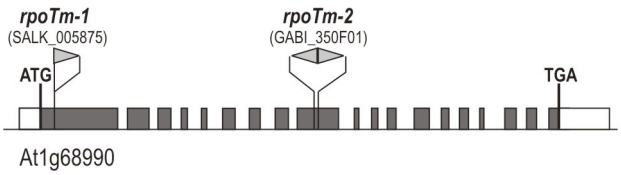
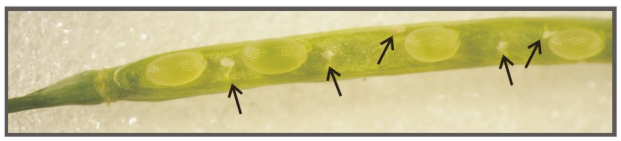
Supplemental data. Kühn et al. (2009). Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of *Arabidopsis thaliana.*

A



B



C

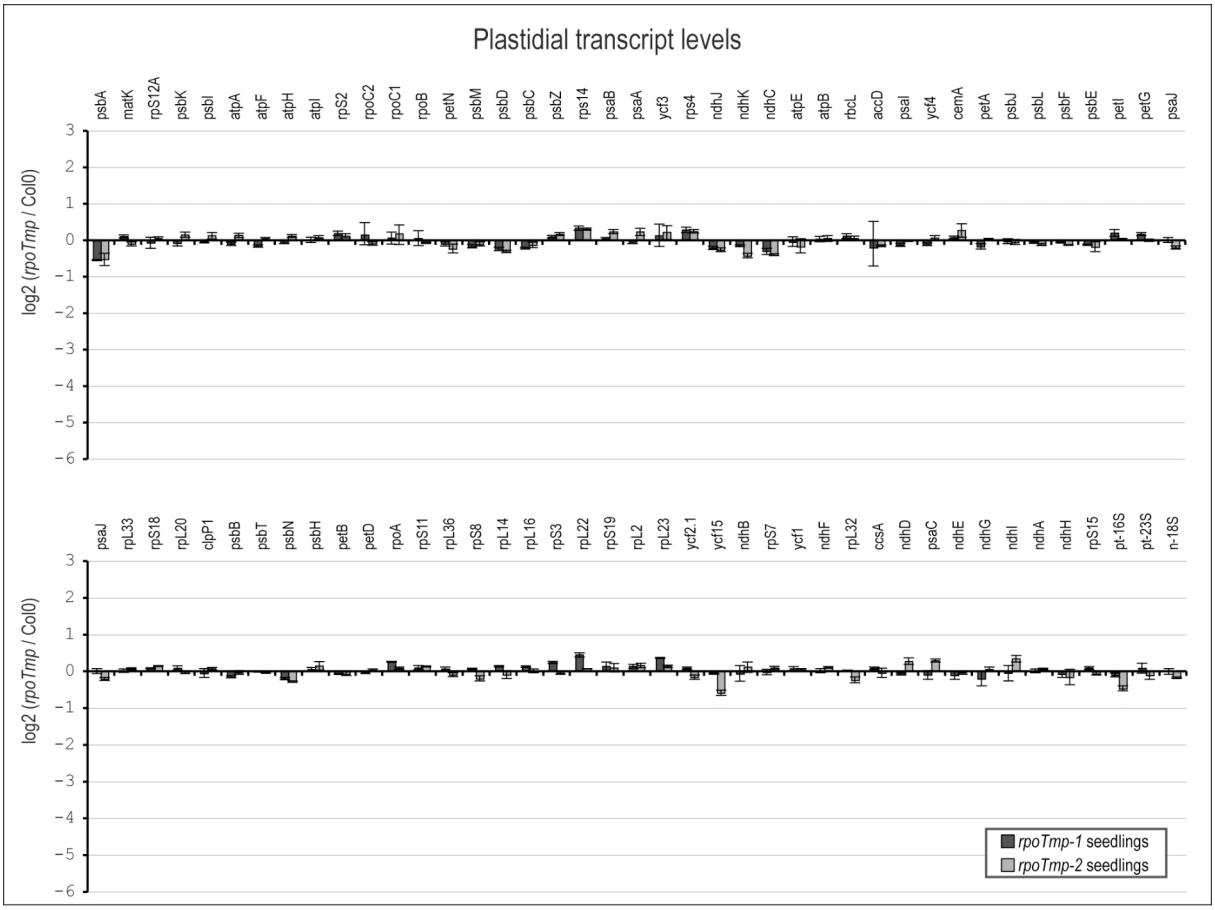
Self-pollination of plants heterozygous for a disrupted *rpoTm* allele:

Genotype of heterozygous parent	Germination	Heterozygous Progeny
RPOTm / rpoTm-1	100%	24/71 (34%)
RPOTm / rpoTm-2	100%	24/72 (33%)

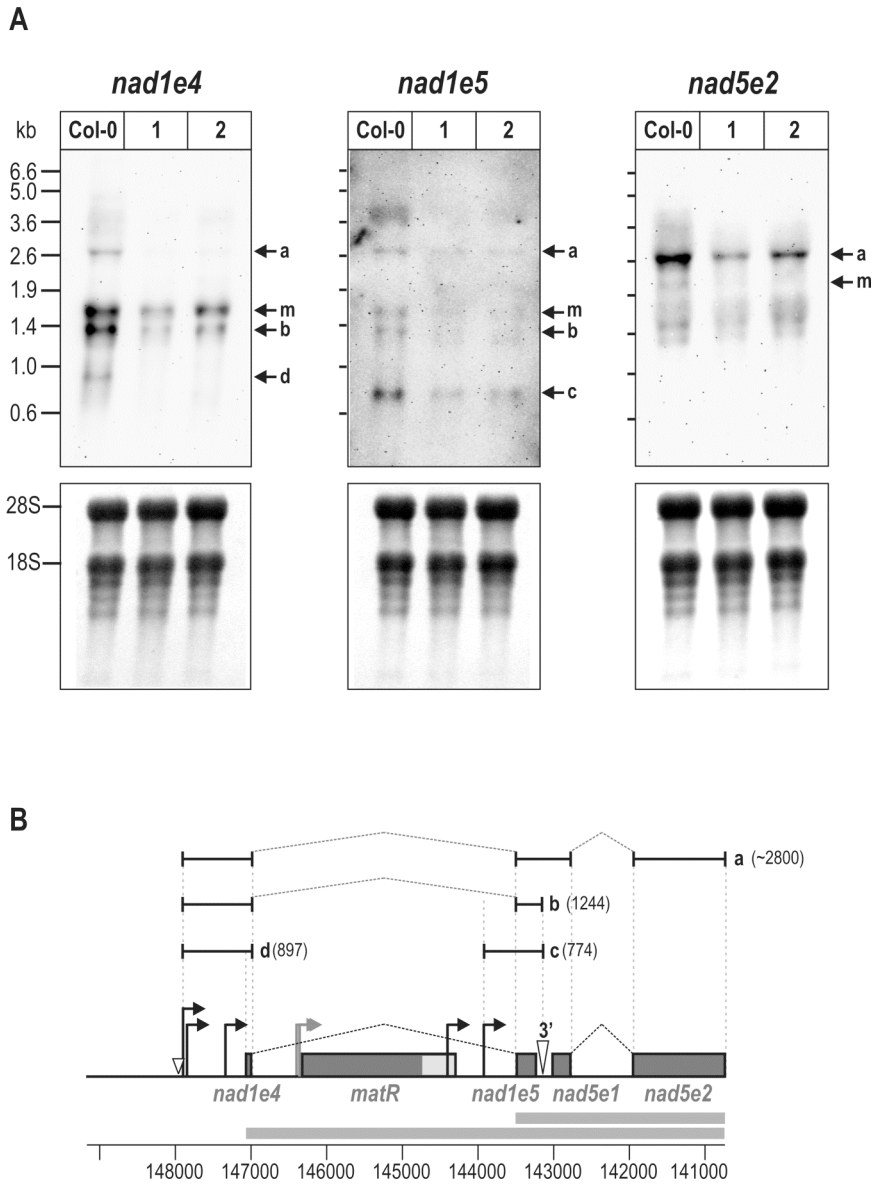
Reciprocal backcrosses:

Female	Male	Germination	Heterozygous Progeny
RPOTm / RPOTm	*RPOTm / rpoTm-1*	100%	25/96 (26%)
RPOTm / rpoTm-1	*RPOTm / RPOTm*	100%	27/96 (28%)
RPOTm / RPOTm	*RPOTm / rpoTm-2*	100%	9/96 (9%)
RPOTm / rpoTm-2	*RPOTm / RPOTm*	100%	35/96 (36%)

Supplemental Figure 1. Lethality of loss of RPOTm.
(A) *RPOTm* gene organisation and positions of T-DNA insertions in two *rpoTm* mutant alleles *rpoTm-1* (SALK_005875, obtained from the ABRC stock centre) and *rpoTm-2* (GABI_350F01, obtained from GABI-Kat). *RPOTm* gene exons are represented as boxes (dark grey, coding sequences; white, untranslated regions). Grey arrowheads indicate T-DNA insertions pointing towards the T-DNA left border. Through PCR with primers annealing to T-DNA borders and their flanking *RPOTm* sequence, the insertion in *rpoTm-2* was found to be present as inverted repeat of the T-DNA, while *rpoTm-1* carried a single insertion. Insertion sites were confirmed by sequencing of PCR products.
(B) Silique of an *RPOTm* / *rpoTm-1* heterozygous plant showing aborted ovules/embryos (arrows). We were unable to obtain plants homozygous for a disrupted *rpoTm* allele.
(C) Segregation analysis of *rpoTm* alleles supporting lethality of the loss of RPOTm and reduced transmission of *rpoTm* alleles. Following self-pollination of *RPOTm* / *rpoTm* heterozygous plants and reciprocal backcrosses of heterozygous plants with wild-type plants (Col-0), progenies were analysed for their genotypes at At1g68990, and the frequency of heterozygous progeny was determined. Observed frequencies differed significantly from i) frequencies expected for lethality of *rpoTm* / *rpoTm* embryos but complete *rpoTm* gamete viability and ii) frequencies expected for male *rpoTm* gamete lethality but female *rpoTm* gamete viability and *vice versa* ($P < 0.01$; chi square test of independence). Therefore, transmission of *rpoTm* alleles by both male and female gametes is reduced, which could be due to decreased male and female *rpoTm* gamete fitness, partial gamete lethality or decreased fitness of heterozygous embryos.



