

Multiple promoters are a common feature of mitochondrial genes in Arabidopsis

Kristina Kühn, Andreas Weihe and Thomas Börner*

Institut für Biologie/Genetik, Humboldt-Universität zu Berlin, Chausseestr. 117, D-10115 Berlin, Germany

Received November 15, 2004; Revised and Accepted December 20, 2004

ABSTRACT

Mitochondrial genes in the plant *Arabidopsis thaliana* are transcribed by two phage-type RNA polymerases encoded in the nucleus. Little is known about cis-elements that are recognized by these enzymes and mediate the transcription of the Arabidopsis mitochondrial genome. Here, 30 transcription initiation sites of 12 mitochondrial genes and gene clusters have been determined using 5'-RACE and ribonuclease protection analysis of primary transcripts labelled *in vitro* by guanylyltransferase. A total of 9 out of 12 genes were found to possess multiple promoters, revealing for the first time that multiple promoters are a common feature of mitochondrial genes in a dicotyledonous plant. No differences in promoter utilization were observed between leaves and flowers, suggesting that promoter multiplicity reflects a relaxed promoter specificity rather than a regulatory role of promoter selection. Nearly half the identified transcription initiation sites displayed immediately upstream a CRTA core sequence, which was mostly seen within the previously described CRTAAGAGA promoter motif or a novel CGTATATAA promoter element. About as many promoters possessed an ATTA or RGTA core. Our data indicate that the majority of mitochondrial promoters in Arabidopsis deviate significantly from the nonanucleotide consensus derived earlier for dicot mitochondrial promoters.

INTRODUCTION

Plant mitochondrial genomes considerably vary in size but contain a fairly stable number of 50–60 genes (1–4). These may be dispersed or organized in gene clusters and predominantly encode components of the respiratory chain and of the translational apparatus. The machinery that transcribes these genes is encoded in the nucleus. While genes for mitochondrial RNA polymerases have been identified in a

range of photosynthetic organisms (5–10), cofactors mediating transcription initiation in plant mitochondria are to date unknown.

In contrast to metazoan mitochondria, which encode fewer, closely spaced genes and initiate transcription of the mitochondrial DNA (mtDNA) at a single unidirectional or bidirectional promoter on each strand, plant mitochondria use numerous transcription initiation sites [reviewed in (11)]. Mitochondrial promoters have been analysed in several plant species through identifying primary 5' termini of mitochondrial transcripts and aligning sequences surrounding transcription initiation sites, thereby revealing conserved promoter motifs (12). Moreover, sequence elements have been defined that are relevant for promoter function in *in vitro* transcription studies (13–18). Sequences of up to 25 nt around the transcription start site, which display the conserved motif YRTA (Y = T or C and R = A or G) immediately upstream of the initiating nucleotide, were found to be required to correctly and efficiently initiate transcription *in vitro* (19). In mitochondrial promoters of dicots, this core motif is embedded in an extended consensus of 9 nt, CRTAAGAGA, with the initiating nucleotide at the penultimate position (20). The majority of higher plant mitochondrial promoters exhibit an A/T-rich sequence element immediately upstream of the promoter core, which has been proven essential for the full function of several dicot and monocot mitochondrial promoters *in vitro* (13,17). Only a few transcription start sites in higher plant mitochondria coincide with sequences lacking any recognizable core motif. In contrast to mitochondrial promoters of *Oenothera berteriana* and potato that conform to the nonanucleotide consensus and are recognized by a pea *in vitro* transcription system, deviating mitochondrial promoters of both plants do not function in the heterologous system, which implies that specific mechanisms are involved in transcription initiation at alternative promoters (16). Moreover, studies on mitochondrial transcription start sites support the idea that the utilization of particular promoters requires a distinct nuclear background (21,22). Comparing transcription initiation sites in a broader range of plant species will thus be required to fully understand the mechanisms directing the transcription of plant mitochondrial genes. Extending the data set on dicot

*To whom correspondence should be addressed. Tel: +49 30 20938140; Fax: +49 30 20938141; Email: thomas.boerner@rz.hu-berlin.de

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use permissions, please contact journals.permissions@oupjournals.org.

mitochondrial promoters might moreover contribute to elucidating possible differences in transcriptional mechanisms and promoter usage between dicots and monocots. Such differences may be expected as the mitochondrial genome of dicots is transcribed by two phage-type RNA polymerases (5–7), while only one enzyme of this type is required in monocot mitochondria (8,9,23).

