6650–6663 *Nucleic Acids Research, 2013, Vol. 41, No. 13* *Published online 8 May 2013*
doi:10.1093/nar/gkt337

The pentatricopeptide repeat MTSF1 protein stabilizes the *nad4* mRNA in Arabidopsis mitochondria

Nawel Haïli[1,2], Nadège Arnal[1,2], Martine Quadrado[1,2], Souad Amiar[1], Guillaume Tcherkez[3,4], Jennifer Dahan[1,2], Pierre Briozzo[1,2], Catherine Colas des Francs-Small[5], Nathalie Vrielynck[1,2] and Hakim Mireau[1,2,*]

[1]INRA, UMR1318, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France, [2]AgroParisTech, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France, [3]Institut de Biologie des Plantes, CNRS UMR 8618, Bâtiment 630, Université Paris-Sud, 91405 Orsay Cedex, France, [4]Institut Universitaire de France, 103 Boulevard Saint-Michel, 75005 Paris, France and [5]Australian Research Council Centre of Excellence in Plant Energy Biology, University of Western Australia, Crawley WA 6009, Australia

Received November 15, 2012; Revised April 5, 2013; Accepted April 9, 2013

ABSTRACT

Gene expression in plant mitochondria involves a complex collaboration of transcription initiation and termination, as well as subsequent mRNA processing to produce mature mRNAs. In this study, we describe the function of the Arabidopsis mitochondrial stability factor 1 (*MTSF1*) gene and show that it encodes a pentatricopeptide repeat protein essential for the 3′-processing of mitochondrial *nad4* mRNA and its stability. The *nad4* mRNA is highly destabilized in Arabidopsis *mtsf1* mutant plants, which consequently accumulates low amounts of a truncated form of respiratory complex I. Biochemical and genetic analyses demonstrated that MTSF1 binds with high affinity to the last 20 nucleotides of *nad4* mRNA. Our data support a model for MTSF1 functioning in which its association with the last nucleotides of the *nad4* 3′ untranslated region stabilizes *nad4* mRNA. Additionally, strict conservation of the MTSF1-binding sites strongly suggests that the protective function of MTSF1 on *nad4* mRNA is conserved in dicots. These results demonstrate that the mRNA stabilization process initially identified in plastids, whereby proteins bound to RNA extremities constitute barriers to exoribonuclease progression occur in plant mitochondria to protect and concomitantly define the 3′ end of mature mitochondrial mRNAs. Our study also reveals that short RNA molecules corresponding to pentatricopeptide repeat-binding sites accumulate also in plant mitochondria.

INTRODUCTION

The production of functional RNA molecules is a multistep process that requires the participation of a high number of protein factors. In plant mitochondria and plastids, this process has increased in complexity throughout evolution, given the considerable number of post-transcriptional events that are required to generate and express mature RNAs. Post-transcriptional modifications include C-to-U RNA editing, cis- and trans-splicing of introns and 5′ and 3′ processing (1). RNA stabilization and ultimately RNA degradation are also important steps in regulating the quality and quantity of organellar RNAs and thus their functions. Most of these RNA modification and expression steps are still poorly understood at the molecular level in higher plant organelles, particularly in mitochondria. Nevertheless, the vast majority of involved proteins are encoded in the nucleus, which has superimposed complex gene expression machinery on the eubacterial core of these organelles throughout the evolution (1). Genetic analyses, especially in maize, Chlamydomonas and Arabidopsis, have revealed the prevalence of the pentatricopeptide repeat (PPR) proteins in organellar RNA metabolism [for review, see (2,3)]. PPR proteins have been associated with all RNA processing and expression steps as well as gene transcription in plant mitochondria and chloroplasts. These proteins are made up of loosely conserved 35 amino acid repeat motifs (4), and PPR proteins appear to be site-specific RNA-binding proteins with a strong preference for single-stranded RNA (5). Their mode of action may imply direct or indirect recruitment of effector proteins to specific RNA sites (6). It has been shown that PPR function can include local RNA structure reorganization to make otherwise hidden binding sites more accessible to

*To whom correspondence should be addressed. Tel: +33 1 30833070; Fax: +33 1 30833319; Email: mireau@versailles.inra.fr

other proteins or complexes (7). One PPR protein protects RNA from degradation in plastids by concealing an endoribonuclease cleavage site (5,8). Another important feature of PPR proteins is their considerable expansion in land plants: there are >400 PPR members in Angiosperms, and most of them are predicted to be transported into mitochondria or chloroplasts (6,9). An evolutionarily distinct subgroup of PPR proteins, called restorers of fertility (Rf), has also evolved notably to suppress the expression of mitochondrial genes involved in cytoplasmic male sterility (10,11).

Transcript end processing and the intricately connected process of RNA stabilization are fundamental steps in the production of mature RNA in plant organelles. In the case of plastids, genes are organized in operons, which give rise to the production of long polycistronic transcripts (12). Therefore, a complex pattern of 5′ and 3′ RNA processing involving both endonucleolytic and exonucleolytic cleavage is necessary to produce monocistronic plastid mRNAs; for review, see (12). There has been recent progress on the understanding of intercistronic RNA processing and mRNA stabilization in plastids. The current underlying model proposes that endoribonucleases cleave long plastid precursor RNAs in a relatively non-specific manner, and exoribonucleases then continue to degrade the RNAs until they encounter proteins, such as PPRs, located at specific sites at the ends of monocistronic mRNAs, thereby stopping the degradation process (3). This process was first revealed on analysis of the maize PPR10 protein (13), but other PPRs and other types of helical repeat proteins associated with the accumulation of processed chloroplastic transcripts likely have similar protective activity against exoribonuclease-mediated degradation (14–22). Housekeeping endoribonucleases, such as RNase E and possibly RNase J, have been proposed to be involved in the initial processing steps, whereas secondary exoribonucleolytic degradation is most likely due to the action of PNPase (3′–5′) or RNase J (5′–3′) (12,23,24).

In contrast to plastids, plant mitochondria have few polycistronic transcripts, and protein-coding genes are generally separated by large and poorly conserved intergenic sequences (25). Therefore, a complex interplay of transcription initiation and termination as well as correct processing of precursor RNAs is required to produce mature mRNA extremities in plant mitochondria. An extensive analysis of mitochondrial mRNA extremities in Arabidopsis indicates that most protein-encoding mRNAs present multiple 5′ ends, whereas almost all transcripts have a single and precise 3′ terminus (26). This analysis also shows that 5′ mRNA ends rarely originate from transcription initiation, being more often generated post-transcriptionally. No particular sequence motif near 5′ or 3′ mitochondrial mRNA termini has been found, although in a few cases, tRNA or tRNA-like structures (t-elements) correlate with mRNA extremities. Three PPR proteins belonging to the Rf-like subfamily and one P-type PPR protein (6) appear to be involved in producing the 5′ ends of some mitochondrial mRNAs (27–30). Surprisingly, the produced 5′ mRNA extremities rarely seem essential for the function of the corresponding mitochondrial mRNAs, as mutants affected in the corresponding *PPR* genes accumulate 5′-extended but equally stable mitochondrial transcripts in most cases analyzed. As there is probably no 5′-to-3′ exonuclease activity in plant mitochondria, the mechanism by which these Rf-like PPR proteins help to define 5′ mRNA ends is currently not understood. Instead, Rf-like PPR proteins may recruit an as-of-yet unknown endonuclease at the 5′ extremities of mature mRNAs. Evidence supporting this endonucleolytic model includes the detection of the upstream cleavage product (28,29). Although almost nothing is known about 3′ transcript formation in plant mitochondria, it may involve exonucleolytic trimming initiated at the 3′ extremity of precursor RNAs. PNPase and RNase R homolog 1 represent two 3′-to-5′ exoribonucleases likely involved in removing unnecessary sequence extensions in the 3′ region of mRNA (31). In chloroplasts, stem-loop structures located in 3′ untranslated regions (UTRs) are important in determining the 3′ ends and the stability of mRNAs (32,33). These structures are only occasionally found at the 3′ extremities of mitochondrial mRNAs (26,34), which suggest that RNA stabilization processes mediated by 3′ inverted repeats are not frequent in plant mitochondria.

Much remains to be discovered on 3′-end formation in plant mitochondria, and, critically, there is currently no evidence for protein-mediated 3′-end stabilization of mitochondrial mRNAs. Here, we examined the function of the mitochondria-targeted mitochondrial stability factor 1 (MTSF1) protein in *Arabidopsis thaliana*. We provide evidence that this PPR protein is required for correct 3′ processing and stability of the mitochondrial *nad4* mRNA. Furthermore, we demonstrate that the MTSF1-binding site matches the 3′ end of the mature *nad4* mRNA. This study thus reveals that the protein-based model of mRNA end processing and stabilization previously demonstrated in plastids is also conserved in higher plant mitochondria.

MATERIALS AND METHODS

Plant materials

Arabidopsis Col-0 plants were obtained from the *Arabidopsis thaliana* Resource Centre at INRA Versailles (http://dbsgap.versailles.inra.fr/portail/). The N654650 (*mtsf1-1*) and N586724 (*mtsf1-2*) Arabidopsis mutant lines were acquired from the European Arabidopsis Stock Centre (http://arabidopsis.info/). The *mtsf1-1* mutant plants were genotyped with the MTSF1-3 and the MTSF1-5 primers combined with the LB-SALK2 primer. The *mtsf1-2* mutant plants were genotyped with the MTSF1-7 and the MTSF1-8 primers combined with the LB-SALK2 primer.

RNA extraction and analysis

Total RNA was extracted from 8-week-old Arabidopsis leaves and flower buds using the TRIzol reagent (Invitrogen) according to the manufacturer's instructions. When needed, total RNA was treated with RNase-free DNase I (RNeasy Mini Kit, Qiagen) before further

analysis. For northern blot analyses, 15 μg of total RNA was loaded on a 1.5% agarose gel containing 3% formaldehyde in 1× MOPS buffer (20 mM MOPS at pH 7, 0.5 mM sodium acetate, 1 mM EDTA). Following electrophoresis for ~5 h at 100 V, RNA samples were transferred onto a nylon membrane (Genescreen) and covalently bound to membranes by a UV treatment (Stratalinker, Stratagene). Hybridizations with radiolabeled, gene-specific probes were performed overnight at 65°C in Church Buffer [7% SDS, 0.25 M Na_2HPO_4 (pH 7.4), 2 mM EDTA, 200 μg/ml heparin]. Blots were washed twice in 1× SSC, 0.1% SDS for 15 min, then in 0.1× SSC, 0.1% SDS for 15 min at 65°C before being exposed to a phosphorimager screen. Primer combinations used to PCR-amplify the corresponding DNA probes are shown in Supplementary Table S1.