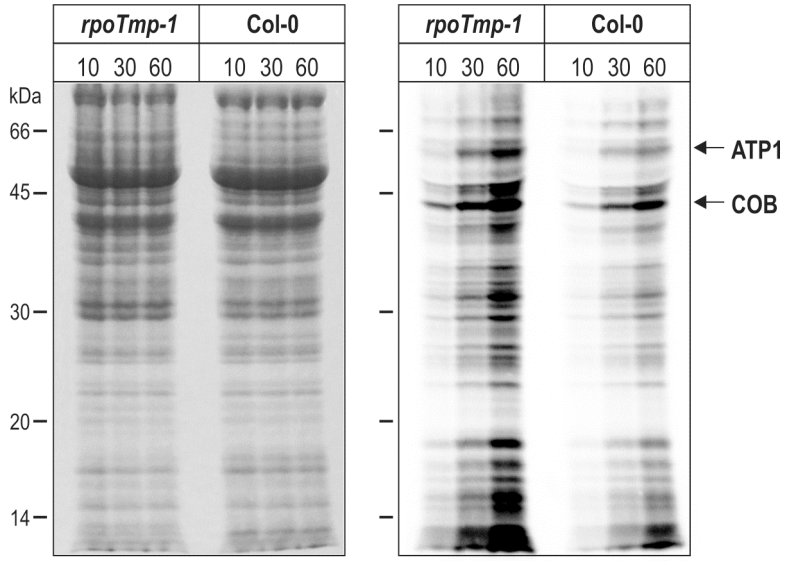
Supplemental Figure 2. Plastidial steady-state transcript levels in *rpoTmp* mutants.
Transcript levels are depicted as the log2 ratio of transcript levels in mutants compared with levels in wild-type plants (Col-0). Transcript abundances of all plastidial protein-encoding and rRNA genes were determined by qRT-PCR in both mutant and wild-type seedlings using the protocol and primers described earlier (Falcon de Longevialle, A., Hendrickson, L., Taylor, N.L., Delannoy, E., Lurin, C., Badger, M., Millar, A.H. and Small, I. (2008) The pentatricopeptide repeat gene *OTP51* with two LAGLIDADG motifs is required for the *cis*-splicing of plastid *ycf3* intron 2 in Arabidopsis thaliana. *Plant J*, **56**, 157-168.). Three technical replicates were averaged per genotype; standard errors are indicated. No differences between mutant and wild-type transcript levels were seen.



Supplemental Figure 3. RNA gel blot hybridizations with probes specific to exons *nad1e4*, *nad1e5* and *nad5e2*.

(**A**) Probes were hybridized to 10 µg of filter-immobilized total leaf RNA isolated from *rpoTmp-1* (lanes 1), *rpoTmp-2* (2) and wild-type (Col-0) seedlings (upper panels). RNA size markers were run alongside samples; marker sizes are indicated. Signals corresponding in size to expected mature mRNAs are indicated by arrows and labeled m. Incompletely processed transcript species are labeled a, b, c and d according to transcript identities as laid out in (B). All identified transcripts are reduced in abundance in mutants compared to the wild type. The greater part of signals corresponds to incompletely processed RNAs. The same filters were stained with methylene blue and are shown as a loading control in lower panels. Slightly more RNA was loaded for *rpoTmp-2* in the *nad1e4* and *nad5e2* blots, which could explain the slightly stronger signals in comparison with *rpoTmp-1*.

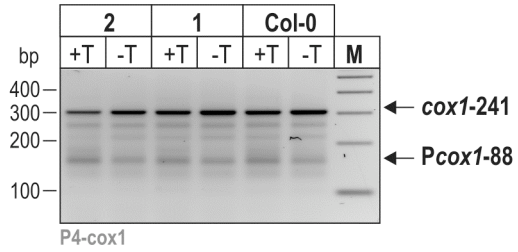
(**B**) Map of mtDNA region comprising nad1e4, matR, nad1e5, nad5e1 and nad5e2 as shown in Figure 6A. The MatR domain X coding region, which in maize mitochondria is transcribed from an additional promoter (Farre, J.C., and Araya, A. (1999). The mat-r open reading frame is transcribed from a non-canonical promoter and contains an internal promoter to co-transcribe exons nad1e and nad5III in wheat mitochondria. Plant Mol Biol 40, 959-967.), is highlighted in light grey. The positions of transcripts detected in (A) are indicated above the map by horizontal lines labelled with identification letters a, b, c and d, and expected transcript sizes. Sizes have been inferred from an indicated transcript 3' end (triangle) defined by Forner et al. (2007) and from our identification of primary transcript 5' ends mapping to P*nad1e4*-838 and P*nad1e5*-427. Splicing of transcripts a and b is indicated by dashed bent lines between exons.

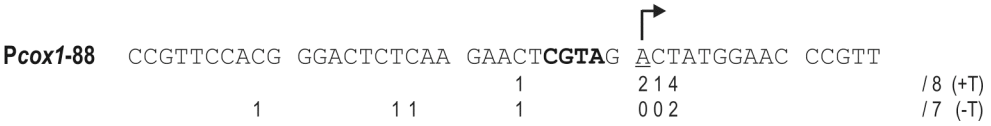


Supplemental Figure 4. In organello protein synthesis by mitochondria isolated from *rpoTmp* and wild-type plants.

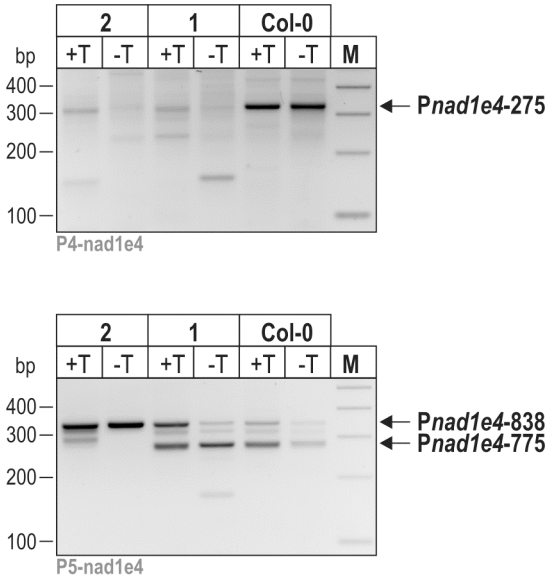
Coomassie-stained gel (left) and autoradiogram (right) of proteins synthesized in organello for 10, 30 and 60 minutes by mitochondria isolated from *rpoTmp-1* and and wild-type (Col-0) rosettes and resolved by SDS-PAGE on a 12% polyacrylamide gel. In organello protein synthesis assays were done as described by Giegé et al. (2005). Two major bands in the autoradiogram have been assigned to ATP1 and COB according to the labelled protein profiles shown by Giegé et al. (2005). The very hydrophobic COX1 and NAD6 proteins, whose transcripts are strongly reduced in *rpoTmp*, do probably not run as clear bands and cannot be distinguished in the autoradiogram. The decrease in translation of these proteins cannot be addressed by the in organello protein synthesis assay.

Although contaminating bacterial protein synthesis is detected in the autoradiogram (e.g. proteins larger than ATP1), major bands are of mitochondrial origin. Protein synthesis appears enhanced in isolated mutant mitochondria. This could be because preparations of mutant mitochondria were of better quality, or it could be related to many transcripts showing elevated levels in the mutants. Mutant assays were set up before wild-type assays and briefly kept on ice until all reactions were transferred to the incubation temperature of 25˚C. It is possible that labelled protein was already made in samples kept at 4˚C, contributing to the increased detection of protein synthesis in mutant mitochondria.





Supplemental Figure 5. 5'-RACE analysis of *cox1* transcripts.
5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products from transcript 5' ends are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for amplification is indicated below the gel image.
Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *cox1* upstream sequence that surrounds the identified TSS is shown, with the CGTA tetranucleotide, which is frequently associated with mitochondrial promoters, printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type to correspond to transcript 5' ends mapping to the respective positions (row +T, clone numbers for products of 5'-RACE following TAP treatment; row –T, clone numbers determined without TAP treatment). Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter. The TSS, which gave rise to a TAP-specific 5'-RACE product, is indicated by a bent arrow.
The *cox1* promoter P*cox1*-355 has been characterized previously; no detailed mapping was therefore done on wild-type transcripts mapping to this site.