Availability of the complete sequence of the *Arabidopsis thaliana* mitochondrial genome (2) renders this plant an excellent object of comprehensive analyses of mitochondrial promoter architecture and distribution. Nonanucleotide motifs identified as elements of mitochondrial promoters in pea, *O. berteriana*, soybean and other dicots are seen upstream of coding regions of several mitochondrial genes in *Arabidopsis* (24,25), yet experimental evidence is limited to a single promoter (26). Here, we have analysed sequences at transcription start sites of selected mitochondrial genes of this plant, thereby revealing a high diversity of mitochondrial promoter sequences as well as a high frequency of multiple promoters of mitochondrial genes, which may be valid also for other dicotyledonous plants.

MATERIALS AND METHODS

RNA isolation

Total RNA was isolated from leaves and flowers of *Arabidopsis* plants as described previously (23).

5'-RACE analysis of RNA

Transcript 5' termini were determined employing a 5'-RACE technique described by Bensing *et al.* (27) with the following modifications. 5' triphosphates were converted to monophosphates by treating 5 µg RNA with 10 U of tobacco acid pyrophosphatase (TAP) (Epicentre Technologies) at 37°C for 1 h in the presence of 40 U of RNase inhibitor (Fermentas GmbH, Germany) in the appropriate buffer. Control reactions were set up without pyrophosphatase. The RNA was subsequently extracted with phenol/chloroform/isoamyl alcohol (25:24:1), precipitated from the aqueous phase by adding 3 vol of ethanol/3 M sodium acetate, pH 5.2 (30:1) and dissolved in water. The RNA was then supplemented with 10 pmol 5' RNA adapter A3 (28), and the ligation of transcripts to the adapter was performed at 37°C for 1 h with 50 U of T4 RNA ligase (Epicentre Technologies) in the presence of 1 mM ATP and 80 U of RNase inhibitor (Fermentas GmbH) in the appropriate buffer. Control reactions were set up without adding the adapter. Following the ligation, the RNA was extracted with phenol/chloroform/isoamyl alcohol (25:24:1), precipitated by adding 3 vol of ethanol/3 M sodium acetate, pH 5.2 (30:1), dissolved in water and then reverse-transcribed using gene-specific primers and Omniscript Reverse Transcriptase (Qiagen, Germany) according to the manufacturer's protocol. The products of reverse transcription were amplified in a first PCR step by using 1–3 µl of the RT reaction, 5 pmol of each adapter-specific forward primer P1a (5'-CGA ATT CCT GTA GAA CGA ACA CTA GAA G-3') and gene-specific reverse primer, 200 µM of each dNTP and 0.5 U of *Taq* DNA polymerase (Qiagen) in 25 µl of the appropriate buffer. Cycling conditions: 94°C for 1 min; 35 cycles of 95°C for 20 s, 58–62°C for 20 s, 72°C for 2 min; 72°C for 10 min. An aliquot

of 0.1–1 µl of the first PCR reaction was used as template for subsequent nested PCRs set up essentially as the first PCR in a volume of 50 µl with 10 pmol of each gene-specific and adapter-specific primer. Gene-specific primers were repeatedly placed upstream of identified transcriptional starts, until no 5'-RACE products reaching further upstream could be detected. PCR reactions were analysed on gels composed of 1% agarose and 2% Nusieve agarose (Biozym, Germany). Products of interest were excised, purified over QIAquick spin columns (Qiagen) and ligated into pDrive (Qiagen). Ligation products were transformed into *Escherichia coli* TOP10 (Invitrogen, Germany). Bacterial clones containing the plasmid insert were identified by colony PCR with vector-specific primers. Colony PCR reactions were set up and performed essentially as the first PCR of the 5'-RACE protocol, and PCR products were purified over QIAquick spin columns (Qiagen) and sequenced employing an ABI 377 automatic DNA Sequencer (Applied Biosystems).

A list of all gene-specific primers used in 5'-RACE reactions is provided in the Supplementary Material.