For the circular reverse transcriptase-polymerase chain reaction (RT-PCR) analysis, 10 μg of total RNA was self-ligated for 2 h at 37°C with 40 U of T4 RNA ligase (Fermentas) according to the manufacturer's instructions. After phenol extraction and precipitation, first strand Complementary DNA (cDNA) synthesis was done for 3 h at 40°C using 400 U of M-MLV reverse transcriptase (Fermentas), 8 mM of random hexamers (Eurofins), 1× M-MLV buffer, 0.5 mM dNTPs and 40 U of Riboblock RNase inhibitor (Fermentas). The cDNAs were diluted four times, and 1/20th of the obtained cDNA solution was used in a single PCR amplification reaction. The Nad4-E5 and Nad4-CRT1 primers were used to map the *nad4* mRNA extremities. Amplified DNA products were then cloned in pTopoII plasmids (Invitrogen) and sequenced.

For the sRNA detection analysis, 30 μg of total Arabidopsis RNA were loaded onto a 15% polyacrylamide gel containing 7.5 M urea and 1× Tris-Borate-EDTA buffer (Sigma). Electrophoresis was carried out at 2 W 280 V for 45 min then at 600 V for 90 min. Gels were then stained with ethidium bromide and electrotransferred to a nylon membrane (HybondNX, Amersham). Blots were hybridized with the AS-3′ UTR-NAD4 antisense RNA synthetic oligonucleotide previously labeled at the 5′ end with [γ-^{32}P]ATP and T4 polynucleotide kinase (Fermentas). A synthetic sense RNA oligonucleotide (150 pg), mimicking the putative MTSF1 footprint was also loaded on the gel as a hybridization control and size marker. Hybridization was carried out at 50°C in a buffer containing 5× SSC, 20 mM Na_2HPO_4 (pH 7.2), 7% SDS and 2× Denhart's PerfectHyb Plus hybridization solution (Sigma). Blots were washed twice in a solution containing 3× SSC, 25 mM Na_2HPO_4 (pH 7.2), 5% SDS then once in a solution containing 1× SSC, 1% SDS before phosphoimaging.

Quantitative RT-PCR analysis measuring steady–state levels of mitochondrial pre-mRNA and mature mRNA were performed as in (35) with a slightly different set of primers (Supplementary Table S1). The reverse transcription step was performed on 3 μg of DNase-treated RNA (Ambion DNase, for 1 hr at 37°C), using Superscript III (Invitrogen) and random hexamer primers.

Primer pairs used for quantifying alternative NAD(P)H dehydrogenase (NDA and NDB) and alternative oxidase (AOX) mRNA levels are listed in Supplementary Table S1. For these genes, real-time PCR was then performed on a Mastercycler Ep Realplex (Eppendorf). Two technical and three biological repeats were analyzed. Expression levels were estimated by the comparative ΔΔCt method and normalized to EF1 expression (36).

Oligonucleotides

The DNA and RNA oligonucleotides used in this study are listed in Supplementary Table S1.

Antibody production and purification

A DNA insert encoding an MTSF1 protein fragment, from amino acid residue 445–847, was PCR amplified using the MTSF1-GW5 and MTSF1-GW7 primers. The AttB1 and AttB2 recombination sites were then completed in a second PCR amplification using the GW5 and GW3 oligonucleotides and recombined by Gateway™ LR reaction in the pDEST-17 expression vector (Invitrogen), yielding a 46 kDa polypeptide bearing an N-terminal poly-histidine tag. The recombinant plasmid was transformed in *Escherichia coli* BL21 (DE3) pLysS cells (Novagen), and protein expression was induced by addition of 1 mM IPTG for 3 h at 37°C. The recombinant MTSF1 fragment was solubilized from inclusion bodies in 50 mM Tris–HCl at pH 7, 150 mM NaCl and 10% N-lauryl sarcosine. The fusion protein was further purified on a nickel-Sepharose column and then dialyzed against a buffer containing 50 mM Tris–HCl at pH 7, 150 mM NaCl and 0.1% N-lauryl sarcosine. The purified protein was injected into rabbits to produce polyclonal antisera (Biogene) that were affinity purified against the antigen used for injection.

Protein extraction and analysis

Arabidopsis mitochondria were isolated from flower buds, and chloroplasts were purified from young leaves as described in (37). Total and organellar proteins were extracted in 30 mM HEPES-KOH (pH 7.7), 10 mM Mg-acetate, 150 mM K-acetate, 10% glycerol, 0.5% CHAPS. Protein concentration was measured using the Bradford colorimetric method (Bio-Rad). Approximately 50 μg of proteins were fractionated by SDS–PAGE, electrotransferred to PVDF membrane (Perkin-Elmer) and hybridized with polyclonal antibodies recognizing either MTSF1 (diluted 1:100), NAD9 [diluted 1:100000; (38)], FDH [diluted 1:2000; (39)] or the ATPC protein (diluted 1:5000; a gift of Jorg Meurer, Ludwig-Maximilians-Universität, Germany), or with monoclonal antibodies recognizing the PORIN protein (diluted 1:500; a gift of David Day, University of Western Australia, Australia). Detection was carried out using peroxidase-conjugated secondary antibodies and enhanced chemiluminescence reagents (Amersham Pharmacia Biotech). The apparent molecular masses of the proteins were estimated with pre-stained molecular mass markers (Fermentas).

At least two individual experiments were performed for each analysis. Analysis of the sub-mitochondrial distribution of MTSF1 protein was performed as described in (37).

Blue Native gel and activity stain

Mitochondrial pellets representing ~500 µg of total mitochondrial proteins were solubilized in a solution containing 50 mM bis-Tris (pH 7), 750 mM 6-aminohexanoic acid, 0.5 mM EDTA and 5 g of digitonin (Sigma) per gram of protein. About 200 µg of solubilized proteins were loaded in each lane of 4–13% acrylamide gradient gels and ran overnight at 100 V followed by 5 h at 200 V. After migration, the gels were stained with Coomassie blue G-250 solution for visualization of protein complexes. The NADH dehydrogenase activity of complex I was revealed in gel by incubation in the presence of 0.2% nitroblue tetrazolium, 0.2 mM of NADH and 0.1 M Tris–HCl (pH 7.4). When sufficient signal was obtained, the staining reaction was stopped by soaking the gel in fixing solution (30% methanol and 10% acetic acid).

Expression of recombinant MTSF1 and RNA-binding assays

The MTSF1 protein deprived of its first 146 amino acids was expressed as a His-tag fusion protein in *E.coli* Rosetta 2 cells (Novagen). The corresponding DNA fragment was first PCR amplified using the MTSF1-GW3 and MTSF1-GW4 primers. The AttB1 and AttB2 recombination sites were then completed in a second PCR amplification using the GW5 and GW3 oligonucleotides. The obtained amplification product was cloned by Gateway^{TM} BP reaction into pDONR207 (Invitrogen) and subsequently transferred into the pDEST17 expression vector by LR reaction (Invitrogen). Following expression at 37°C, the 6xHis-MTSF1 protein appeared to be highly insoluble and was solubilized from *E.coli* inclusion bodies in 150 mM NaCl, 50 mM Tris (pH 7.5) and 10% N-lauryl sarcosine, 0.5 µg/ml leupeptin and 0.06 µg/ml pepstatin. The obtained protein solution was subsequently dialyzed overnight against the same buffer in which the N-lauryl sarcosine concentration was changed to 0.01%. The protein was then enriched to high purity using a Ni-Sepharose column (GE Healthcare) and stored at 4°C before use. Gel mobility shift (GMS) assays used gel-purified synthetic RNA derived from the 3′ UTR and the coding sequence of the *nad4* mRNA (Eurofins MWG Operon). Purified RNA oligonucleotides were labeled at the 5′-end with T4 polynucleotide kinase (Fermentas) and [γ-^{32}P]ATP according to manufacturer's instructions. Before each GMS assay, 100 pM of purified oligonucleotides were heated to 95°C for 1 min and snap-cooled on ice. The RNA-binding reactions contained 40 mM Tris–HCl (pH 7.5), 100 mM NaCl, 0.1 mg/ml BSA, 10% glycerol, 1 mg/ml heparin, 10 U Riboblock (Fermentas), 100 pM radiolabeled RNA and protein concentrations indicated in figure legends. The reactions were carried out 30 min at room temperature, and samples were deposited on a 5% polyacrylamide (29:1 acrylamide/bis-acrylamide) running gel prepared in 1× THE [66 mM HEPES, 34 mM Tris (pH 7.7), 0.1 mM EDTA]. The gels were run at 10 W (constant power), dried and phosphoimaged (FLA-5000 Fujifilm).

Gas exchange measurements

Photosynthetic and respiratory CO_2 gas exchange was measured with a portable open system Li-Cor Li-6400 xt (Li-Cor) using the 2 cm^2 leaf chamber. Humidity was strictly controlled at 80% relative humidity using a dew-point generator Li-610 connected to the inlet of the Li-6400. Measurements were carried out on 10-week-old (wild-type) or 15-week-old (*mtsf1*) plants grown on soil under long days (16/8 h day/night). Net CO_2 assimilation was first measured under standard conditions (400 ppm of CO_2, 300 µmol/m^2/s of Photosynthetically Active Radiation (PAR), 21% O_2, 23°C), and then under high light (800 µmol/m^2/s PAR, at 400 ppm CO_2) or high CO_2 (1000 ppm CO_2, at 300 µmol/m^2/s PAR). Photosynthetic measurements are expressed relative to leaf surface area. Three biological replicates were analyzed per genotype. Respiratory O_2 consumption was measured using a closed system made of the gas-phase oxygen electrode (Hansatech Electrode), in which the decrease in the O_2 mole fraction monitored. Owing to the rather low sensitivity of the electrode, several leaves were placed in the cuvette, and thus respiratory O_2 uptake is expressed relative to dry mass rather than surface area.

