mugford, S.T., Tanz, S., Lee, B.R., et al. (2010) Identification of a pentatricopeptide repeat protein implicated in splicing of intron 1 of mitochondrial *nad7* transcripts. *J. Biol. Chem.* 285: 32192–32199.
- Kühn, K., Carrie, C., Giraud, E., Wang, Y., Meyer, E.H., Narsai, R., et al. (2011) The RCC1 family protein RUG3 is required for splicing of *nad2* and complex I biogenesis in mitochondria of *Arabidopsis thaliana*. *Plant J.* 67: 1067–1080.
- Kühn, K., Obata, T., Feher, K., Bock, R., Fernie, A.R. and Meyer, E.H. (2015) Complete mitochondrial complex I deficiency induces an up-regulation of respiratory fluxes that is abolished by traces of functional complex I. *Plant Physiol.* 168: 1537–1549.
- Lee, K., Han, J.H., Park, Y.I., Colas Des Fracs-Small, C., Small, I. and Kang, H. (2017) The mitochondrial pentatricopeptide repeat protein PPR19 is involved in the stabilization of NADH dehydrogenase 1 transcripts and is crucial for mitochondrial function and *Arabidopsis thaliana* development. *New Phytol.* 215: 202–216.
- Lee, B.H., Lee, H., Xiong, L. and Zhu, J.K. (2002) A mitochondrial complex I defect impairs cold-regulated nuclear gene expression. *Plant Cell* 14: 1235–1251.
- Ligas, J., Pineau, E., Bock, R., Huynen, M.A. and Meyer, E.H. (2019) The assembly pathway of complex I in *Arabidopsis thaliana*. *Plant J.* 97: 447–459.
- Li-Pook-Than, J. and Bonen, L. (2006) Multiple physical forms of excised group II intron RNAs in wheat mitochondria. *Nucleic Acids Res.* 34: 2782–2790.
- Liu, Y., He, J., Chen, Z., Ren, X., Hong, X. and Gong, Z. (2010) ABA over-sensitive 5 (ABO5), encoding a pentatricopeptide repeat protein required for cis-splicing of mitochondrial *nad2* intron 3, is involved in the abscisic acid response in Arabidopsis. *Plant J.* 63: 749–765.
- Lurin, C., Andrés, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyère, C., et al. (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16: 2089–2103.
- Maclean, A.E., Hertle, A.P., Ligas, J., Bock, R., Balk, J. and Meyer, E.H. (2018) Absence of complex I is associated with diminished respiratory chain function in European mistletoe. *Curr. Biol.* 28: 1614–1619.
- Matsuura, M., Noah, J.W. and Lambowitz, A.M. (2001) Mechanism of maturase-promoted group II intron splicing. *EMBO J.* 20: 7259–7270.
- Meinke, D., Muralla, R., Sweeney, C. and Dickerman, A. (2008) Identifying essential genes in *Arabidopsis thaliana*. *Trends Plant Sci.* 13: 483–491.
- Meinke, D.W. (2020) Genome-wide identification of EMBRYO-DEFECTIVE (EMB) genes required for growth and development in Arabidopsis. *New Phytol.* 226: 306–325.
- Meyer, E.H. (2012) Proteomic investigations of complex I composition: how to define a subunit? *Front. Plant Sci.* 3: 1–7.
- Meyer, E.H., Solheim, C., Tanz, S.K., Bonnard, G. and Millar, A.H. (2011) Insights into the composition and assembly of the membrane arm of plant complex I through analysis of subcomplexes in Arabidopsis mutant lines. *J. Biol. Chem.* 286: 26081–26092.
- Meyer, E.H., Tomaz, T., Carroll, A.J., Estavillo, G., Delannoy, E., Tanz, S.K., et al. (2009) Remodeled respiration in *ndufs4* with low phosphorylation efficiency suppresses Arabidopsis germination and growth and alters control of metabolism at night. *Plant Physiol.* 151: 603–619.

- Mohr, S., Matsuura, M., Perlman, P.S. and Lambowitz, A.M. (2006) A DEAD-box protein alone promotes group II intron splicing and reverse splicing by acting as an RNA chaperone. *Proc. Natl. Acad. Sci. USA* 103: 3569–3574.
- Nakagawa, N. and Sakurai, N. (2006) A mutation in At-nMat1a, which encodes a nuclear gene having high similarity to group II intron maturase, causes impaired splicing of mitochondrial NAD4 transcript and altered carbon metabolism in *Arabidopsis thaliana*. *Plant Cell Physiol.* 47: 772–783.