Analysis of *in vitro*-cappable transcripts

In vitro capping reactions were set up in a volume of 100 µl with 100 µg of total RNA isolated from flowers and 5 U guanylyltransferase (Ambion) in the appropriate buffer in the presence of 650 µM *S*-adenosyl methionine, 12.5 U RNase inhibitor (Fermentas GmbH), 100 µCi [α -³²P]GTP (3000 Ci/mmol) and incubated at 37°C for 75 min. After 30 min, another 7.5 U of guanylyltransferase were added. The RNA was purified with 2 vol of water, 4 vol of TRIzol (Invitrogen) and 0.8 vol of chloroform, and precipitated from the aqueous phase by adding 0.8 vol of isopropanol and washed twice with 70% ethanol. Transcripts were then dissolved in 30 µl hybridization buffer (Roche Applied Science, Germany) together with 0.5 µg of complementary riboprobe and subjected to ribonuclease protection using the RNase Protection Kit (Roche Applied Science) according to the manufacturer's protocol, except that hybridizations were carried out overnight at 45°C (*rrn 18*, *rrn26*) or 65°C (all other genes). Protected transcripts were separated in 5% polyacrylamide gels.

Preparation of riboprobes

Fragments of mitochondrial genes were amplified from total *Arabidopsis* DNA with primer pairs Cap-*rrn26*-F (5'-AAA GGC GTT ATT GCT GTG CT-3') and Cap-*rrn26*-R (5'-TTT TCA ACT CGT AAA GGC AAA GA-3'), Cap-*rrn18*-F (5'-GAG ACC GAT CCA GGA ACC CTA C-3') and P4-*rrn18*-b (5'-TCG TGA ACC GGG CGT ACT AC-3'), Cap-*atp1*-F (5'-GCT AGG CTG GCA CTT AGG A-3') and Cap-*atp1*-R (5'-TGA TAG GTT TAT TTC GCA CTT TAG-3'), Cap-*atp6-1*-F (5'-CAC CGC ACG AGT CAG ACC T-3') and Cap-*atp6-1*-R (5'-TCC AGA CAG CTT CAC TCC GTC-3'), Cap-*atp6-2*-Fb (5'-CGT GTT CCG TCG ATC AC-3') and Cap-*atp6-2*-Rb (5'-CTT ACG TCA AGC CTC TAG GAG T-3'), Cap-*atp9*-Fb (5'-TTG GGA TAA GTG AAA TCG TAT-3'), and Cap-*atp9*-Rb (5'-CGA CAA AGA GAA GTA CAA GC-3'), Cap-*cox2*-F (5'-TGC CTT GCC TTA CCA CAC C-3') and P3-*cox-d* (5'-TTC TTC TTC TTC TTA CAA TAT TTT GAG TTA GAT G-3'), and ligated into pDrive

(Qiagen) in the appropriate orientation. Riboprobes were generated through *in vitro* transcription of antisense strands of the cloned fragments and a subsequent DNase digest using the MAXIScript kit (Ambion). Transcripts were then extracted with phenol/chloroform/isoamyl alcohol (25:24:1), precipitated from the aqueous phase by adding 3 vol of ethanol/3 M sodium acetate, pH 5.2 (30:1) and dissolved in water.

RESULTS

Identification of transcription initiation sites by 5'-RACE

To learn about promoter specificities of the mitochondrial transcription machinery in Arabidopsis, we experimentally determined mitochondrial transcription initiation sites using a 5'-RACE technique first described by Bensing *et al.* (27), which since has been applied to define primary transcript 5' termini in different groups of bacteria (28,29). In bacteria as in mitochondria, primary transcript 5' ends carry triphosphates, while processed transcripts have monophosphates at their 5' ends. Only the latter are a substrate to RNA ligase, and are in the experimental procedure selectively ligated to an RNA oligonucleotide, to which a forward primer will anneal in a subsequent 5'-RACE step. Primary 5' termini may be ligated only after removal of a 5' pyrophosphate through TAP. Consequently, 5'-RACE will yield products from TAP-treated RNA for both primary and processed transcripts, whereas without exposure to TAP, products resulting from primary transcript termini will be significantly reduced or absent. Comparison of 5'-RACE products obtained from TAP-treated and non-treated RNA (lanes +T and -T in Figures 1 and 3) would thus identify primary transcripts.