Synchrotron radiation circular dichroism

An MTSF1 protein solution at 2.3 mg/ml was measured in 100 mM Tris–HCl buffer (pH 9.0), 150 mM Na_2SO_4 and 0.05% (w/v) N-lauryl-sarcosine, on DISCO beamline at Soleil synchrotron (Gif-sur-Yvette, France). Four measurements, from 170 to 280 nm, with 1 nm intervals per second, were performed. For data processing, spectra of dialysis buffer were subtracted from those of corresponding samples. Spectra were analyzed down to 175 nm: this corresponds to a photomultiplier high tension remaining below half of its total variation. The 260–270 nm region was set to zero, and the resulting spectrum was calibrated with D-10-camphorsulfonic acid using the CDtool software (40). Secondary structure determination was performed using ContinLL program in Dichroweb (41). The reference set used was SP175 (42). Normalized root-mean-square deviation (NRMSD) gave insight into the most accurate fit for the spectrum. ContinLL analysis indicated 42% α-helix, 16% β-strand, 15% turn and 27% unordered with an NRMSD of 0.1.

RESULTS

Arabidopsis *mtsf1* mutants are affected in a mitochondria-targeted P-type PPR protein

The *mtsf1-1* mutant was originally identified in a series of *Arabidopis thaliana* T-DNA mutants having insertions in nuclear genes encoding putative mitochondria-targeted PPR proteins. The *mtsf1-1* mutant plants were classified among slow-growing mutants that can be cultivated

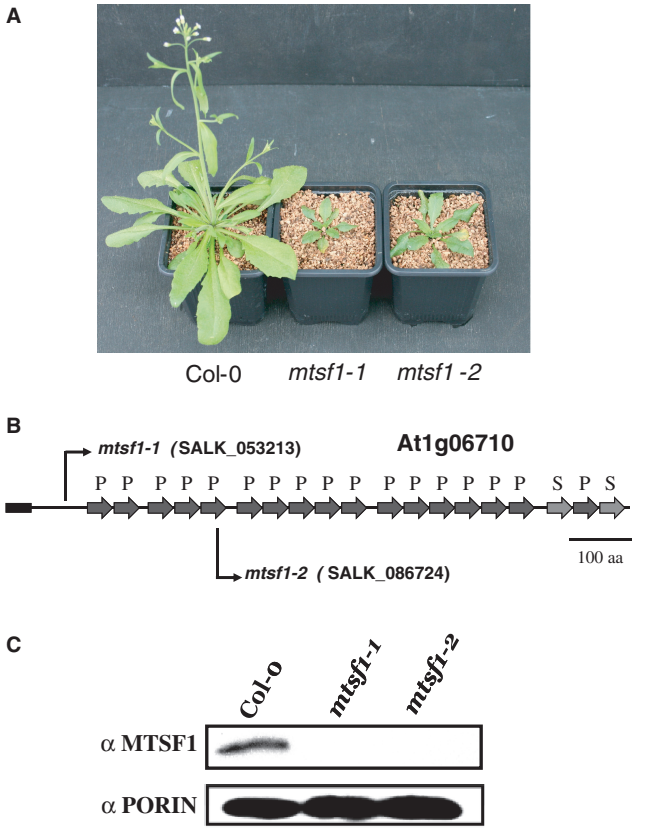

Figure 1. The Arabidopsis *mtsf1* mutants grow slowly on soil, but are fertile. (**A**) Photograph taken 65 days after germination showing the aerial growth of the *mtsf1-1* and *mtsf1-2* mutants compared with the wild-type Col-0 plant. (**B**) Diagram of the MTSF1 protein showing the predicted 17 P and the 2 S PPR repeats and the location of the corresponding T-DNA insertion sites present in the *mtsf1-1* and *mtsf1-2* mutants. (**C**) Immunodetection of the MTSF1 and PORIN proteins in mitochondrial protein extracts prepared from wild-type (Col-0) and *mtsf1-1* and *mtsf1-2* mutant plants. Fifty micrograms of total mitochondrial proteins were deposited on a denaturing gel, transferred to PVDF membrane and labeled with antibodies directed against the indicated proteins.

relatively easily on soil under conventional greenhouse conditions (Figure 1A). The implicated PPR gene corresponds to the AT1G06710 gene and encodes a 111 kDa protein comprising 19 PPR repeats [17 P-type and 2 S-type, (6)] according to motif predictions available on the FlagDB++ database (Figure 1B). A second *mtsf1* mutant allele was identified and *mtsf1-2* homozygous plants showed identical phenotypes to those observed in the *mtsf1-1* mutant line, confirming that the growth alterations were due to the inactivation of gene AT1G06710 (Figure 1A and B).

After 3 months of growth on soil in long-day conditions, *mtsf1* mutants reached about half the size of wild-type plants and had curly rosette leaves (Supplementary Figure S1A). The plants were also late flowering, but not sterile (Supplementary Figures S1B and S2). Seeds derived from *mtsf1* mutants were dark brown, flattened and most contained embryos showing incomplete development (Supplementary Figure S3A and B). Nevertheless, ~20% of the seeds produced by *mtsf1* mutants germinated *in vitro* and yielded plantlets that could be further cultivated on soil (Supplementary Figure S1C).

A polyclonal antibody raised against the MTSF1 protein was used to hybridize mitochondrial proteins extracted from wild-type plants and *mtsf1* mutants. A specific signal at the expected size (~100 kDa) was observed in wild-type, but not in mutant protein extracts, strongly supporting that both identified mutant lines represent null mutants (Figure 1C).

MTSF1 encodes a mitochondria-targeted PPR protein

A protein fusion comprising the first 42 amino acids of the MTSF1 protein and the green fluorescent protein (GFP) was transiently expressed in tobacco cells (*Nicotiana benthamiana*) along with a mitochondria-targeted mCherry fusion protein as a control (43). Transformed tobacco cells were observed using confocal microscopy, and the GFP fluorescence appeared as a punctuated signal within the cells that overlapped with the mitochondria-localized mCherry red signal (Figure 2A). The cellular distribution of MTSF1 was further analyzed by probing total, chloroplastic and mitochondrial protein preparations with the immunopurified anti-MTSF1 antibody (Figure 2B). The MTSF1 protein was strongly enriched in the mitochondrial protein preparation, as attested with the anti-NAD9 and anti-ATPC control antibodies. The sub-mitochondrial distribution of MTSF1 was further determined by analyzing soluble and membrane-bound mitochondrial proteins for the presence of MTSF1. MTSF1 was mostly detected in the membrane protein fraction but also in soluble proteins, indicating that MTSF1 may exist both as a soluble and membrane-bound protein or that it is loosely attached to mitochondrial membranes (Figure 2C). Mitochondrial integral and peripheral membrane proteins were further separated by sodium carbonate treatment, which has been shown to solubilize extrinsic membrane proteins (44). Most of the MTSF1 signal was found in the extrinsic membrane proteins, supporting a peripheral association of MTSF1 with mitochondrial membranes (Figure 2D).

The *mtsf1-1* and *mtsf1-2* mutants show altered respiratory and photosynthetic activities

The respiratory and photosynthetic activities of *mtsf1* mutants were determined by gas exchange measurements. Under standard conditions (400 ppm CO_2, 300 µmol photon.m^{-2}.s^{-1}), both mutants had lower net CO_2 assimilation compared with wild-type plants (Figure 3A). At saturating light (800 µmol photon.m^{-2}.s^{-1}), net CO_2 assimilation significantly increased in both *mtsf1* mutant plants but did not reach values measured in control plants. At a saturating CO_2 level (1000 ppm CO_2), net CO_2 assimilation increased and reached the value of wild-type plants under the same conditions. This type of photosynthetic response to CO_2 indicates that mutants had a photosynthetic capacity similar to the wild-type, but had higher photorespiratory CO_2 release (lower CO_2 levels at carboxylation sites) at ordinary CO_2 levels.

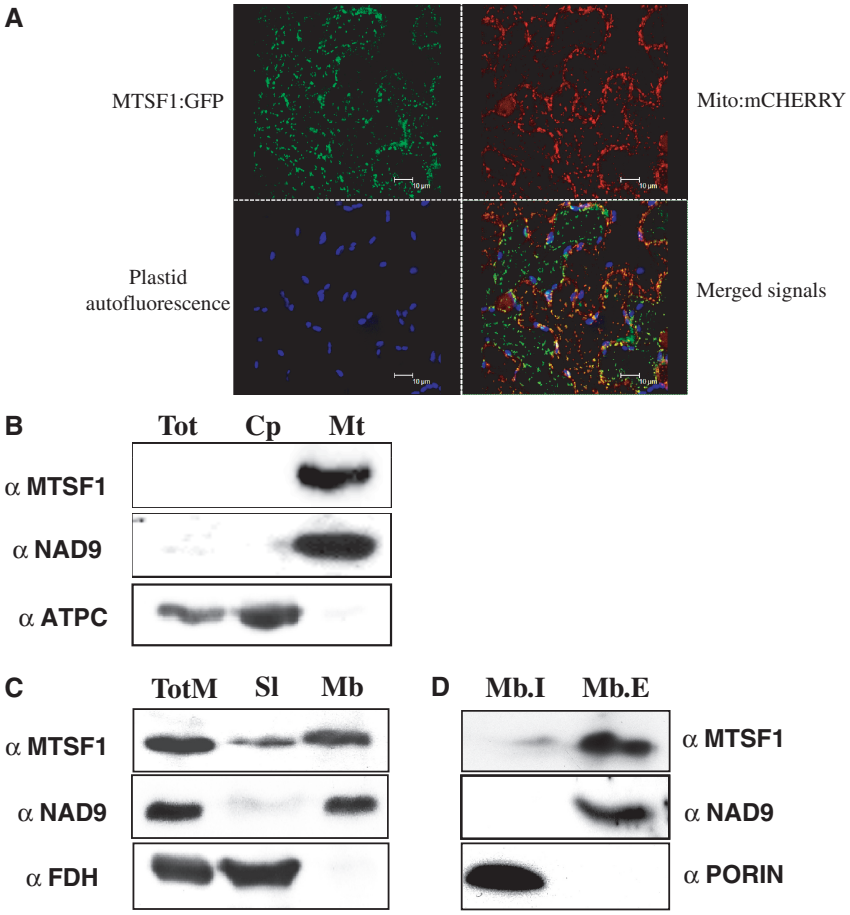

Figure 2. Sub-cellular and sub-mitochondrial localization of MTSF1. (**A**) Confocal images showing the cellular distribution of an MTSF1-GFP fusion protein (green) and a mitochondria-targeted mCHERRY marker (red) in *Nicotiana benthamiana* mesophyll cells transiently transformed by agro-infiltration. The location of chloroplasts was revealed by chlorophyll autofluorescence (blue). The good overlap of both GFP and mCHERRY fluorescence supports a mitochondrial localization of the MTSF1-GFP fusion (shown in the bottom right-hand image). (**B**) Probing of total (Tot), chloroplastic (Cp) and mitochondrial (Mt) protein preparations with the anti-MTSF1 antibody. Antibodies against marker proteins NAD9 and ATPC were used to check the purification of mitochondria and chloroplasts, respectively. (**C**) Immunoblot analysis of total (TotM), soluble (Sl) and membrane (Mb) mitochondrial protein fractions. Protein gel blots were probed with anti-MTSF1 antibodies (αMTSF1). Replicate blots were also probed with antibodies directed against formate dehydrogenase (αFDH, a known soluble matrix protein) and NAD9 (αNAD9, an extrinsic protein of the inner membrane of mitochondria). Each lane was loaded with ~50 µg of the protein preparation. (**D**) Analysis of protein content in sodium carbonate-treated mitochondrial membranes. Purified mitochondrial membranes were treated with Na_2CO_3 to release and solubilize peripheral membrane proteins Mb.I corresponds to membrane proteins that could not be solubilized by sodium carbonate, such as integral membrane proteins. Mb.E corresponds to proteins that could be extracted by sodium carbonate, such as extrinsic membrane proteins. Approximately 50 µg of each protein preparation was loaded onto each lane, and blots were probed with the indicated antibodies. The αPORIN antibody detects an intrinsic protein located in the outer membrane of mitochondria.