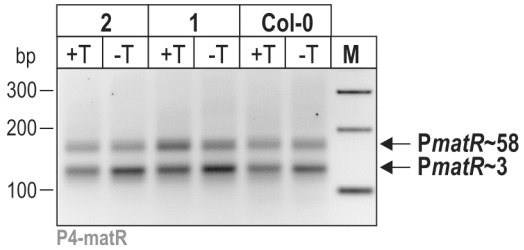



```
Pnad1e4-275  AATTTTCGGA ATCTACAAAA ACATTCTAGG AAAGGGCTCG TAATG
                                    62 2                  / 10 (+T)
                          1  1 1    1211                  /  8 (-T)

Pnad1e4-775  TCATAGAACC GATCGACTGC TATCTAAGAG AAGATAAGTA GTTCT
                                  4  3  2                 /  9 (+T)
                                  7  0  7        1        / 15 (-T)

Pnad1e4-838  GGTGCTACGA GATGGCAATT TATCGTATAG AAGAATAGAT CCGTG
                                    316   2               / 12 (+T)
                                    000  12    1          / 13 (-T)
```

Supplemental Figure 6. 5'-RACE analysis of *nad1e4* transcripts.
5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products from transcript 5' ends are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for amplification is indicated below the gel image. No signal is seen in the displayed 5'-RACE experiments for P*nad1e4*-775 in *rpoTmp-2* but has been detected in replicates of this experiment (not shown).
Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *nad1e4* upstream sequence that surrounds the identified TSS is shown, with the TCTA and CGTA tetranucleotides, which are found associated with mitochondrial promoters, printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type to correspond to transcript 5' ends mapping to the respective positions (row +T, clone numbers for products of 5'-RACE following TAP treatment; row –T, clone numbers determined without TAP treatment). Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter. The TSS, which gave rise to a TAP-specific 5'-RACE product, is indicated by a bent arrow; the triangle corresponds to the site of a likely processed end.



PmatR~3 TTTCCCCCCT TTACAATATG AAGAAAGAAA GAA*GGGTTGA AGTTT*

1 3 2 1 2 3 2 / 14 (+T)

1 1 1 1 0 1 1 0 4 / 13 (-T)

PmatR~58 ACGGCAGGAG GGGGCCTTTG **ATTA**TAAAAA AGGGGAGCAA AAAAC

3 4 4 1 / 12 (+T)

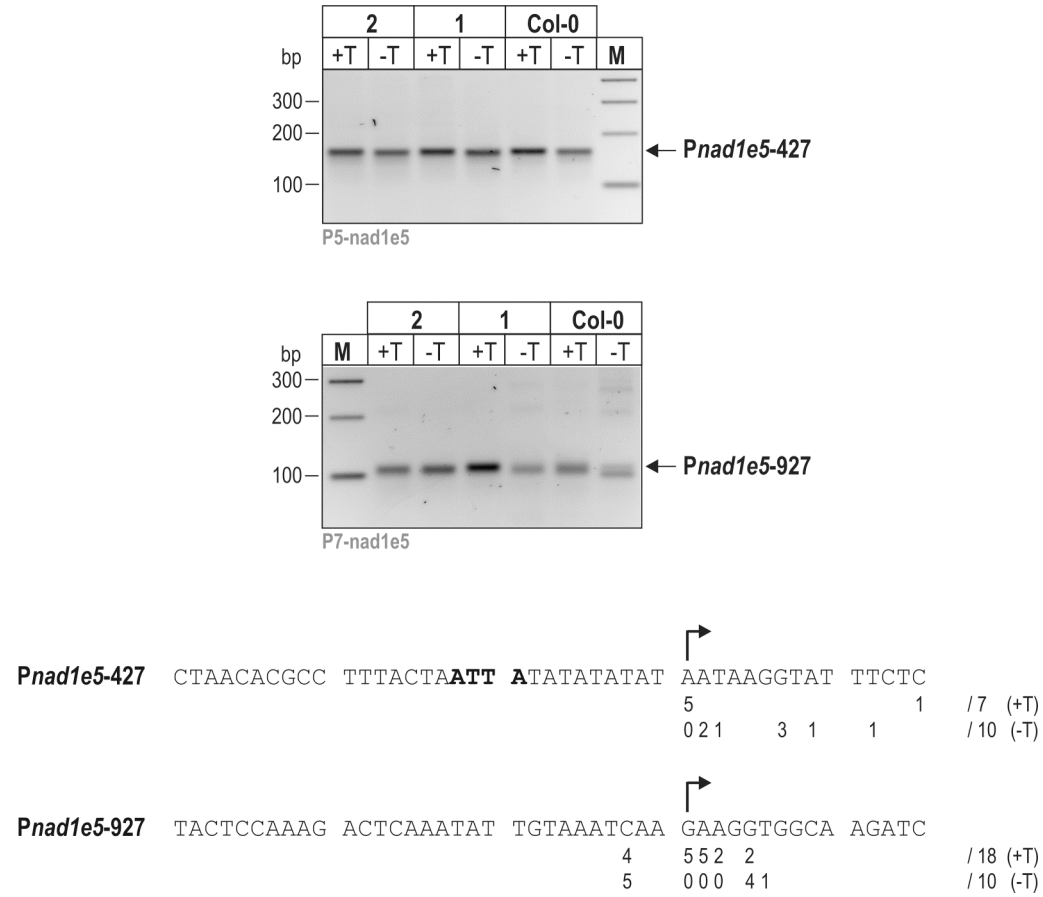
Supplemental Figure 7. 5'-RACE analysis of *matR* transcripts.

5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products from transcript 5' ends are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for amplification is indicated below the gel image.

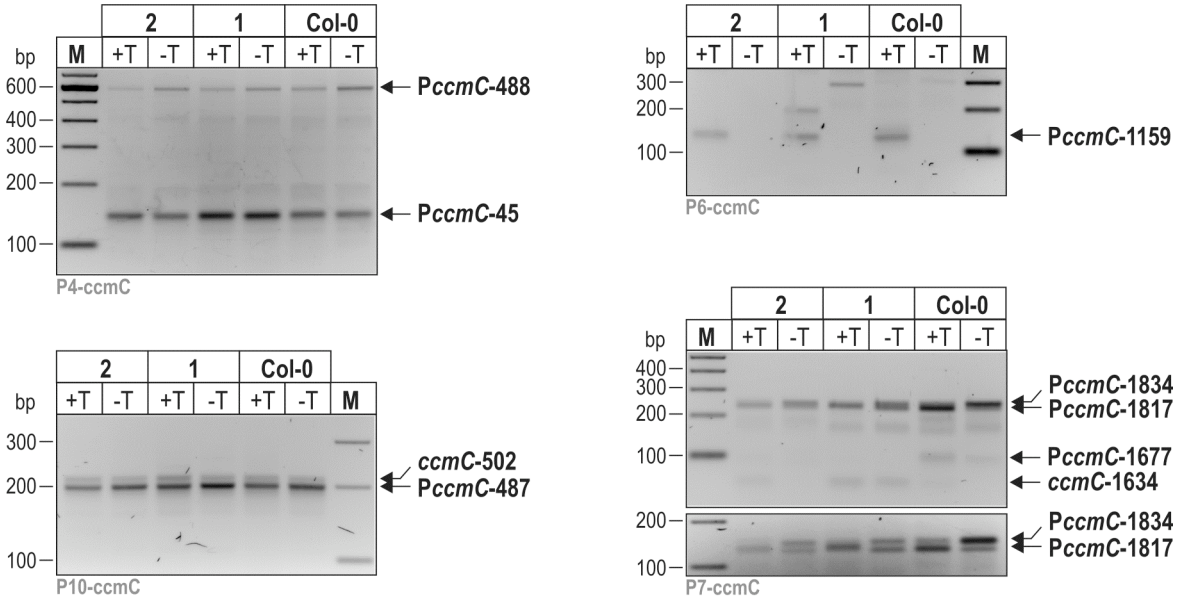
Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *matR* upstream sequence that surrounds the suggested TSS is shown, with the ATTA tetranucleotide, which is frequently associated with mitochondrial promoters, printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type to correspond to transcript 5' ends mapping to the respective positions (row +T, clone numbers for products of 5'-RACE following TAP treatment; row –T, clone numbers determined without TAP treatment). Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter.

The identified *matR* 5' ends were highly diverse and showed no differences between TAP-treated and untreated samples in gel images. A TAP-specific 5' end was detected mapping to position -3 in a highly A-rich sequence stretch; yet, because of the diversity of ends mapping to positions immediately upstream of the annotated MATR coding region (coding nucleotides are italicized), transcription initiation at position -3 is not certain.

No transcripts mapping to P*matR*-58 were analyzed from RNA not exposed to TAP. This site is considered a putative promoter because of its ATTA element and the generally A/T-rich sequence.



Supplemental Figure 8. 5'-RACE analysis of *nad1e5* transcripts.
5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products from transcript 5' ends are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for amplification is indicated below the gel image. Products in lanes marked +T are migrating more slowly than products in lanes –T, which is typical for primary transcripts that are post-transcriptionally clipped at the 5' ends by a few nucleotides.
Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *nad1e5* upstream sequence that surrounds the identified TSS is shown, with the ATTA tetranucleotide, which is found associated with mitochondrial promoters, printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type to correspond to transcript 5' ends mapping to the respective positions (row +T, clone numbers for products of 5'-RACE following TAP treatment; row –T, clone numbers determined without TAP treatment). Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter. The TSS, which gave rise to a TAP-specific 5'-RACE product, is indicated by a bent arrow.