The only transcription start site that has so far been experimentally defined in Arabidopsis mitochondria is located upstream of the *rrn18* gene and coincides with a conserved nonanucleotide sequence motif (26). Including *rrn18* as a control, we first investigated the *rrn18*, *cox2* and *atp9* genes, for which promoters have been characterized in several dicots (16,26,30,31). To look for possible tissue-specific variations in promoter utilization, analysis of transcript 5' ends was performed on RNA isolated from leaves and from flowers of Arabidopsis plants.

An *rrn18* transcription start site was identified that mapped to position -156 with respect to the mature 18S rRNA 5' end (Figure 1). The initiating nucleotide was found to be part of the motif CGTATATAA (initiating nucleotide underlined), which has not yet been described as a promoter motif. The previously determined primary end of this transcript at position -69 appeared to result from processing rather than transcription initiation, as it gave rise to a PCR product that was not enhanced after TAP treatment of transcripts, compared with the control. In the following, transcriptional starts and their surrounding sequences, which in plant mitochondria encompass the promoter (13-15,17), will be specified with the letter P ('promoter'), followed by the gene name and position of the initiating nucleotide with respect to the start of the coding sequence or the mature RNA, e.g. *Prrn18*-156.

In the *cox2* upstream region, two transcriptional starts were detected by 5'-RACE. Although TAP-treated and non-treated

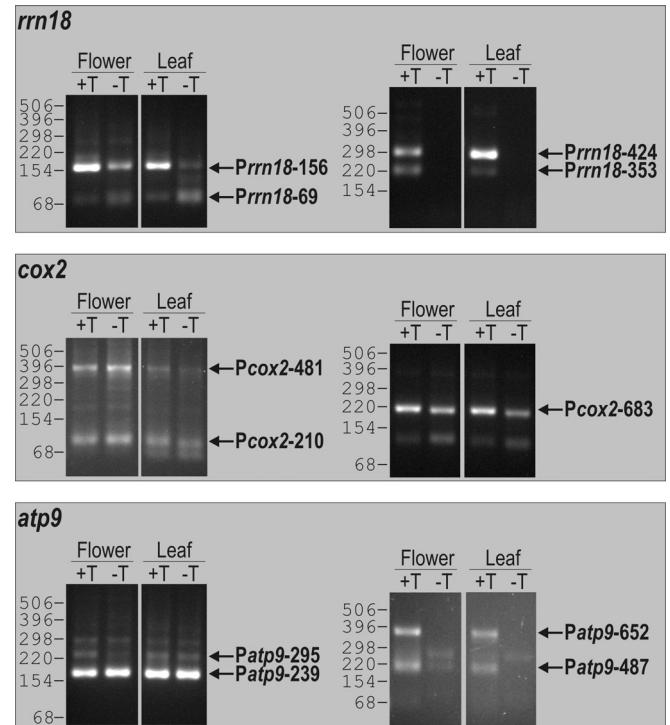


Figure 1. 5'-RACE analysis of the mitochondrial *rrn18*, *cox2* and *atp9* transcripts. 5'-RACE products were separated on agarose gels alongside molecular weight markers; sizes are given in base pairs (marker lane not displayed). Products corresponding to primary transcript 5' ends are indicated by arrows and labelled with the name of the corresponding promoter as listed in Table 1. In most cases, TAP-specific products were obtained, which correspond to 5' termini resulting from transcription initiation (lanes designated +T; see also Figure 2 and Table 1). Initiation at *Patp9*-239 and *Patp9*-295 required confirmation through ribonuclease protection analysis of cap-labelled transcripts (compare Figures 2 and 3). Control experiments were performed by 5'-RACE from RNA mock-treated in TAP buffer without TAP (lanes -T).

RNAs lead to similar band patterns (Figure 1), extensive sequencing of cloned PCR products revealed that among products of similar lengths, particular 5' ends were significantly enriched or exclusively present in the TAP-treated sample (Table 1 and Figure 2) and are thus bona fide primary ends. While a nonanucleotide sequence at *Pcox2*-210 matched the motif found at *Prrn18*-156 exactly, only limited similarity to any known plant mitochondrial promoter was seen for *Pcox2*-481.