All plant lines had a similar stomatal conductance (data not shown); therefore, our results suggest that the internal mesophyll CO_2 conductance was lower in mutants.

Respiratory activity was measured as CO_2 production and oxygen uptake in the dark (Figure 3A and B). In both cases, measurements revealed respiratory activity that was two to three times higher in *mtsf1* mutants compared with wild-type plants. To better understand the origin of this increase, the expression of alternative components of the respiratory chain was analyzed. Quantitative RT-PCR was conducted to determine the steady–state levels of mRNAs encoding alternative mitochondrial NAD/NADH dehydrogenases (NDA and NDB) and oxidases (AOX) in mutant and wild-type plants. Most of these genes were overexpressed between 2- to 4-fold in the mutants, except for the *NDB4* gene, whose mRNA abundance was 60 times higher compared with wild-type plants (Figure 4A). The overexpression of AOX was further estimated by probing total mitochondrial protein extracts with an antibody to AOX. A highly significant increase in the AOX signal was detected in both mutants, confirming that leaf respiration relied heavily on the alternative pathway in *mtsf1* plants (Figure 4B).

mtsf1 plants are complex I respiratory mutants

To further clarify the respiratory perturbations present in *mtsf1* mutants, we analyzed the steady–state levels of the different respiratory complexes in comparison with the wild-type. Mitochondria prepared from mutant and

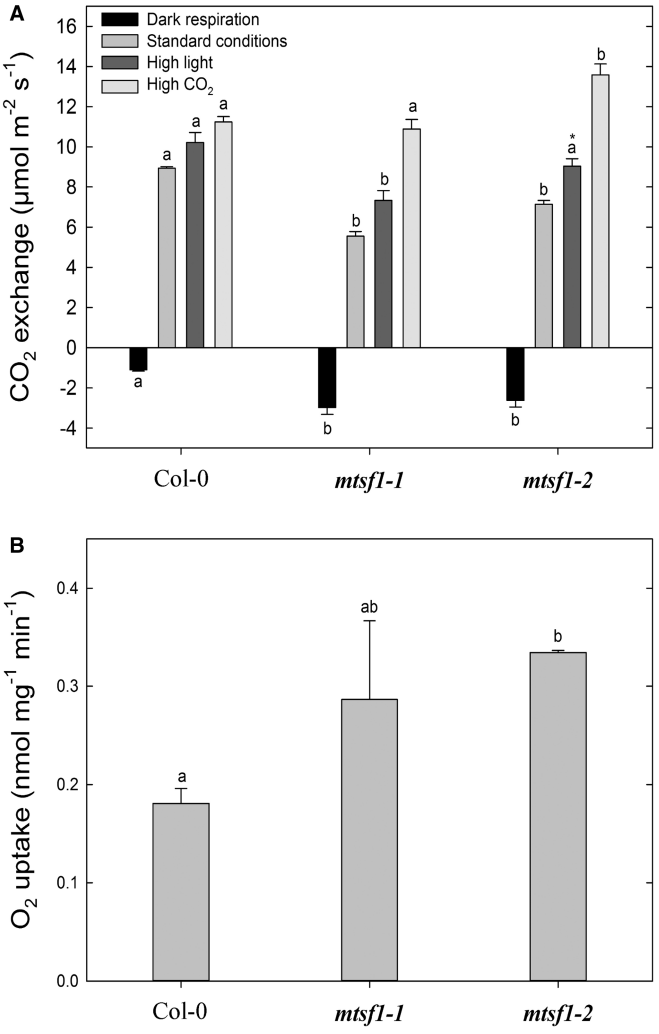

Figure 3. Analysis of photosynthetic and respiratory activities in *mtsf1* mutants. (**A**) Net CO_2 assimilation in the light and CO_2 flux in the dark as measured in Arabidopsis rosette leaves (still attached). Means ± SE ($n = 3$). (**B**) Oxygen consumption by dark respiration of rosette leaf fragments. Means ± SE ($n = 3$). Statistically significant differences are indicated by different letters (a, b: $P < 0.01$, a*: $P < 0.07$).

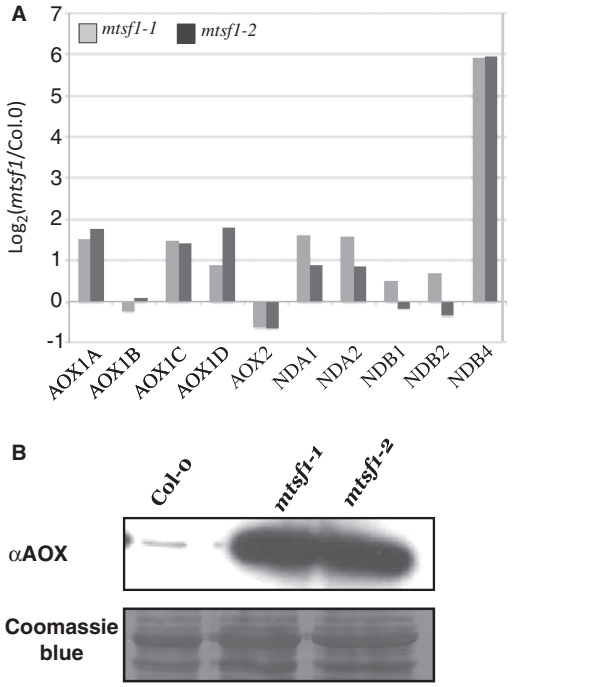

Figure 4. Genes associated with alternative electron transport in the mitochondrion are overexpressed in *mtsf1* mutants. (**A**) Quantitative RT-PCR analysis measuring the relative steady–state level of alternative NADH dehydrogenase (NDA and NDB) and alternative oxidase (AOX) mRNAs in *mtsf1* mutants. (**B**) Immunoblot analysis on total mitochondrial protein prepared from wild-type (Col-0) and *mtsf1* plants and probed with an antibody to AOX. Approximately 50 µg of protein samples were loaded in each lane.

wild-type plants were lysed in the presence of digitonin, and respiratory complexes were separated by blue-native polyacrylamide gel electrophoresis (BN-PAGE). This approach revealed that the complex I and the super-complex I + III were missing from *mtsf1* mutants (Figure 5A). No other major differences were observed in BN-PAGE of mutant and control plants. An in-gel activity assay confirmed the lack of full-length complex I in both mutants and revealed two lower bands of smaller sizes compared with wild-type mitochondria (Figure 5B). These two lower bands were likely caused by the accumulation of a truncated version of complex I that conserved its ability to associate with complex III in mitochondrial membranes. These truncated versions of complex I and I + III were not detected on BN-PAGE gels stained with Coomassie blue, indicating that they accumulated only at low levels.

Stability of *nad4* mRNA and splicing of *nad2* mRNA are compromised in *mtsf1* mutants

Owing to the role of PPR proteins in the processing of organellar RNAs, we hypothesized that the processing of one or several mitochondrial mRNAs encoding complex I subunit was defective in *mtsf1* mutants. RNA gel blot analysis was first carried out to compare the accumulation levels of mitochondrial mRNAs encoding complex I subunits in mutant and wild-type plants. The major difference involved the *nad4* and *nad2* genes. No mature or precursor *nad4* mRNA was detected at all in the *mtsf1* mutants, suggesting strong destabilization of *nad4* mRNA (Figure 6). The mature *nad2* mRNA accumulated at a lower level compared with control plants and the concomitant increase in abundance of a precursor mRNA indicated that the splicing of *nad2* mRNA was partly affected in *mtsf1* plants (Figure 6). No other significant differences were observed for any of the other *nad* mRNAs, other than a tendency to over accumulate in both mutants (Supplementary Figure S4). Quantitative RT-PCR analysis then confirmed that mature *nad4* was almost undetectable in *mtsf1* mutants (Supplementary Figure S5A) and showed that *nad2* splicing defects involved mostly the first intron and, to a lesser extent, the second intron (Supplementary Figure S5B). Consistently, a concomitant over accumulation of *nad2* mRNA precursors with unspliced introns was detected