- Ostersetzer-Biran, O. (2016) Respiratory complex I and embryo development. *J. Exp. Bot.* 67: 1205–1207.
- Perales, M., Eubel, H., Heinemeyer, J., Colaneri, A., Zabaleta, E. and Braun, H.-P. (2005) Disruption of a nuclear gene encoding a mitochondrial gamma carbonic anhydrase reduces complex I and supercomplex I + III<sub>2</sub> levels and alters mitochondrial physiology in *Arabidopsis*. *J. Mol. Biol.* 350: 263–277.
- Pétriaccq, P., de Bont, L., Genestout, L., Hao, J., Laureau, C., Florez-Sarasa, I., et al. (2017) Photoperiod affects the phenotype of mitochondrial complex I mutants. *Plant Physiol.* 173: 434–455.
- Pineau, B., Layoune, O., Danon, A. and De Paepe, R. (2008) L-galactono-1,4-lactone dehydrogenase is required for the accumulation of plant respiratory complex I. *J. Biol. Chem.* 283: 32500–32505.
- Pla, M., Mathieu, C., De Paepe, R., Chétrit, P. and Vedel, F. (1995) Deletion of the last two exons of the mitochondrial *nad7* gene results in lack of the NAD7 polypeptide in a *Nicotiana sylvestris* CMS mutant. *Mol. Gen. Genet.* 248: 79–88.
- Ren, Z., Fan, K., Fang, T., Zhang, J., Yang, L., Wang, J., et al. (2019) Maize empty pericarp602 encodes a P-Type PPR protein that is essential for seed development. *Plant Cell Physiol.* 60: 1734–1746.
- Ren, X., Pan, Z., Zhao, H., Zhao, J., Cai, M., Li, J., et al. (2017) EMPTY PERICARP11 serves as a factor for splicing of mitochondrial *nad1* intron and is required to ensure proper seed development in maize. *J. Exp. Bot.* 68: 4571–4581.
- Ren, R.C., Wang, L.L., Zhang, L., Zhao, Y.J., Wu, J.W., Wei, Y.M., et al. (2020) DEK43 is a P-type pentatricopeptide repeat (PPR) protein responsible for the Cis-splicing of *nad4* in maize mitochondria. *J. Integr. Plant Biol.* 62: 299–313.
- Schmitz-Linneweber, C., Lampe, M.-K., Sultan, L.D. and Ostersetzer-Biran, O. (2015) Organellar maturases: a window into the evolution of the spliceosome. *Biochim. Biophys. Acta* 1847: 798–808.
- Schmitz-Linneweber, C. and Small, I. (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci.* 13: 663–670.
- Senkler, J., Rugen, N., Eubel, H., Hegermann, J. and Braun, H.-P. (2018) Absence of complex I implicates rearrangement of the respiratory chain in European mistletoe. *Curr. Biol.* 28: 1606–1613.
- Senkler, J., Senkler, M. and Braun, H.-P. (2017) Structure and function of complex I in animals and plants—a comparative view. *Physiol. Plant.* 161: 6–15.
- Shen, C., Zhang, D., Guan, Z., Liu, Y., Yang, Z., Yang, Y., et al. (2016) Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. *Nat. Commun.* 7: 11285.
- Small, I. and Peeters, N. (2000) The PPR motif—a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* 25: 46–47.
- Subrahmanian, N., Remacle, C. and Hamel, P.P. (2016) Plant mitochondrial Complex I composition and assembly: a review. *Biochim. Biophys. Acta* 1857: 1001–1014.
- Sultan, L.D., Mileshina, D., Grewe, F., Rolle, K., Abudraham, S., Głodowicz, P., et al. (2016) The reverse-transcriptase/RNA-maturase protein MatR is required for the splicing of various group II introns in Brassicaceae mitochondria. *Plant Cell* 28: 2805–2829.
- Sun, T., Bentolila, S. and Hanson, M.R. (2016) The unexpected diversity of plant organelle RNA editosomes. *Trends Plant Sci.* 21: 962–973.
- Sun, F., Xiu, Z., Jiang, R., Liu, Y., Zhang, X., Yang, Y.Z., et al. (2019) The mitochondrial pentatricopeptide repeat protein EMP12 is involved in the splicing of three *nad2* introns and seed development in maize. *J. Exp. Bot.* 70: 963–972.
- Sun, F., Zhang, X., Shen, Y., Wang, H., Liu, R., Wang, X., et al. (2018) The pentatricopeptide repeat protein EMPTY PERICARP8 is required for the splicing of three mitochondrial introns and seed development in maize. *Plant J.* 95: 919–932.