P*ccmC*-45 GTTGGGGGAA ATCCTTTGAT TG**CGTATTA**T AGATCCATGT CTTTC
4 2 1 8 1 / 16 (+T)
0 2 5 2 3 1 1 / 15 (-T)

-502
P*ccmC*-487 ATGAGTTGAT CGATATCGAT CGG**CGTA**AGA AAAGATTCTA CTATC
1 1 1 2 1 2 6 3 / 17 (+T)
1 1 3 0 2 1 1 3 3 1 / 16 (-T)

P*ccmC*-1159 CAGCTAATCA ATTGATTCGT CAT**GGTA**GAG AAGAAAAACG GCGCA
6 / 6 (+T)

P*ccmC*-1677 CTTTCCTTTC TGGCTACCAA ACGTGCTAAT AAAAGAGCTT ATTCT
8 1 / 9 (+T)
0 0 5 2 / 7 (-T)

P*ccmC*-1817 TACTATATAA GATAAGTAAT TC**ATTA**TAAG ATA**AGTA**ATT CATTG
1 3 9 / 13 (+T)
0 0 4 2 2 / 8 (-T)

P*ccmC*-1834 CGAAGAGAAG GTTACACTAC TATATAAGAT AAGTAATTCA TTATA
2 / 2 (+T)
1 4 1 / 6 (-T)

Supplemental Figure 9. 5'-RACE analysis of *ccmC* transcripts.

5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products from transcript 5' ends are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for amplification is indicated below the gel image.

Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *ccmC* upstream sequence that surrounds the identified TSS or putative TSS is shown; tetranucleotide elements that are frequently associated with mitochondrial promoters are printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type to correspond to transcript 5' ends mapping to the respective positions (row +T, clone numbers for products of 5'-RACE following TAP treatment; row –T, clone numbers determined without TAP treatment). Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter. The TSS and putative TSS are indicated by bent arrows.

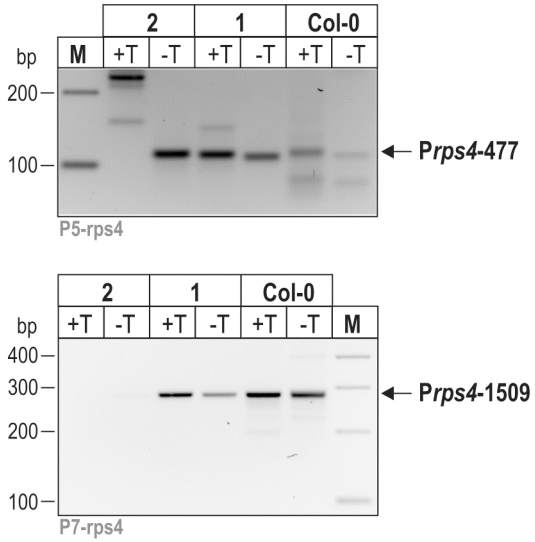
While no differences were seen in P*ccmC*-45 5'-RACE gel patterns between +TAP and –TAP samples, sequencing of cloned 5'-RACE products identified a 5' end that was exclusive to TAP-treated RNA and therefore represents a primary end.

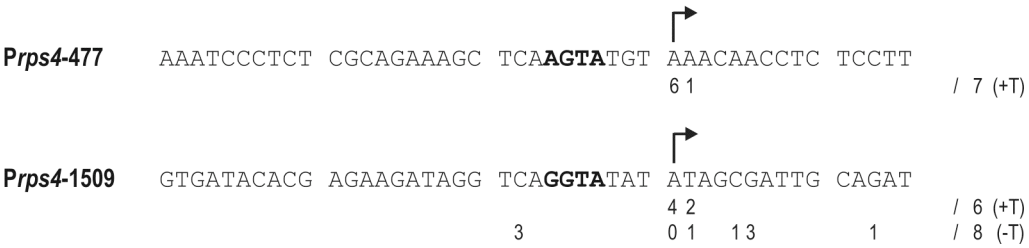
No major TAP-specific 5' end was found for P*ccmC*-487, which showed a variety of ends and putative upstream (position -502) and downstream (triangle; this 5' end has also been mapped by Forner et al. 2007). Because of the presence of a CGTA tetranucleotide, P*ccmC*-487 might function as TSS despite the lack of experimental proof.

No products mapping to P*ccmC*-1159 were amplified from –TAP samples, identifying this site as a TSS.

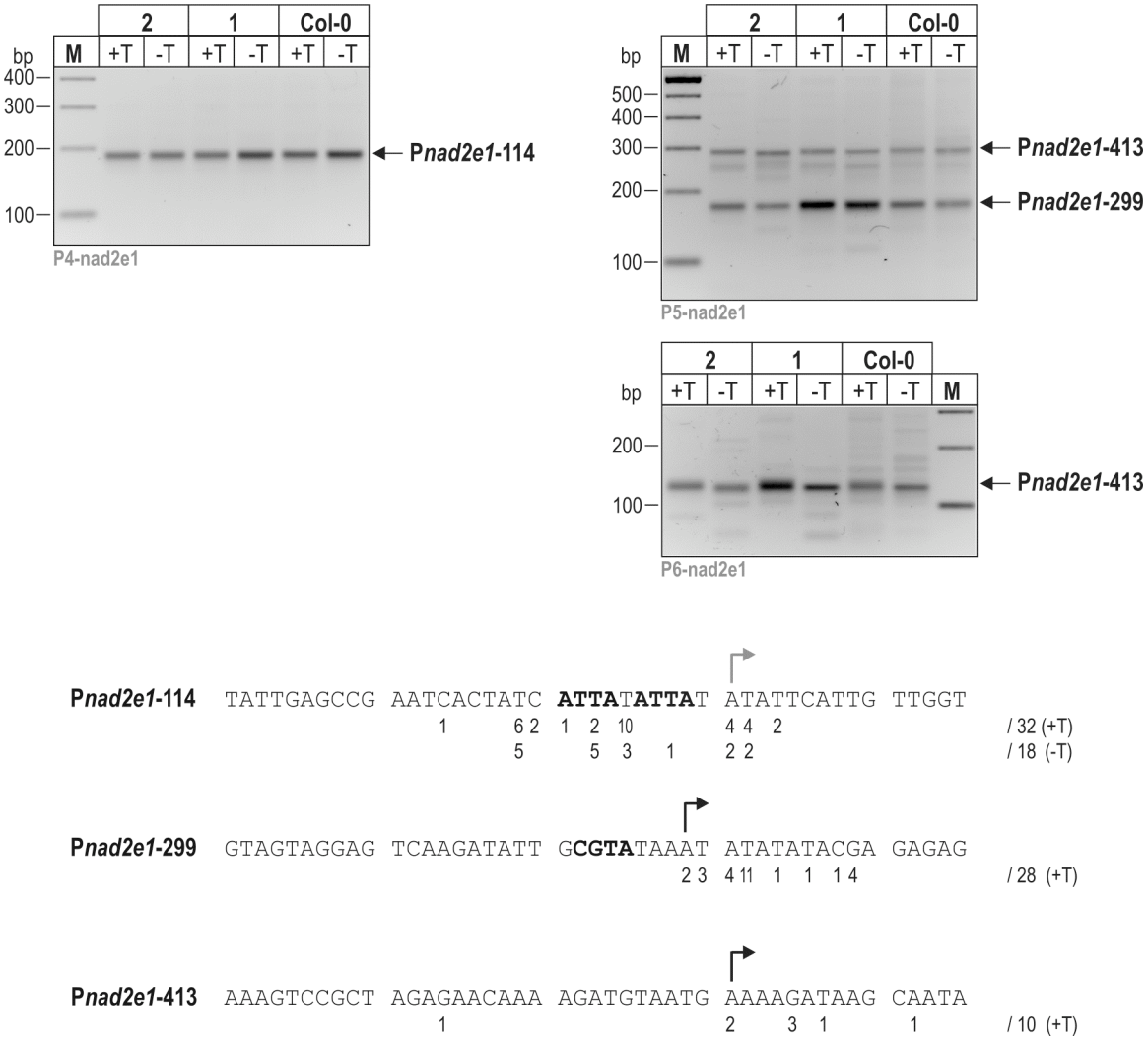
Sequencing of products mapping to P*ccmC*-1677 and P*ccmC*-1817 detected TAP-specific, and therefore primary, transcript 5' ends.

Transcripts mapping to position -1834 appear as processed transcripts in gel patterns, and sequencing of cloned 5' ends was insufficient to support these transcripts being primary. However, because the region is A/T-rich, P*ccmC*-1834 is suggested as a putative TSS.