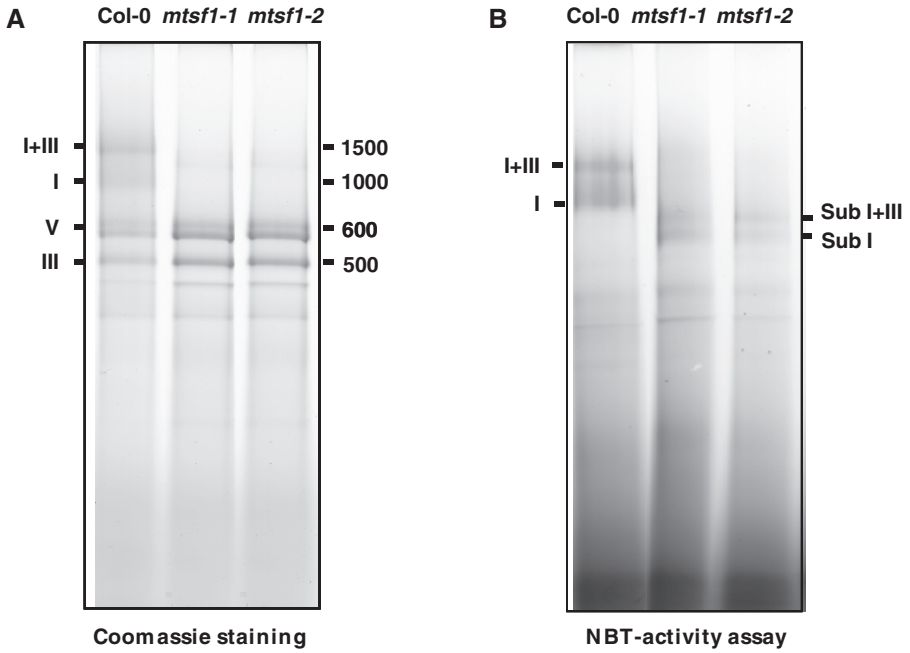

Figure 5. *mtsf1* mutants accumulate a truncated form of respiratory complex I. (**A**) Blue native PAGE gel analysis of the different respiratory complexes present in wild-type (Col-0) and *mtsf1* mutant plants. About 200 µg of digitonin-treated mitochondrial proteins were loaded in each lane. After electrophoresis, the gel was stained with Coomassie blue to reveal the different respiratory complexes, which are identified by their roman numeral in the figure. Numbers on the right-hand side indicate positions of bands corresponding to molecular weight marker (in kDa) (**B**) In-gel staining revealing complex I activity in both wild-type (Col-0) and *mtsf1* mutant plants. Gel running conditions were identical to that presented in panel A Sub I corresponds to the truncated form of complex I detected in *mtsf1* mutants.

(Figure 6). For *nad4*, no significant increase of any precursor mRNA species in the mutants was observed (Figure 6 and Supplementary Figure S5), indicating that there was no splicing defect for *nad4*. Altogether, these results show that the major effect of the *mtsf1* mutations is a strong destabilization of mitochondrial *nad4* mRNA.

3′-truncated *nad4* mRNAs accumulate in *mtsf1* mutants

To better understand the molecular events leading to the destabilization of the mitochondrial *nad4* mRNA in *mtsf1* plants, its 5′ and 3′ extremities were mapped by circular RT-PCR. This PCR-based approach was chosen because it appeared to be appropriate for mapping the extremities of the *nad4* mRNA found only in low abundance in *mtsf1* mutants. On agarose gels, major amplification products of different sizes were detected in wild-type and *mtsf1-1* plants (Figure 7A). These DNA fragments were subcloned and sequenced. As expected, the major band obtained from wild-type plants corresponded to the previously determined 3′ end of the stable *nad4* mRNA, which is located 30 nucleotides downstream of the stop codon (Figure 7B). However, we observed that the shorter band derived from *mtsf1-1* plants corresponded to 3′-truncated *nad4* mRNA molecules, and that the most frequently identified 3′ extremity of these transcripts was located 13–15 nucleotides upstream of the UGA stop codon (37 of 53 clones analyzed, Figure 7B). In contrast, the 5′ end of these shorter mRNA molecules was unchanged and identical to the wild-type. This analysis revealed that the rare *nad4* mRNA molecules present in *mtsf1* plants bear truncated 3′ extremities in most cases. These truncated transcripts likely correspond to degradation intermediates whose destabilization is initiated from their 3′-end in a 3′-to-5′ direction.

The MTSF1 protein binds specifically to the 3′ end of *nad4* mRNA

We then investigated the possibility that the MTSF1 protein associates with the 3′ end of the mitochondrial *nad4* mRNA to stabilize it *in vivo*. We used a soluble and active source of MTSF1. The MTSF1 protein deprived of its first 146 amino acids was fused to a 6xHis N-terminal extension and expressed in *E.coli*. This slightly truncated version of MTSF1 did not contain the poorly conserved mitochondrial targeting region but contained all PPR repeats present in the wild-type protein. The MTSF1 protein was solubilized from *E.coli* inclusion bodies in the presence of a high concentration of N-lauryl sarcosine. The concentration of detergent was decreased to 0.1% (w/v) by dialysis, and, finally, the MTSF1 protein was recovered at high purity by metal chelate chromatography (Figure 8A). The correct fold of the purified MTSF1 protein was attested by synchrotron radiation circular dichroism analysis, which indicated that the protein was mainly α-helical (Figure 8B). Different synthetic RNA oligonucleotides mapping to the 3′-coding and non-coding region of the *nad4* mRNA were then tested in a GMS assay for their ability to associate with the MTSF1 protein. MTSF1 was able to associate specifically with a 30 nucleotide RNA covering the entire *nad4* 3′ UTR (Figure 8C and D). A shorter version comprising only the last 20 nucleotides of the 3′ UTR was recognized by

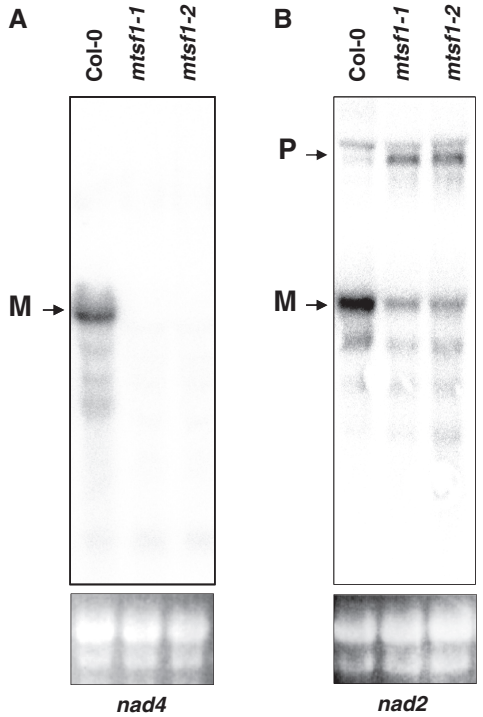

Figure 6. MTSF1 is required to produce stable *nad4* transcript and mature *nad2* mRNA. (**A**) RNA gel blot analysis showing the steady–state accumulation of *nad4* mRNA species in wild-type (Col-0) and *mtsf1* mutant plants; 15 µg of total RNAs were loaded in each lane. The probe used corresponded to *nad4* cDNA. (**B**) RNA gel blot analysis showing the steady–state accumulation of *nad2* mRNA species in wild-type (Col-0) and *mtsf1* mutant plants; 15 µg of total RNAs were loaded in each lane. The probe used corresponded to *nad2* cDNA. M, mature mRNA; P, precursor mRNA.

MTSF1 with similar affinity. In contrast, MTSF1 showed no binding to synthetic RNA oligonucleotides corresponding either to the last 20 nucleotides of the *nad4* mRNA coding sequence or to the first 30 nucleotides located just downstream of *nad4* mRNA 3′ extremity (Figure 8C and D). These results strongly suggest that the last 20 nucleotides of the *nad4* 3′ UTR represents the MTSF1-binding site on the *nad4* mRNA *in vivo*.

A short RNA representing the binding site of MTSF1 on *nad4* mRNA accumulates *in vivo* in an MTSF1-dependent manner

Several abundant short RNA sequences corresponding to binding sites of different RNA-stabilizing proteins (including PPR proteins) accumulate in relatively high levels in plastids (21,45). These short segments of RNAs are protected from degradation by a tight association with a corresponding PPR protein involved in mRNA stability. These PPR footprints can be found in the large RNA sequence libraries produced by the high-throughput small RNA sequencing programs. To check whether an RNA footprint corresponding to MTSF1 could be detected *in vivo*, we looked for the presence of short RNAs homologous to the 3′ UTR of *nad4* in several of these small RNA sequence collections. A short RNA corresponding to the last 20 nucleotides of *nad4* mRNA was identified, and its presence *in vivo* was tested by RNA gel blot analysis in wild-type and mutant plants. An antisense RNA probe to this short RNA sequence revealed a specific hybridization signal at the expected size (Figure 9). In addition, the observed signal required the presence of an active copy of the *MTSF1* gene, strongly suggesting that the mRNA was protected from degradation through the binding to the MTSF1 protein. These observations indicate that the short RNA of 20 nucleotides likely represents the *nad4* RNA segment with which the MTSF1 protein associates *in vivo*, corroborating the GMS assay results.

DISCUSSION

In this study, we examined the function of a new mitochondria-targeted PPR protein in Arabidopsis and showed that it is essential for the stabilization and the 3′-end processing of mitochondrial *nad4* mRNA. Several lines of evidence strongly support these interconnected roles of MTSF1. In an *mtsf1* mutant background, *nad4* mRNA appeared to be highly destabilized and was almost undetectable (Figure 6 and Supplementary Figure S5). The few *nad4* mRNA molecules that accumulated in the absence of MTSF1 bore 3′ extremities that were truncated upstream of the wild-type 3′ extremity, suggesting degradation via the action of a 3′-to-5′ exoribonuclease (Figure 7). Additionally, we demonstrated that the MTSF1 protein binds with high affinity to the last 20 nucleotides of *nad4* mRNA *in vitro*, and that a short RNA corresponding to this region accumulates *in vivo* in an MTSF1-dependent manner (Figures 8 and 9). These results provide strong evidence that the 3′ extremity of mature *nad4* mRNA corresponds to the MTSF1-binding site *in vivo* and that MTSF1 stabilization and processing activity requires an association of the protein with this region of *nad4* mRNA. The perfect sequence conservation of the MTSF1-binding sites in dicot species strongly suggests that the protective function of MTSF1 on *nad4* mRNA is conserved (Figure 10). Interestingly, a 7 nucleotide insertion was found at the extremity of *nad4* 3′ UTR in monocots compared with the sequence found in dicots (Figure 10). This may indicate that potential MTSF1 orthologs in monocots have evolved to recognize a slightly different *nad4* mRNA 3′-end sequence.