5'-RACE analysis of *atp9* transcripts identified one major and one minor 5' end, the latter mapping to position -295 within the motif CGTATATAA and the former mapping to position -239 within the sequence CATAAGAGA which, based on sequence comparisons with the experimentally defined *atp9* promoter in pea mitochondria, had been predicted to function as a promoter upstream of *atp9* and several other genes in Arabidopsis mitochondria (24). However, PCR products resulting from either 5' end were equally abundant after amplification from TAP-treated and non-treated RNA (Figure 1), and transcripts were found to start with the nucleotides underlined in Figure 2, regardless of the application of TAP (Figure 2). Thus, both 5' ends were carrying 5' monophosphates and therefore resulting from processing events, despite the perfect nonanucleotide motifs. Alternatively, mixed populations of primary and processed

Table 1. Transcription initiation sites detected by 5'-RACE and *in vitro* capping

Gene	Promoter	Sequence	No. of clones (+TAP)	No. of clones (-TAP)	<i>In vitro</i> cappable
<i>rrn18</i>	<i>Prrn18-156</i>	TAGAATAATA CGTA TATAA TCAGAA	20/23	4/7	+
<i>orf291</i>	<i>Porf291-307</i>	TGGAATAATA CGTA TATAA CAGAT	7/10	n.d.	n.d.
<i>atp6-1</i>	<i>Patp6-1-200</i>	GCCAATAATA CGTA TATAA GAAGAG	3/14	n.d.	+
<i>atp9</i>	<i>Patp9-295</i>	CTGGTGCTCT CGTA TATAA GAGAAG	8/8	10/11	+
<i>atp1</i>	<i>Patp1-1947</i>	CTGGTGGTAT CGTA TATAA GAGAGA	8/13	10/15	+
<i>cox2</i>	<i>Pcox2-210</i>	ATGTTGGTTT CGTA TATAA GAAGAC	5/39	0/31	+
<i>tRNA-fMet</i>	<i>PrrnM-98^b</i>	TTTGAATAT CGTA AGAGAAGAAAG	12/12	n.d.	n.d.
<i>rrn26</i>	<i>Prrn26-893^b</i>	CTATCAATTT CATA AGAGAAGAAAAG	12/13	0/23	+
<i>atp1</i>	<i>Patp1-1898^b</i>	CTATCAATTT CATA AGAGAAGAAAAG	13/13	n.d.	+
<i>atp9</i>	<i>Patp9-239^b</i>	CTATCAATTT CATA AGAGAAGACGA	21/21	12/13	+
<i>atp6-1</i>	<i>Patp6-1-156^b</i>	CTATCAATCT CATA AGAGAAGAAAAT	5/14	n.d.	+
<i>atp6-2</i>	<i>Patp6-2-148^b</i>	CTATCAATCT CATA AGAGAAGAAAAT	7/13	n.d.	+
<i>atp8</i>	<i>Patp8-157^b</i>	CTATCAATCT CATA AGAGAAGAAAAT	14/22	n.d.	n.d.
<i>rrn18</i>	<i>Prrn18-69^a</i>	AGTGAATTC AATA AGAGAAGAAAAG	n.d.	6/8	+
<i>atp8</i>	<i>Patp8-999</i>	ATAAAATTA AATA AGAGCAAAAAT	9/12	n.d.	n.d.
<i>atp8</i>	<i>Patp8-228/226^c</i>	CATACCATA CATA TATA GAATCGA	1/28, 6/28	0/14, 0/14	n.d.
<i>rrn18</i>	<i>Prrn18-353</i>	TACTTTCCATCT TATA TATAAATGAA	10/12	n.d.	n.d.
<i>atp6-1</i>	<i>Patp6-1-916/913^c</i>	AGCCCTTTAT TATTA TATAA TAAAGC	2/23, 11/23	0/23, 1/23	n.d.
<i>cox1</i>	<i>Pcox1-355</i>	AATTTATTC AATTA TATAA TAATAA	18/23	19/30	n.d.
<i>cox2</i>	<i>Pcox2-481</i>	ATGAATATTC ATTAG ATAATAGATT	13/43	1/34	n.d.
<i>rps3</i>	<i>Prps3-1133</i>	TAGAAAAAAT TATTA GTAATACGTA	6/26	0/15	n.d.
<i>rrn18</i>	<i>Prrn18-424</i>	TCAAATCCTC GGTA TATAA AGAGAA	9/10	n.d.	n.d.
<i>cox2</i>	<i>Pcox2-683</i>	GACACGTAA GGTA AAAATAAGAACTCT	6/12	0/8	n.d.
<i>rps3</i>	<i>Prps3-1053</i>	TTTTTTATTT GGTAGGTA ACATCGC	12/14	n.d.	n.d.
<i>atp9</i>	<i>Patp9-487</i>	ATGTCATT GGTA TATGTGATACAAG	13/14	n.d.	n.d.
<i>atp9</i>	<i>Patp9-652</i>	AGAAGATTGA AGTA AGGAGCAGGTT	7/16	0/29	n.d.
<i>atp6-2</i>	<i>Patp6-2-436</i>	TCTTGAATTA AGTA TATA GAAAAGA	5/20	n.d.	n.d.
<i>atp6-2</i>	<i>Patp6-2-507</i>	GATAAATTA AGTA TAGTAATAAGAA	9/12	n.d.	n.d.
<i>atp8</i>	<i>Patp8-710</i>	ATCGGAGCTGCCAATA AGCTAATCC	4/12	0/13	n.d.
<i>tRNA-fMet</i>	<i>PrrnM-574/573^c</i>	CTAATTTATATAAAAA AGACCCGGGA	9/18, 9/18	n.d.	n.d.