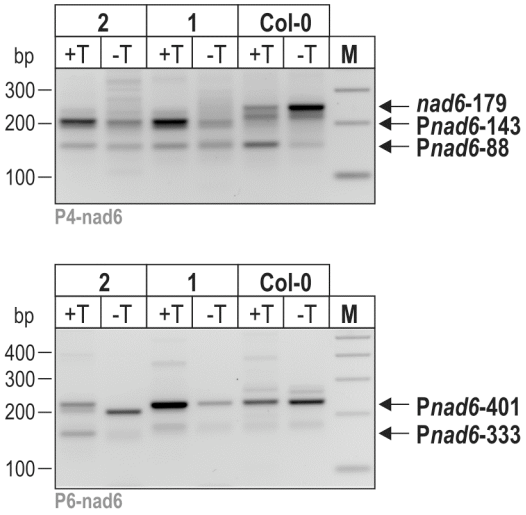




Supplemental Figure 10. 5'-RACE analysis of *rps4* transcripts.
5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products from transcript 5' ends are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for amplification is indicated below the gel image. In 5'-RACE assays shown here for *rpoTmp-2*, no specific products were amplified from TAP-treated RNA. In a replicate of this experiment, specific products were seen for *rpoTmp-2* but not *rpoTmp-1*. We think that the low abundance of 5' ends mapping to P*rps4*-477 and Prps4-1509 in mutants effects on the reproducibility of the experiment.
Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *rps4* upstream sequence that surrounds the identified TSS is shown, with the AGTA and GGTA tetranucleotides, which are frequently associated with mitochondrial promoters, printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type to correspond to transcript 5' ends mapping to the respective positions (row +T, clone numbers for products of 5'-RACE following TAP treatment; row –T, clone numbers determined without TAP treatment). Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter. The TSS, which gave rise to a TAP-specific 5'-RACE product, is indicated by a bent arrow. For P*rps4*-477, products in lanes marked +T are migrating more slowly than products in lanes –T, which is typical for primary transcripts that are post-transcriptionally clipped at the 5' ends by a few nucleotides (compare Supplemental Figure 8). We therefore did not sequence cloned 5'-RACE products from RNA not treated with TAP.



Supplemental Figure 11. 5'-RACE analysis of *nad2e1_e2* transcripts.
5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products from transcript 5' ends are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for amplification is indicated below the gel image.
Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *nad2e1* upstream sequence that surrounds the identified TSS and putative TSS is shown, with the ATTA and CGTA tetranucleotides, which are frequently associated with mitochondrial promoters, printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type to correspond to transcript 5' ends mapping to the respective positions (row +T, clone numbers for products of 5'-RACE following TAP treatment; row –T, clone numbers determined without TAP treatment). Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter. The TSS, or putative TSS in the case of P*nad2e1*-114, is indicated by a bent arrow.
Among the highly diverse products mapping to the sequence surrounding position -114, no TAP-specific products were identified. Because the site shows promoter-like sequence elements, we have listed P*nad2e1*-114 as a putative promoter.
For P*nad2e1*-299 and P*nad2e1*-413, products in lanes marked +T are migrating more slowly than products in lanes –T, which is typical for primary transcripts that are post-transcriptionally clipped at the 5' ends by a few nucleotides (compare Supplemental Figure 9). We therefore did not sequence cloned 5'-RACE products from RNA not treated with TAP.



Pnad6-88 GAAGGGCTTG GAAGAAGAAA ATGAAATAGG AACAACCGCG CTGGT

2 10 3 1 / 16 (+T)
1 0 1 4 1 / 7 (-T)

Pnad6-143 ATTAAATGGA TGCATTGAGA TTC**CGTA**AGT AACTCAGTGA GTGCT

9 4 2 1 / 16 (+T)

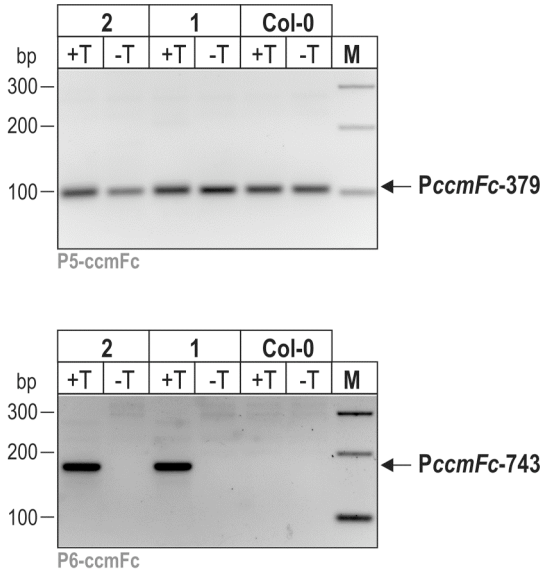
Pnad6-333 TGATTTACTG CGCGATGATT ATGTTGAAAA AGATCGAAGC CGCGG

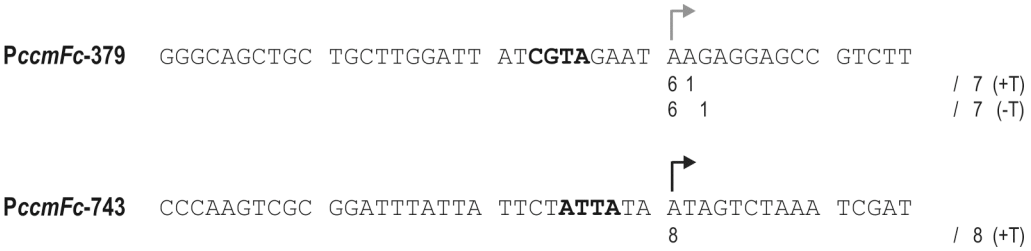
8 / 8 (+T)
0 7 / 7 (-T)

Pnad6-401 GATCATATTC ACGCGGGTAC **AGTAGTAGGT A**AACTTGAAG GAGAC

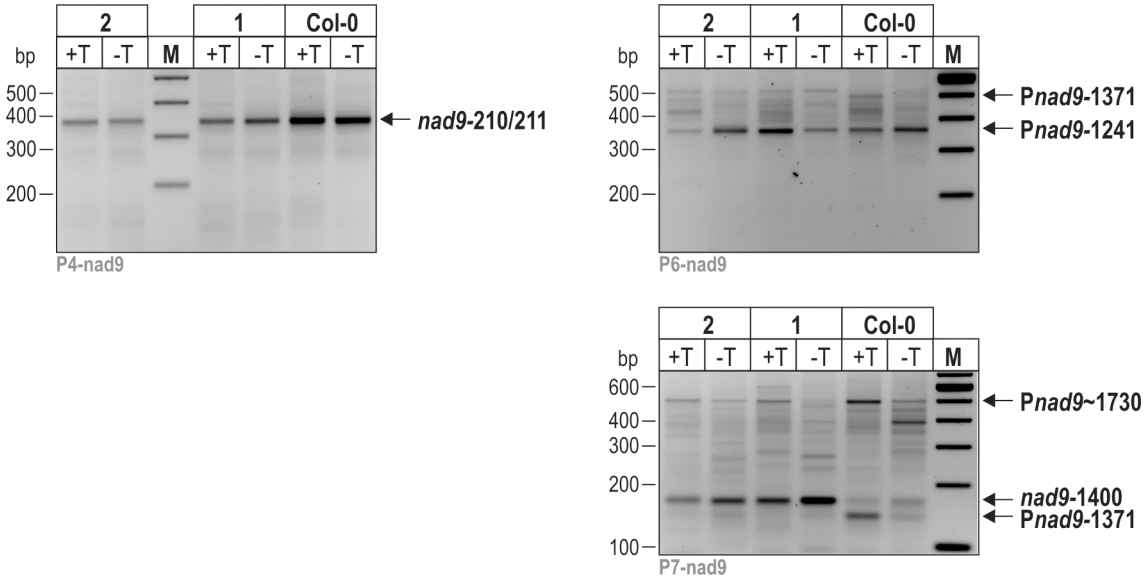
18 1 / 19 (+T)
2 6 / 8 (-T)

Supplemental Figure 12. 5'-RACE analysis of *nad6* transcripts.
5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products from transcript 5' ends are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for amplification is indicated below the gel image.
Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *nad6* upstream sequence that surrounds the identified TSS is shown, with the CGTA, AGTA and GGTA tetranucleotides, which are frequently associated with mitochondrial promoters, printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type (in *rpoTmp-2* for P*nad6*-333 and P*nad6*-143) to correspond to transcript 5' ends mapping to the respective positions (row +T, clone numbers for products of 5'-RACE following TAP treatment; row –T, clone numbers determined without TAP treatment). Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter. The TSS, which gave rise to a TAP-specific 5'-RACE product, is indicated by a bent arrow.
Products mapping to P*nad6*-333 were not detected in the wild type or *rpoTmp-1* but only in *rpoTmp-2*. The discrepancy between the two mutant lines, and also between *rpoTmp-2* and the wild type, could be due to very low abundance of the P*nad6*-333 5' end. The apparent difference at P*nad6*-333 between *rpoTmp-2* and the wild type is therefore considered insignificant.