The functions identified for MTSF1 are reminiscent of the way other helical repeat proteins stabilize and process plastid RNAs (7,13,22), and MTSF1 likely acts in a similar way. We believe that the definition of the *nad4* mRNA 3′ end and its protection involve an early association between MTSF1 and the 3′ UTR of precursor *nad4* mRNA. The fact that MTSF1 binds to a region located upstream and not downstream of the end of the *nad4* 3′ UTR strongly suggests that MTSF1 is not involved in recruitment of an endoribonuclease, but in preventing exoribonuclease progression. Unnecessary *nad4* mRNA 3′ extensions are likely removed by the action of 3′-to-5′ exoribonucleases until they reach the MTSF1 protein, which acts as a physical barrier impeding further degradation of the mRNA. We also showed that *mtsf1* mutants

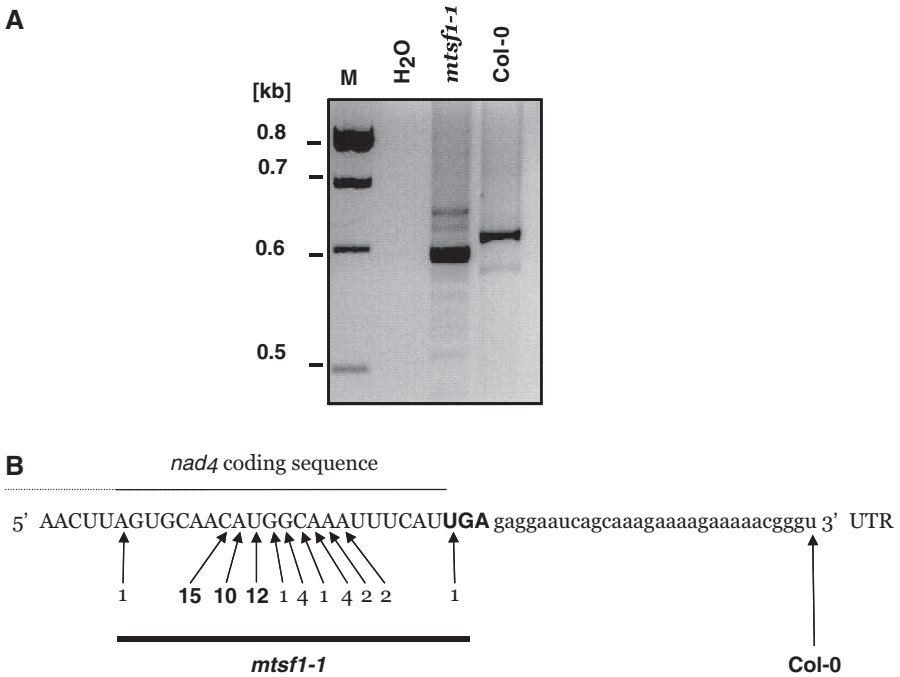

Figure 7. 3′ truncated *nad4* mRNAs accumulate in *mtsf1* plants. (**A**) Agarose gel showing circular RT-PCR amplification products for the *nad4* gene in the wild-type (Col-0) and *mstf1-1* mutant. M: DNA size marker, H_2O: negative control. (**B**) Partial RNA sequence showing the last 60 nucleotides of *nad4* mRNA. The region corresponding to the coding sequence is shown in capital letters. The rest corresponds to the 3′ UTR. The UGA stop codon is shown in bold. The arrows indicate the positions of the *nad4* mRNA 3′ ends found in the wild-type (Col-0) and *mtsf1-1* plants. Numbers on arrows represent how often a particular end was found after sub-cloning and sequencing the major band shown in panel A.

are moderately deficient in the splicing of the first and second introns of *nad2*. A minor role of MTSF1 in *nad2* splicing is therefore also plausible. Nevertheless, no obvious sequence homology was detected between *nad2* introns 1 and 2 and *nad4* 3′ UTR, indicating that MTSF1 is unlikely to bind these introns *in vivo*. Additionally, *nad2* splicing has been shown to be affected in several unrelated complex I mutants (46,47). Therefore, the effect of the *mtsf1* mutations on *nad2* splicing could simply be a pleiotropic effect of the loss of complex I in Arabidopsis.

Our results clearly demonstrate that the simultaneous mechanisms of RNA stabilization and transcript-end processing owing to the presence of protective proteins on mature RNA ends also occur in plant mitochondria. As in plastids, mitochondria-targeted PPR proteins have evolved to associate specifically with mitochondrial RNA extremities and protect them against exonucleases involved in bulk RNA decay. This conserved mechanism probably accounts for the stabilization of most mitochondrial mRNAs, which are generally devoid of protective stable stem-loop structures (26). Although our data suggest that the basic mechanism of RNA stabilization and transcript-end processing involving PPR proteins is conserved among plastids and mitochondria in plants, some important differences remain. In contrast to plastids, the processing and protective activity of the MTSF1 protein is not associated with intercistronic processing. This difference is not surprising because mitochondrial genes are not transcribed as long polycistronic precursors (25), and thus no intercistronic cleavage is required to release monocistronic mRNA. Additionally, this protein-based protective mechanism also accounts for the formation of 5′ mRNA ends in plastids (7,13,22). In mitochondria, a different mechanism of 5′-end cleavage involving an as-yet-unknown endoribonuclease activity and a peculiar subgroup of rapidly evolving PPR proteins as well as classical PPR-P protein has been proposed (28–30). However, in these cases, the PPR proteins involved do not correspond to stability factors *per se*—their inactivation does not lead to the disappearance of their target mitochondrial mRNA—but rather to an increase in the abundance of larger RNA precursors. It therefore appears that RNA stability and 5′-end processing correspond to separate processes in mitochondria, possibly involving different molecular players.

As indicated earlier in the text, our study reveals that short RNA molecules corresponding to a PPR-binding site can be detected in plant mitochondria, as previously shown in plastids (21,45). This result demonstrates that PPR protein-binding footprints, at least for the ones involved in mRNA stabilization, accumulate in plant mitochondria. The identification of short RNAs matching transcript 3′ ends may therefore reveal mitochondrial mRNAs that are potentially stabilized by an association with a PPR protein (or other types of RNA-binding protein) at their 3′ extremity. Nevertheless, as no 5′-to-3′ exoribonuclease seems to be active in plant mitochondria, this discovery raises the question of how the 5′ extremity of these short RNAs is formed *in vivo*. It is currently difficult to answer this question, and further clarifications are needed to determine whether

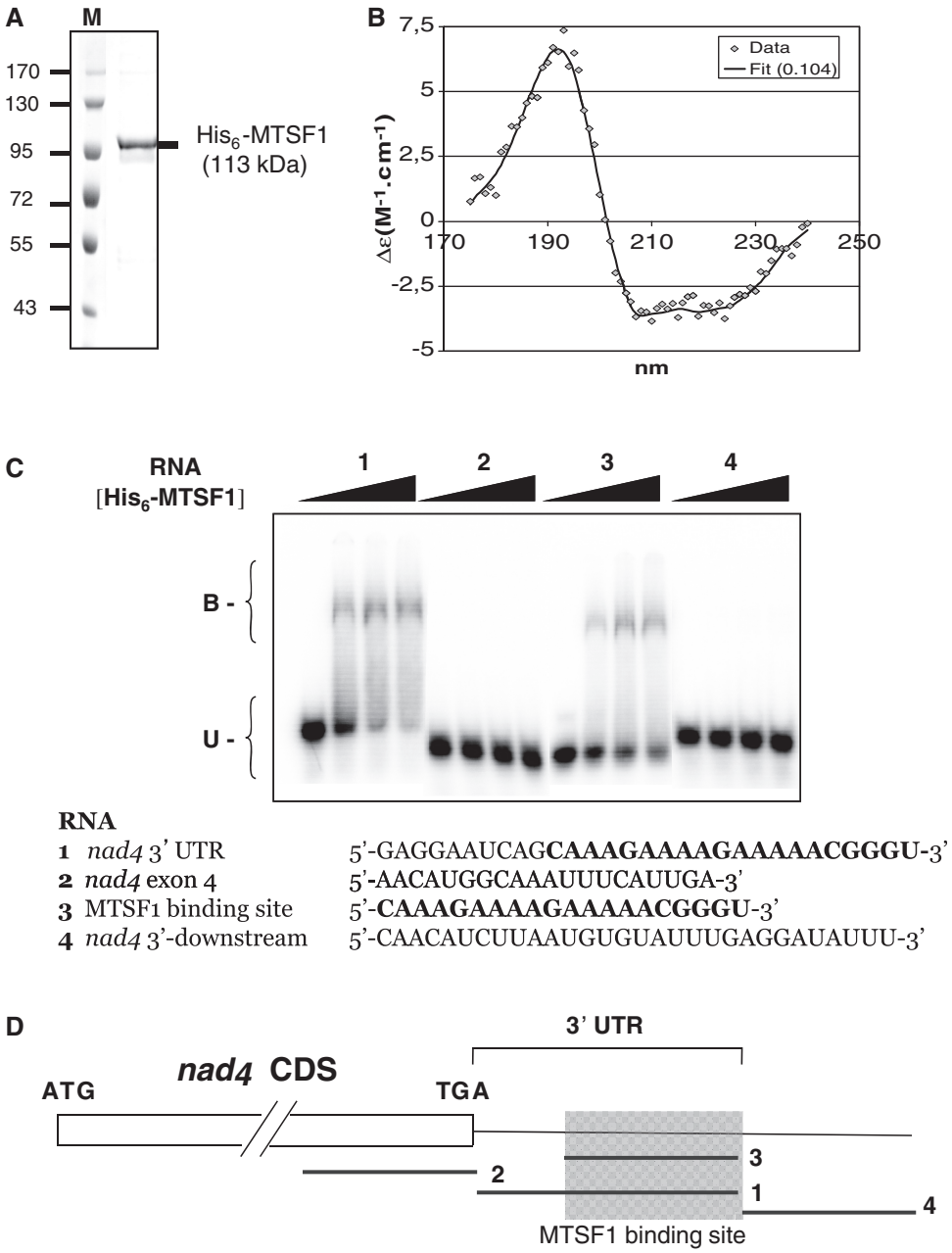

Figure 8. The MTSF1 protein binds specifically to the last 20 nucleotides of the *nad4* mRNA *in vitro*. (**A**) Protein gel stained with Coomassie blue showing the purity of the His_6-MTSF1 protein overexpressed and purified from *E.coli*; 15 microgram of the purified protein was loaded on the gel. M, Protein size marker. Sizes are indicated in kDa. (**B**) Biophysical analysis of recombinant MTSF1. Synchrotron radiation circular dichroism spectrum of purified MTSF1 protein at 2.6 mg/ml. (**C**) GMS assays performed using the purified His_6-MTSF1 protein and different synthetic RNA oligonucleotides in the 3′ region of *nad4* mRNA. In all, 40 pM of labeled RNA oligonucleotide and 0, 100, 200 or 400 nM of His_6-MTSF1 protein were used. The sequence of the synthetic RNA probes and the MTSF1 RNA-binding domain are shown. B, Bound; U, Unbound. (**D**) Diagram of the *nad4* coding sequence and 3′ UTR with the relative positions of the RNA probes (numbered from 1 to 4) used in gel shift assays. The MTSF1-binding site is shaded in gray.

endoribonucleases are involved in the formation of the 5′ ends of these short RNA molecules in plant mitochondria.