- Sung, T.Y., Tseng, C.C. and Hsieh, M.H. (2010) The SLO1 PPR protein is required for RNA editing at multiple sites with similar upstream sequences in *Arabidopsis* mitochondria. *Plant J.* 63: 499–511.
- Takenaka, M., Zehrmann, A., Brennicke, A. and Graichen, K. (2013) Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. *PLoS One* 8: e65343.
- Tzafir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., et al. (2004) Identification of genes required for embryo development in *Arabidopsis*. *Plant Physiol.* 135: 1206–1220.
- Unsel, M., Marienfeld, J.R., Brandt, P. and Brennicke, A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.* 15: 57–61.
- Wang, C., Aubé, F., Planchard, N., Quadrado, M., Dargel-Graffin, C., Nogué, F., et al. (2017) The pentatricopeptide repeat protein MTSF2 stabilizes a *nad1* precursor transcript and defines the 3' end of its 5'-half intron. *Nucleic Acids Res.* 45: 6119–6134.
- Wang, C., Aubé, F., Quadrado, M., Dargel-Graffin, C. and Mireau, H. (2018) Three new pentatricopeptide repeat proteins involved in splicing of mitochondrial transcripts and complex I biogenesis in *Arabidopsis thaliana*. *J. Exp. Bot.* 69: 5131–5140.
- Weissenberger, S., Soll, J. and Carrie, C. (2017) The PPR protein SLOW GROWTH 4 is involved in editing of *nad4* and affects the splicing of *nad2* intron 1. *Plant Mol. Biol.* 93: 355–368.
- Wittig, I., Braun, H.-P. and Schägger, H. (2006) Blue-native PAGE. *Nat. Protoc.* 1: 418–428.
- Wydro, M.M., Sharma, P., Foster, J.M., Bych, K., Meyer, E.H. and Balk, J. (2013) The evolutionarily conserved iron-sulfur protein INDH is required for complex I assembly and mitochondrial translation in *Arabidopsis*. *Plant Cell* 25: 4014–4027.
- Xiu, Z., Sun, F., Shen, Y., Zhang, X., Jiang, R., Bonnard, G., et al. (2016) EMPTY PERICARP 16 is required for mitochondrial *nad2* intron 4 cis-splicing, complex I assembly and seed development in maize. *Plant J.* 85: 507–519.
- Yan, J., Yinying, Y., Sixing, H., Yan, Y., Cuicui, S., Qunxia, Z., et al. (2019) Delineation of pentatricopeptide repeat codes for target RNA prediction. *Nucleic Acids Res.* 47: 3728–3738.
- Yang, Y.Z., Ding, S., Wang, Y., Wang, H.-C., Liu, X.-Y., Sun, F., et al. (2019) PPR20 is required for the cis-splicing of mitochondrial *nad2* intron 3 and seed development in maize. *Plant Cell Physiol.* 61: 370–380.
- Yang, L., Zhang, J., He, J., Qin, Y., Hua, D., Duan, Y., et al. (2014) ABA-mediated ROS in mitochondria regulate root meristem activity by controlling PLETHORA expression in *Arabidopsis*. *PLoS Genet.* 10: e1004791.
- Yin, P., Li, Q., Yan, C., Liu, Y., Liu, J., Yu, F., et al. (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* 504: 168–171.
- Zhang, Y. and Lu, C. (2019) The enigmatic roles of PPR-SMR proteins in plants. *Adv. Sci.* 6: 1900361.
- Zhou, W., Lu, Q., Li, Q., Wang, L., Ding, S., Zhang, A., et al. (2017) PPR-SMR protein SOT1 has RNA endonuclease activity. *Proc. Natl. Acad. Sci. USA* 114: E1554–E1563.
- Zmudjak, M., Colas Des Francs-Small, C., Keren, I., Shaya, F., Belasov, E., Small, I., et al. (2013) mCSF1, a nucleus-encoded CRM protein required for the processing of many mitochondrial introns, is involved in the biogenesis of respiratory complexes I and IV in *Arabidopsis*. *New Phytol.* 199: 379–394.
- Zmudjak, M. and Ostersetzer-Biran, O. (2017) Mitochondrial RNA metabolism and transcript regulation. *Ann. Plant Rev.* 50: 261–309.
- Zmudjak, M., Shevtsov, S., Sultan, L.D., Keren, I. and Ostersetzer-Biran, O. (2017) Analysis of the roles of the *Arabidopsis* nMAT2 and PMH2 proteins provided with new insights into the regulation of group II intron splicing in land-plant mitochondria. *Int. J. Mol. Sci.* 18: 2428–2425.
- Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31: 3406–3415.