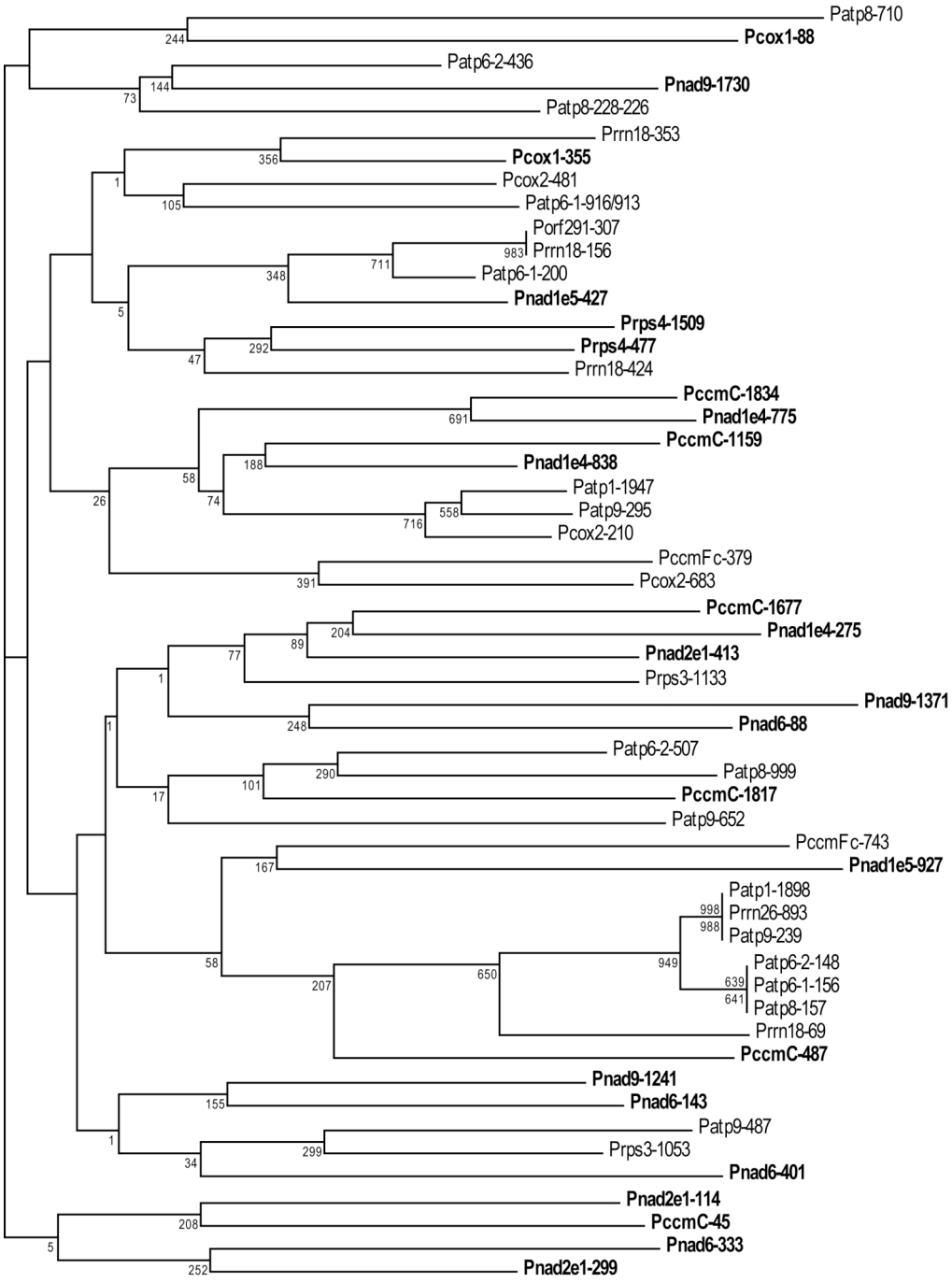




Supplemental Figure 13. 5'-RACE analysis of *ccmFc* transcripts.
5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products from transcript 5' ends are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for amplification is indicated below the gel image.
Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *ccmFc* upstream sequence that surrounds the identified TSS and putative TSS is shown, with the ATTA and CGTA tetranucleotides, which are frequently associated with mitochondrial promoters, printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type (in *rpoTmp-2* for P*ccmFc*-743) to correspond to transcript 5' ends mapping to the respective positions (row +T, clone numbers for products of 5'-RACE following TAP treatment; row –T, clone numbers determined without TAP treatment). Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter. The TSS, or putative TSS in the case of P*ccmFc*-379, is indicated by a bent arrow.
Among the products mapping to the sequence surrounding position -379, no TAP-specific products were identified. Because the site shows promoter-like sequence elements, we have listed P*ccmFc*-379 as a putative promoter.
TAP-specific products mapping to P*ccmFc*-743 were consistently detected in mutants but not in the wild type. Although this indicates that P*ccmFc*-743 is inactive in the wild type, we are unsure if this is truly the case, or if technical difficulties prevent its detection.



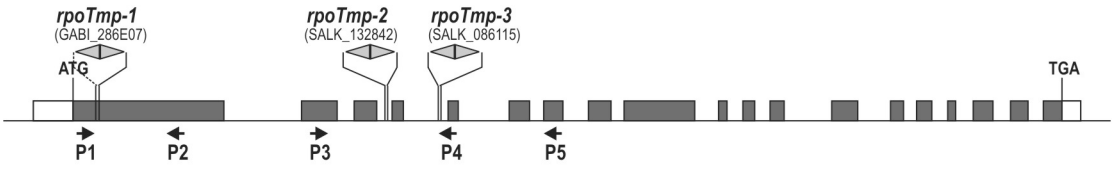
Supplemental Figure 14. 5'-RACE analysis of *nad9* transcripts.
5'-RACE was performed as detailed in Figure 6, and products obtained from TAP-treated (+T) and untreated RNA (-T) were separated on agarose gels alongside a molecular weight marker (M). Products are labelled with the name of the corresponding promoter or processing site as listed in Table I. The reverse primer used for 5'-RACE is indicated below the gel image.
Transcript 5' termini were analyzed through cloning and sequencing of marked 5'-RACE products. The part of the *nad9* upstream sequence that surrounds the identified TSS or putative TSS is shown; tetranucleotide elements that are frequently associated with mitochondrial promoters are printed bold. Numbers written below nucleotide positions indicate frequencies of clones that were found in the wild type to correspond to transcript 5' ends mapping to the respective positions when amplified from TAP-treated RNA. Numbers behind slashes indicate the numbers of clones that were sequenced in total for a promoter. The TSS and putative TSS are indicated by bent arrows. nad9 transcript 5' ends were not characterized in as much detail as those of other genes in this study; therefore only P*nad9*-1371 which shows a TAP-specific product in 5'-RACE gel patterns can be considered an experimentally verified promoter. The putative promoter P*nad9*~1730 has been inferred from the length of a TAP-specific 5'-RACE product corresponding to a promoter-like sequence in the *nad9* upstream region; no sequencing of 5'-RACE products was done. The grey filled triangle marks a nad9 5' end determined by Forner et al. (2007).



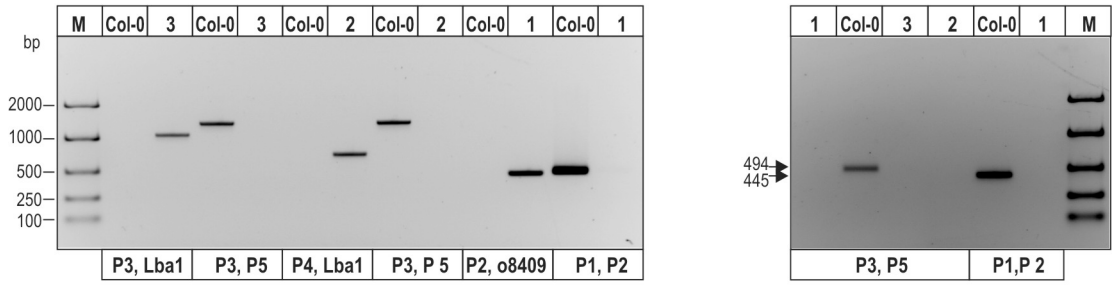
Supplemental Figure 15. Neighbor joining distance tree of mitochondrial promoter sequences in *Arabidopsis*.

The tree has been generated from promoter sequences aligned with respect to their TSS, as displayed in Figure 9, using ClustalW (Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22, 4673-4680). Bootstrap values are based on 1000 bootstrap repetitions. Sequences spanning 18 nucleotides, from position -14 to +4 around the TSS, have been analysed. This corresponds to the region required for promoter function in dicot plant mitochondria, as defined by Dombrowski et al. (2008). Promoters of genes showing reduced transcription in *rpoTmp* mutants are printed bold; they do not cluster as an obvious subgroup. Deriving a tree from an alignment of mitochondrial promoters generated by ClustalW, which ignored the positions of TSS, did similarly not result in clustering of *rpoTmp*-affected promoters (not shown).