Our study also revealed that *mtsf1* mutants fail to synthesize a fully assembled complex I and instead accumulate low levels of a truncated version of this complex. This subcomplex I had a molecular mass of ~800 kDa, retains NADH-dehydrogenase activity and was still able to associate with respiratory complex III (Figure 5). As *mtsf1* mutants accumulated only low levels of mature *nad2* and *nad4* mRNAs, the production of this truncated complex I is likely caused by an insufficient synthesis of one, or both, of these complex I subunits. It is currently difficult to determine which subunit is most limiting and impacts complex I assembly in *mtsf1* plants. Nevertheless, the effect of *mtsf1* mutations on mature *nad4* mRNA accumulation is clearly stronger than the one on *nad2* transcripts.

The *mtsf1* plants do accumulate only trace levels of mature *nad4* mRNA, whereas our quantitative RT-PCR data indicated that mature *nad2* mRNA is simply reduced to ~25% of wild-type levels in *mtsf1* plants (Figure 6 and Supplementary Figure S5A). This amount of mature *nad2* mRNA is clearly not limiting for complex I assembly because other Arabidopsis mutants that accumulate much fewer mature *nad2* mRNAs compared with *mtsf1* still produce detectable amounts of fully assembled respiratory complex I (48). Additionally, a sub-complex I of similar size to the one observed in *mstf1* mutants and also showing NADH-dehydrogenase activity has been identified in other NAD4-deficient plants (49,50). These observations clearly favor the hypothesis that the major molecular event leading to the *mtsf1* mutant phenotype is linked to the deregulation of *nad4* expression and not to *nad2*, although additional molecular analyses are necessary to confirm these conclusions.

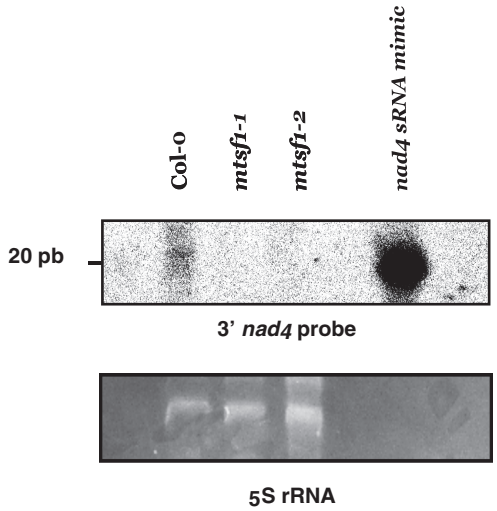

Figure 9. A short RNA corresponding to the MTSF1-binding site on *nad4* mRNA accumulates in mitochondria in an MTSF1-dependent manner. Total RNA extracted from the indicated genotypes (30 µg) was fractionated on a polyacrylamide gel before being transferred to nylon membrane. Blots were hybridized with a radiolabeled oligonucleotide complementary to the last 20 nucleotides of the *nad4* mRNA. A synthetic RNA oligonucleotide mimicking the MTSF1-binding site was also loaded in the last lane of the gel as indicated.

Among the Arabidopsis mutants showing impaired respiratory complex I, the *mtsf1* plants are typical examples of severely affected mutants. Effectively, *mtsf1* plants were extremely delayed in their development (Figure 1A). Homozygous *mtsf1* embryos did not reach complete maturity and <20% germinated (Supplementary Figures S1 and S3). Nevertheless, pollen development proceeded normally, and the overall fertility of *mtsf1* mutants decreased only moderately under all growth conditions tested (Supplementary Figure S2). Likewise, none of the Arabidopsis complex I mutants identified so far have been reported to show significant alteration in male or female gametophyte development, even for those like *mtsf1* that show severe growth retardation or in which no fully assembled complex I can be detected (46,47,50–55). In contrast, complex I-defective *cmsII* and *nms1* mutants in *Nicotiana sylvestris* have been described to be male sterile and recover partial male fertility only under high light (56,57). Nevertheless, our study showed that *mtsf1* mutants have abnormal respiratory and photosynthetic activities that are highly similar to the ones previously measured in *cmsII* plants (58). We effectively observed that the *mtsf1* mutations also lead to enhanced respiratory activity and induced pleiotropic effects on photosynthesis, with a lower rate of net CO_2 assimilation (Figure 3). Although the specific origin of these modifications is presently not well understood, they likely stem from both a higher commitment of electrons to the AOX pathway and inefficient electron donation, thereby leading to an uncoupling effect. Regarding the effect on photosynthesis, the larger respiratory CO_2 release in the light is not the cause of this alteration, as it can be compensated by a higher CO_2 mole fraction. These analyses revealed neither any major differences with the measurements done in the *cmsII* mutant of *N.sylvestris* nor any explanation for the male fertility of *mtsf1* mutants. The relationship between better gametophyte development, and complex I dysfunction in Arabidopsis is currently unclear, but it may involve higher induction capacity of the non-proton pumping alternative NAD(P)H dehydrogenases in reproductive tissues. Further investigations are needed to explore other possibilities and explain why male gametophyte development is less sensitive to complex I inactivation in Arabidopsis.

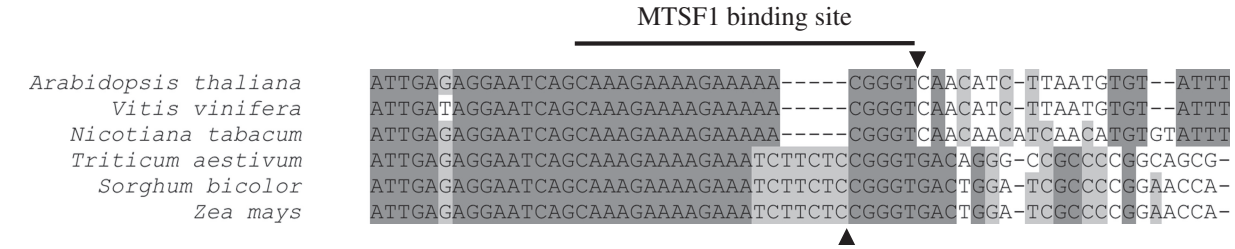

Figure 10. The Arabidopsis MTSF1-binding site is strictly conserved in dicots but slightly divergent in monocots. Multiple sequence alignment of the *nad4* 3′ UTR from a representative panel of dicot (*Arabidopsis thaliana, Vitis vinifera and Nicotiana tabacum*) and monocot (*Triticum aestivum, Sorghum bicolor, Zea mays*) plant species. The downward arrowhead indicates the end of the *nad4* mRNA 3′ UTR identified in *A.thaliana*, and the upward arrowhead indicates the one identified in *T.aestivum*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–5.

ACKNOWLEDGEMENTS

The authors are thankful to Emilie Elvira-Matelot and Katia Belcram for technical assistance. The authors also thank Dr Frank Wien (DISCO beamline at the SOLEIL synchrotron facility, Saclay, France) for his invaluable help during SRCD experiments.

FUNDING

Institut National de la Recherche Agronomique [INRA UMR 1318]; Agence National de la Recherche [ANR-09-BLAN-0244]; French *Ministère de l'Enseignement et de la Recherche* (to N.H.). Funding for open access charge: [IJPB, INRA UMR1318].

Conflict of interest statement. None declared.