A



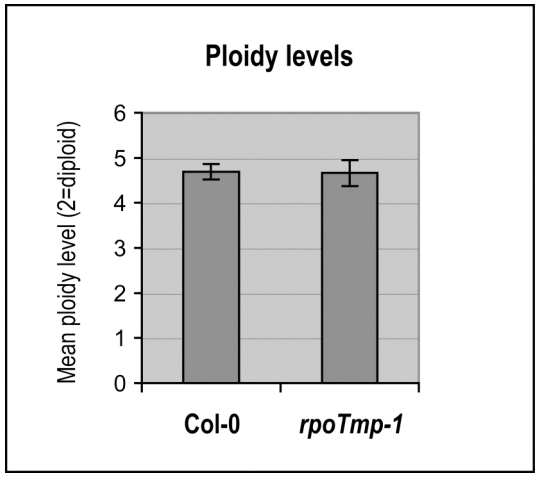
B



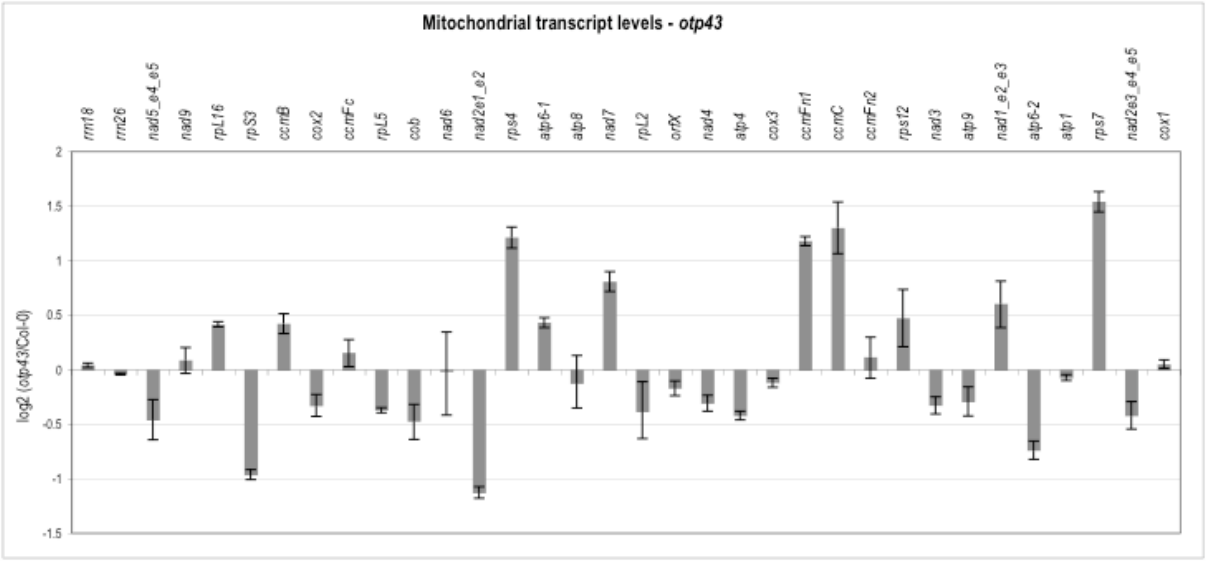
C

Primer	Sequence (5'→3')
P1	TAAACAAGCAATTTCGAGGTCAG
P2	AGCAGCTCTTTCCCATTCTTC
P3	GTTGATAAAGAAACAGAAGTTGTCAGCAG
P4	AGCGACATTCACATTTCCAAC
P5	ATTTAGTACTTCCAAGCGTATCCAGG
LBa1	TGGTTCACGTAGTGGGCCATCG
o8409	ATATTGACCATCATACTCATTGC

Supplemental Figure 16. Identification of *rpoTmp* mutant lines.
(A) Annealing sites of gene-specific primers (arrows) used to verify T-DNA insertions and the absence of a wild-type *RPOTmp* allele from lines *rpoTmp-1*, *rpoTmp-2* and *rpoTmp-3*.
(B) PCR analysis of the At5g15700 locus (left gel image) in *rpoTmp-1* (lanes labelled 1), *rpoTmp-2* (2) and *rpoTmp-3* (3) and the wild type (Col-0) using gene-specific primers and T-DNA left border primers LBa1 and o8409; primer combinations are indicated below the gel. The absence of a functional *RPOTmp* mRNA in mutants was confirmed by RT-PCR (right gel image).
(C) Genotyping primer sequences.



Supplemental Figure 17. Nuclear genome ploidy levels.
Nuclear DNA content in seven-day-old wild-type (Col-0) and *rpoTmp-1* seedlings was measured by flow cytometry as described by Zoschke et al. (2007). Error bars correspond to standard deviations (n = 4). This data was provided by E. Cincu, J. Fuchs and T. Börner.



Supplemental Figure 18. Mitochondrial steady-state transcript levels in the *Arabidopsis otp43* mutant.
Transcript levels determined by qRT-PCR as described in the main manuscript are depicted as the log2 ratio of transcript levels in *otp43* compared with levels in wild-type (Col-0) plants. The mean of *rrn18* and *rrn26* transcript levels was used for data normalization. Three technical replicates were averaged; standard errors are indicated.

Supplemental Table 1. Sequence motifs significantly overrepresented among promoters of genes showing reduced transcription in *rpoTmp*.

	Motif	N (subset1)	% (subset1)	N (subset2)	% (subset2)	Z
6-mers	AGAAGA	3	12	0	0	1.919558773
	AAAAGA	2	8	0	0	1.552169697
	AAATAT	2	8	0	0	1.552169697
	AAAGAT	2	8	0	0	1.552169697
	AATGAA	2	8	0	0	1.552169697
	AAGATA	2	8	0	0	1.552169697
	ATATTG	2	8	0	0	1.552169697
	ATAGAT	2	8	0	0	1.552169697
	ATTGCG	2	8	0	0	1.552169697
	ATGAAA	2	8	0	0	1.552169697
5-mers	AAGAT	4	16	0	0	2.238570973
	GAAAA	4	16	0	0	2.238570973
	AAAAG	3	12	0	0	1.919558773
	AAGAA	3	12	0	0	1.919558773
	TGAAA	3	12	0	0	1.919558773
	TGTAA	3	12	0	0	1.919558773
	TCAAG	3	12	0	0	1.919558773
	GAAGA	3	12	0	0	1.919558773
	GCGTA	3	12	0	0	1.919558773
	GTAAA	4	16	1	3.448275862	1.586668598
4-mers	GAAA	6	24	0	0	2.798213716
	AAAC	4	16	0	0	2.238570973
	AGAT	7	28	2	6.896551724	2.074874321
	ATCA	3	12	0	0	1.919558773
	TGTA	3	12	0	0	1.919558773
	GCGT	3	12	0	0	1.919558773
	CAAG	3	12	0	0	1.919558773
	TAGG	4	16	1	3.448275862	1.586668598
	AATG	2	8	0	0	1.552169697
	AACT	2	8	0	0	1.552169697

N equals the number of occurrences of a given motif among the subset of promoters of genes with reduced transcript levels in *rpoTmp* (subset 1) or among the subset of promoters of genes showing unaltered or increased transcription *in rpoTmp* (subset 2).
Sequences were searched for the presence of one or more of all 4096 6-mers, 1024 5-mers and 256 4-mers of the bases A, T, G and C. The proportions p_1 and p_2 of sequences which contain each motif in the two datasets were compared using the two-sample z-statistic

$$z = \frac{p_1 - p_2}{\sqrt{p_p(1 - p_p)(1/n_1 + 1/n_2)}}$$

where p_p is the pooled sample proportion. Statistical significance was assessed at the 99% level ($|z| > 2.58$).
Only the tetranucleotide GAAA (highlighted green) was significantly overrepresented in subset 1.

Supplemental Table 2. Primers used for quantitative RT-PCR analysis.
Primer sequences are given in 5'→3' orientation.

Transcript	Primer 1	Primer 2
nad5e4_e5	AACATTGCAAAGGCATAATGA	GTTCCTGCGTTTCGGATATG
nad9	GGATGACCCTCGAAACCATA	CACGCATTCGTGTACAAACC
rpl16	GAGCATTTGCCAAACTCACA	CGGACACTTTCATCGTGCTA
rps3	CCGATTTCGGTAAGACTTGG	AGCCGAAGGTGAGTCTCGTA
ccmB	TCTTGGAATCACATCCAGCA	CGAGACCGAAATTGGAAAAA
cox2	TGATGCTGTACCTGGTCGTT	TGGGGGATTAATTGATTGGA
ccmFc	CACATGGAGGAGTGTGCATC	GTGGGTCCATGTAAATGATCG
rpl5	AAGGGGTTCGACAGGAAAGT	CGTATTTCGACCGGAAAATC
cob	TGCCGGAATGGTATTTCCTA	GCCAAAAGCAACCAAAACAT
nad6	TATGCCGGAAAGGTACGAAG	GTGAGTGGGTCAGTCGTCCT
nad2e1_e2	GGATCCTCCCACACATGTTC	GCGAGCAGAAGCAAGGTTAT
rps4	ACCCATCACAGAGATGCACA	TCACACAAACCCTTCGATGA
atp6-1	TCTTTTGCGAGTCAATGCAC	TCTCGCGTATCTCACATTGC
atp8	CCGTCGACTTATTGGGAAAA	TTCCTTGGCCATGTACAACA
nad7	ACTGTCACTGCACAGCAAGC	CATTGCACAATGATCCGAAG
nad5e1_e2	TGGACCAAGCTACTTATGGATG	CCATGGATCTCATCGGAAAT
nad1e5	AGCCCGGGATCTTCTTGA	ACGGAGCTGCATCCCTACT
matR	AATTTTTGCGAGAGCTGGAA	TTGAACCCCGTCCTGTAGAC
nad1e4	AAAAGAGCAGACCCCATTGA	GGGAGCTGTATGAGCGGTAA
rpl2	CCGAAGACGGATCAAGGTAA	CGCAATTCATCACCATTTTG
orfX	GGGGTCTTTCTTTGGAAACC	TCTCCCTCATTCCACTCGTC
nad4	AATACCCATGTTTCCCGAAG	TGCTACCTCCAATTCCCTGT
atpB	GGATCAGCTTGCGAATTTGT	GCAAATTGCTTCCCCACTAA
nad4L	GGGGAATCCTCCTTAATAGACG	AACGAAAATGGCTAACCCAATA
cox3	CCGTAACTTGGGCTCATCAT	AAACCATGAAAGCCTGTTGC
ccmFn1	AGCTCTTGGCATTGCTTTGT	AGTGCCACAATCCCATTCAT
ccmC	AGCTACGCGCAAATTCTCAT	GCCGTGGCGATATAAACAAT
ccmFn2	CGTGTCGTTCGTAATGGAAA	TGATAAGCCCACCAACTTCC
rpsl2	AGCCAAAGTACGGTTGAGCA	TTTGGGTTTTTCTGCACCAT
nad3	CGAATGTGGTTTCGATCCTT	GCACCCCTTTTCCATTCATA
atp9	CATTCCCTCTGACGTCGAAT	TCGTCGATTCTTACCCTCGT
nad1e2_e3	TCTGCAGCTCAAATGGTCTC	ATTCAGCTTCCGCTTCTGG
atp6-2	TTCTGCCAGGAGTGCCTATT	TTCTCGCTGGAGGTTCCTAA
atp1	TCACTTCGACACGTCTTTGC	GGAATGGCCTTGAATCTTGA
rps7	CTCGAACTGAACGCGATGTA	AAGCTGCTTCAAGGATCCAA
nad2e2_e3_e4	TATTTGTTCTTCGCCGCTTT	CAAAGGAGAGGGGTATAGCAA
cox1	GTAGCTGCGGTGAAGTAGGC	CTGCCTGGATTCGGTATCAT
rrn18	CGTCACCTGGGTCAAAAACT	GCTTGAAAACCGAAGTGAGC
rrn26	GACGAGACTTTCGCCTTTTG	CTTGGAGCGAATTGGATGAT
nuclear 18S rrna (At3g41768)	AAACGGCTACCACATCCAAG	ACTCGAAAGAGCCCGGTATT
nuclear ACT2,8 (At1g49240, At3g18780)	GGTAACATTGTGCTCAGTGGTGG	AACGACCTTAATCTTCATGCTGC

Supplemental Table 3. Primers used for 5'-RACE.
The following oligonucleotides were used in 5'-RACE experiments for reverse transcription (application "A"), the first RACE PCR ("B") or subsequent nested PCRs ("C"). Primer sequences are given in 5'→3' orientation.

Gene	Primer	Sequence	Application	Detected 5' end
cox1	P2-cox1	TGTGCCCATCACTCCAG	A	
	P3-cox1	ACCGAAAATGAAATAGAGAGTCCCT	B	
	P4-cox1	TGTGGTTTGTGGAGAACAGCC	C	P*cox1*-88 *cox1*-241
	P5-cox1-b	ATCGTCCTACAAAAGATAATGCTCTCAC	C	P*cox1*-355
matR, *nad1e4*	P2-matR-a	GAGTGGCAGCCTTGACC	A	
	P3-matR	CGGAAATGCGATGTGTCTGG	B	
	P4-matR	TGGAAACTCGGGATCGTAAATG	C	P*matR*-14 P*matR*-58
	P3-nad1e4	CGCACCTCATTAAGATCATATTGG	B	
	P4-nad1e4	AAAAGAGCAGACCCCATTGAAG	C	P*nad1e4*-275
	P5-nad1e4	CTCTCGATGTGAGATCAGCAGC	C	P*nad1e4*-775 P*nad1e4*-838
	P6-nad1e4	CCCTTGCCCTTGAACGATAGA	C	P*nad1e4*-775 P*nad1e4*-838 *nad1e4*-892
nad1e5	P2-nad1e5	AAAAGACCAGAAACAGGGAC	A	
	P3-nad1e5	TTTCCGGCCAAGTCCCA	B	
	P5-nad1e5	GTCCAAGCGATAGCGTAAAAGC	C	P*nad1e5*-427
	P7-nad1e5	TCAGGTCTTGACCGGGTCC	C	P*nad1e5*-927
nad5e1_e2	P2-nad5e2	GAAATCCCAAGAGCTAATCC	A	
	P3-nad5e2	AAGACCTACTCCCTCCCATCCC	B	
	P4-nad5e2	CTATGCGGATCCTCGGACA	C	-
	P5-nad5e1	CGAAGAGAATGAAACGCACGTAG	C	-
nad2e1_e2	P2-nad2e1	AAGTAATCCAAGCCAACCC	A	
	P3-nad2e1	TCCAAGCCAACCCACATTACTG	B	
	P4-nad2e1	GTGGGAGGATCCGAACGA	C	P*nad2e1*-114
	P5-nad2e1	GAAAAGAAAGGGCGGAATAGCA	C	P*nad2e1*-297 P*nad2e1*-413/409
nad6	P2-nad6	CGCAATACTTCTTCGTGAA	A	
	P3-nad6	GATGGGAAACAAAACGGAATGT	B	
	P4-nad6	GACCAAAGCAGGGCTCGAC	C	P*nad6*-88 P*nad6*-143 *nad6*-179
	P7-nad6	AGCACTCACTGAGTTACTTACGGAATC	C	*nad6*-179
	P6-nad6	CACGGAGCGGTAGACTGAACAC	C	P*nad6*-333 P*nad6*-401
	P5-nad6	AAACCAACAATCTGCACGCTTAG	C	-
nad9	P2-nad9	AAAGCACAACAATTGAAATAGGT	A	
	P3-nad9	TGAAATAGGTAGTCCGTATTGGTATCAG	B	
	P4-nad9	CGTATTGGTATCAGATCTATTCCCATGT	C	*nad9*-211/210
	P6-nad9	TTTGCCAAACTCACACCATCC	C	*nad9*-1241
	P7-nad9	TTCGGATATAGCACGTCTCCCT	C	P*nad9*-1371 *nad9*-1400 P*nad9*~1730
rps4	P2-rps4-b	GTTTTGTTCCTCTGTGCATC	A	
	P3-rps4-b	TGCGGCGTTGTATTATTGTCAGT	B	
	P4-rps4-b	CAAGTTTGAAATCTTAATGGGGACA	C	*rps4*+2
	P5-rps4	CATGACTCGACCTTTAGTTTTGTGAC	C	P*rps4*-477
	P7-rps4	GAAAAGGTCGGGACAGTGGC	C	P*rps4*-1509
ccmC	P2-ccmC-a	AGCGAAGATAAAGGGGATG	A	
	P2-ccmC-a	AGCGAAGATAAAGGGGATG	A	

	P3-ccmC	TGTTAAGAACAACCAAGACCCAATG	B	
	P4-ccmC	AATGAGAATTTGCGCGTAGCTT	C	P*ccmC*-45
	P10-ccmC	CGGGAAGATCATAGCGAAGAAG	C	P*ccmC*-488/484
	P5-ccmC	TGCTGCAAAAGACTGGGATG	B	
	P6-ccmC	GGGATGGGAAACAAAACAGAATG	C	P*ccmC*-1159
	P7-ccmC	CAGAACAAAGCAGGTAGACCAAGC	C	*ccmC*-1634 P*ccmC*-1677 P*ccmC*-1811 P*ccmC*-1834
ccmFc	P2-ccmFc	TGAAGAACGGGAACGAA	A	
	P3-ccmFc	TCCACTTTTTGCTCCGTCCA	B	
	P4-ccmFc	CCATTTCAATAGTACGGGTGCTG	C	*ccmFc*-124
	P5-ccmFc	ACGAGTAGGTAGGCGGCATCT	C	P*ccmFc*-379
	P6-ccmFc	TAGGAATCTTAGGTTACCTCCAGCC	C	P*ccmFc*-743

Supplemental Table 4. Gene-specific primers used to amplify DNA probes for the detection of mitochondrial run-on transcripts.
Primer sequences are given in 5'→3' orientation.

Transcript	**Primer 1**	**Primer 2**
rrn26	TCACCGAAGCGACGAGACC	ACTCGCGACTATGGCAGGG
rps3	CTTTCTACGGCGGGGTCACT	CGTCATTTTCTATCAGGAACGCT
ccmFc	TTTATGGTCGTGCCTTGTGG	TTCGTCGTTGCTCATTCCC
nad6	TCGACTTTCATGCTGAAGCAAG	CATTTCGTCGGAATACATCCTG
nad2e1_e2	CTAATAACCTTGCTTCTGCTCGC	TGTCCGGTCGTACCCAAAC
	TTCCTAGCCGTCGCCTAGAGT	TCCAAGCCAACCCACATTACTG
rps4	TCAAAATGTGGCGGAGCTAGA	TCTCTTGTGCTCAGCGAAGGA
atp6-1	CAATTAATCCAATATGCGACGAATC	TGAAAGGCTTCCCCGGATAC
cox1	TTCCAGGAAAGCGCGAGAG	ATACCGAATCCAGGCAGAATGAG
rrn26	TCACCGAAGCGACGAGACC	ACTCGCGACTATGGCAGGG
rps3	CTTTCTACGGCGGGGTCACT	CGTCATTTTCTATCAGGAACGCT
ccmFc	TTTATGGTCGTGCCTTGTGG	TTCGTCGTTGCTCATTCCC