REFERENCES

1. Gagliardi,D. and Binder,S. (2007) In: Logan,D. (ed.), *Plant Mitochondria*. Blackwell Publishing, Ames, IA, pp. 50–96.
2. Schmitz-Linneweber,C. and Small,I. (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci.*, **13**, 663–670.
3. Barkan,A. (2011) Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. *Plant Physiol.*, **155**, 1520–1532.
4. Small,I.D. and Peeters,N. (2000) The PPR motif - a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.*, **25**, 46–47.
5. Williams-Carrier,R., Kroeger,T. and Barkan,A. (2008) Sequence-specific binding of a chloroplast pentatricopeptide repeat protein to its native group II intron ligand. *RNA*, **14**, 1930–1941.
6. Lurin,C., Andrés,C., Aubourg,S., Bellaoui,M., Bitton,F., Bruyère,C., Caboche,M., Debast,C., Gualberto,J., Hoffmann,B. *et al*. (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell*, **16**, 2089–2103.
7. Prikryl,J., Rojas,M., Schuster,G. and Barkan,A. (2011) Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proc. Natl Acad. Sci. USA*, **108**, 415–420.
8. Beick,S., Schmitz-Linneweber,C., Williams-Carrier,R., Jensen,B. and Barkan,A. (2008) The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. *Mol. Cell. Biol.*, **28**, 5337–5347.
9. O'Toole,N., Hattori,M., Andrés,C., Iida,K., Lurin,C., Schmitz-Linneweber,C., Sugita,M. and Small,I. (2008) On the expansion of the pentatricopeptide repeat gene family in plants. *Mol. Biol. Evol.*, **25**, 1120–1128.
10. Chase,C.D. (2007) Cytoplasmic male sterility: a window to the world of plant mitochondrial–nuclear interactions. *Trends Genet.*, **23**, 81–90.
11. Fujii,S., Bond,C.S. and Small,I.D. (2011) Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. *Proc. Natl Acad. Sci. USA*, **108**, 1723–1728.
12. Stern,D.B., Goldschmidt-Clermont,M. and Hanson,M.R. (2010) Chloroplast RNA Metabolism. *Annu. Rev. Plant Biol.*, **61**, 125–155.
13. Pfalz,J., Bayraktar,O.A., Prikryl,J. and Barkan,A. (2009) Site-specific binding of a PPR protein defines and stabilizes 5′ and 3′ mRNA termini in chloroplasts. *EMBO J.*, **28**, 2042–2052.
14. Fisk,D.G., Walker,M.B. and 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.
15. Boudreau,E., Nickelsen,J., Lemaire,S.D., Ossenbühl,F. and Rochaix,J.D. (2000) The Nac2 gene of Chlamydomonas encodes a chloroplast TPR-like protein involved in psbD mRNA stability. *EMBO J.*, **19**, 3366–3376.
16. Vaistij,F.E., Boudreau,E., Lemaire,S.D., Goldschmidt-Clermont,M. and Rochaix,J.D. (2000) Characterization of Mbb1, a nucleus-encoded tetratricopeptide-like repeat protein required for expression of the chloroplast psbB/psbT/psbH gene cluster in *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA*, **97**, 14813–14818.
17. Murakami,S. (2005) A spontaneous tRNA suppressor of a mutation in the *Chlamydomonas reinhardtii* nuclear MCD1 gene required for stability of the chloroplast petD mRNA. *Nucleic Acids Res.*, **33**, 3372–3380.
18. Loiselay,C., Gumpel,N.J., Girard-Bascou,J., Watson,A.T., Purton,S., Wollman,F.-A. and Choquet,Y. (2008) Molecular identification and function of cis- and trans-acting determinants for petA transcript stability in *Chlamydomonas reinhardtii* chloroplasts. *Mol. Cell. Biol.*, **28**, 5529–5542.
19. Hattori,M. and Sugita,M. (2009) A moss pentatricopeptide repeat protein binds to the 3′ end of plastid clpP pre-mRNA and assists with mRNA maturation. *FEBS J.*, **276**, 5860–5869.
20. Johnson,X., Wostrikoff,K., Finazzi,G., Kuras,R., Schwarz,C., Bujaldon,S., Nickelsen,J., Stern,D.B., Wollman,F.-A. and Vallon,O. (2010) MRL1, a conserved Pentatricopeptide repeat protein, is required for stabilization of rbcL mRNA in Chlamydomonas and Arabidopsis. *Plant Cell*, **22**, 234–248.
21. Zhelyazkova,P., Hammani,K., Rojas,M., Voelker,R., Vargas-Suarez,M., Borner,T. and Barkan,A. (2011) Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. *Nucleic Acids Res.*, **40**, 3092–3105.
22. Hammani,K.K., Cook,W.B.W. and Barkan,A.A. (2012) RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. *Proc. Natl Acad. Sci. USA*, **109**, 5651–5656.
23. Sharwood,R.E., Halpert,M., Luro,S., Schuster,G. and Stern,D.B. (2011) Chloroplast RNase J compensates for inefficient transcription termination by removal of antisense RNA. *RNA*, **17**, 2165–2176.
24. Germain,A., Herlich,S., Larom,S., Kim,S.H., Schuster,G. and Stern,D.B. (2011) Mutational analysis of Arabidopsis chloroplast polynucleotide phosphorylase reveals roles for both RNase PH core domains in polyadenylation, RNA 3′-end maturation and intron degradation. *Plant J.*, **67**, 381–394.
25. Kubo,T. and Newton,K.J. (2008) Angiosperm mitochondrial genomes and mutations. *Mitochondrion*, **8**, 5–14.
26. Forner,J., Weber,B., Thuss,S., Wildum,S. and Binder,S. (2007) Mapping of mitochondrial mRNA termini in Arabidopsis thaliana: t-elements contribute to 5′ and 3′ end formation. *Nucleic Acids Res.*, **35**, 3676–3692.
27. Jonietz,C., Forner,J., Hölzle,A., Thuss,S. and Binder,S. (2010) RNA PROCESSING FACTOR2 is required for 5′ end processing of nad9 and cox3 mRNAs in mitochondria of Arabidopsis thaliana. *Plant Cell*, **22**, 443–453.
28. Hölzle,A., Jonietz,C., Törjek,O., Altmann,T., Binder,S. and Forner,J. (2011) A RESTORER OF FERTILITY-like PPR gene is required for 5′-end processing of the nad4 mRNA in mitochondria of Arabidopsis thaliana. *Plant J.*, **65**, 737–744.
29. Jonietz,C., Forner,J., Hildebrandt,T. and Binder,S. (2011) RNA PROCESSING FACTOR3 is crucial for the accumulation of mature ccmC transcripts in mitochondria of Arabidopsis accession Columbia. *Plant Physiol.*, **157**, 1430–1439.
30. Hauler,A., Jonietz,C., Stoll,B., Stoll,K., Braun,H.-P. and Binder,S. (2013) RNA PROCESSING FACTOR 5 is required for efficient 5′ cleavage at a processing site conserved in RNAs of three different mitochondrial genes in Arabidopsis thaliana. *Plant J.*, doi:10.1111/tpj.12143.
31. Perrin,R., Meyer,E.H., Zaepfel,M., Kim,Y.J., Mache,R., Grienenberger,J.M., Gualberto,J.M. and Gagliardi,G. (2004)

AtmtPNPase is required for multiple aspects of the 18S rRNA metabolism in Arabidopsis thaliana mitochondria. *Nucleic Acids Res.*, **32**, 5174–5182.
32. Monde,R.A., Greene,J.C. and Stern,D.B. (2000) The sequence and secondary structure of the 3′-UTR affect 3′-end maturation, RNA accumulation, and translation in tobacco chloroplasts. *Plant Mol. Biol.*, **44**, 529–542.
33. Yehudai-Resheff,S.S., Hirsh,M.M. and Schuster,G.G. (2001) Polynucleotide phosphorylase functions as both an exonuclease and a poly(A) polymerase in spinach chloroplasts. *Mol. Cell. Biol.*, **21**, 5408–5416.
34. Schuster,W., Hiesel,R., Isaac,P.G., Leaver,C.J. and Brennicke,A. (1986) Transcript termini of messenger RNAs in higher plant mitochondria. *Nucleic Acids Res.*, **14**, 5943–5954.
35. Colas des Francs-Small,C., Kroeger,T., Zmudjak,M., Ostersetzer-Biran,O., Rahimi,N., Small,I. and Barkan,A. (2012) A PORR domain protein required for rpl2 and ccmFC intron splicing and for the biogenesis of c-type cytochromes in Arabidopsis mitochondria. *Plant J.*, **69**, 996–1005.
36. Czechowski,T., Stitt,M., Altmann,T., Udvardi,M.K. and Scheible,W.-R. (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol.*, **139**, 5–17.
37. Uyttewaal,M., Arnal,N., Quadrado,M., Martin-Canadell,A., Vrielynck,N., Hiard,S., Gherbi,H., Bendahmane,A., Budar,F. and Mireau,H. (2008) Characterization of Raphanus sativus pentatricopeptide repeat proteins encoded by the fertility restorer locus for Ogura cytoplasmic male sterility. *Plant Cell*, **20**, 3331–3345.
38. Lamattina,L., Gonzalez,D., Gualberto,J. and Grienenberger,J.M. (1993) Higher plant mitochondria encode an homologue of the nuclear-encoded 30-kDa subunit of bovine mitochondrial complex I. *Eur. J. Biochem.*, **217**, 831–838.
39. Colas des Francs-Small,C., Ambard-Bretteville,F., Darpas,A., Sallantin,M., Huet,J.C., Pernollet,J.C. and Rémy,R. (1992) Variation of the polypeptide composition of mitochondria isolated from different potato tissues. *Plant Physiol.*, **98**, 273–278.
40. Lees,J.G., Smith,B.R., Wien,F., Miles,A.J. and Wallace,B.A. (2004) CDtool—an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving. *Anal. Biochem.*, **332**, 285–289.
41. Whitmore,L. and Wallace,B.A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.*, **32**, 668–673.
42. Lees,J.G., Miles,A.J., Wien,F. and Wallace,B.A. (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics*, **22**, 1955–1962.
43. Nelson,B.K., Cai,X. and Andreas,N. (2007) A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. *Plant J.*, **51**, 1126–1136.
44. Fujiki,Y., Hubbard,A.L., Fowler,S. and Lazarow,P.B. (1982) Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. *J. Cell Biol.*, **93**, 97–102.
45. Ruwe,H. and Schmitz-Linneweber,C. (2011) Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? *Nucleic Acids Res.*, **40**, 3106–3116.
46. Koprivova,A., Colas des Francs-Small,C., Calder,G., Mugford,S.T., Tanz,S., Lee,B.-R., Zechmann,B., Small,I. and Kopriva,S. (2010) Identification of a pentatricopeptide repeat protein implicated in splicing of intron 1 of mitochondrial nad7 transcripts. *J. Biol. Chem.*, **285**, 32192–32199.
47. de Longevialle,A.F., Meyer,E.H., Andrés,C., Taylor,N.L., Lurin,C., Millar,A.H. and Small,I.D. (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.
48. Kühn,K., Carrie,C., Giraud,E., Wang,Y., Meyer,E.H., Narsai,R., des Francs-Small,C.C., Zhang,B., Murcha,M.W. and Whelan,J. (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.
49. Karpova,O.V. and Newton,K.J. (1999) A partially assembled complex I in NAD4-deficient mitochondria of maize. *Plant J.*, **17**, 511–521.
50. 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.
51. 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.
52. Liu,Y., He,J., Chen,Z., Ren,X., Hong,X. and Gong,Z. (2010) ABA overly-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.
53. Zhu,Q., Dugardeyn,J., Zhang,C., Takenaka,M., Kühn,K., Craddock,C., Smalle,J., Karampelias,M., Denecke,J., Peters,J. *et al.* (2012) SLO2, a mitochondrial PPR protein affecting several RNA editing sites, is required for energy metabolism. *Plant J.*, **71**, 836–49.
54. Yuan,H. and Liu,D. (2012) Functional disruption of the pentatricopeptide protein SLG1 affects mitochondrial RNA editing, plant development, and responses to abiotic stresses in Arabidopsis. *Plant J.*, **70**, 432–444.
55. Keren,I., Tal,L., Colas des Francs-Small,C., Araújo,W.L., Shevtsov,S., Shaya,F., Fernie,A.R., Small,I. and Ostersetzer-Biran,O. (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.
56. Gutierres,S., Combettes,B., De Paepe,R., Mirande,M., Lelandais,C., Vedel,F. and Chetrit,P. (1999) In the Nicotiana sylvestris CMSII mutant, a recombination-mediated change 5′ to the first exon of the mitochondrial nad1 gene is associated with lack of the NADH:ubiquinone oxidoreductase (complex I) NAD1 subunit. *Eur. J. Biochem.*, **261**, 361–370.
57. Brangeon,J., Sabar,M., Gutierres,S., Combettes,B., Bove,J., Gendy,C., Chetrit,P., Colas des Francs-Small,C., Pla,M., Vedel,F. *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.
58. Priault,P., Tcherkez,G., Cornic,G., De Paepe,R., Naik,R., Ghashghaie,J. and Streb,P. (2006) The lack of mitochondrial complex I in a CMSII mutant of Nicotiana sylvestris increases photorespiration through an increased internal resistance to CO2 diffusion. *J. Exp. Bot.*, **57**, 3195–3207.