n.d., not determined.

Initiating nucleotides are underlined; repeatedly observed promoter cores are written bold and the frequent TATATA(A) motif is highlighted. The number of clones that were sequenced for each promoter is given together with the frequency of the respective primary transcript 5' end as determined from TAP-treated flower RNA, and for selected promoters from flower RNA not exposed to TAP.

^aConsistent with primer extension results in Giese *et al.* (26).

^bConsistent with previous predictions of Arabidopsis mitochondrial promoters (24).

^cTranscription initiation was found to occur at two different nucleotides in one promoter region; frequencies of transcript 5' termini are given first for the upstream nucleotide.

transcript 5' ends might have been present, starting with identical nucleotides but carrying either tri- or monophosphates. In order to unambiguously determine whether transcription initiated at positions -239 and -295, the species of *atp9* transcript ends mapping to these positions were tested for the presence of *in vitro*-cappable 5' termini (see below).

With the aim of identifying additional promoters of the genes *rrn18*, *cox2* and *atp9*, we analysed their 5' regions by 5'-RACE through repeatedly placing reverse primers upstream of identified transcriptional starts, until no further transcript ends could be detected. All three genes were found to possess additional upstream promoters (right panels in Figure 1), of which none matched any known plant mitochondrial promoter sequence (Table 1).

We screened the Arabidopsis mitochondrial genome for additional occurrences of sequence motifs coinciding with experimentally defined transcriptional starts. Of the genes displaying a promoter motif in their upstream regions, we selected *rrn26*, *atp1*, *atp6-1*, *atp6-2* and *atp8* for an experimental verification of their predicted transcriptional starts. Notably, the *atp1* and *atp6-1* 5' regions like *atp9* displayed the motif CGTATATAA ~50 bp upstream of the hypothetical CATAAGAGA promoter sequence. We also included

tRNA-fMet with the predicted promoter motif CGTAAGAGA (24), which had been found to be an element of the *atp9* promoter in pea and soybean (16,30). Of those genes not possessing any conserved promoter motif upstream of their coding sequence, *rps3* and *cox1* were selected for transcript 5' end mapping.

Table 1 provides a summary of determined transcription start sites and their surrounding sequences. Besides 5' ends that were unambiguously identified as primary ends in the 5'-RACE, such as those mapping to *Prrn26-893* and *Prps3-1053*, we detected various 5' termini for which 5'-RACE products were not enhanced following TAP treatment of RNAs but which, as described above for *Pcox2-210*, nevertheless coincided with genomic sequences exhibiting strong similarity to bona fide promoters. For these transcripts, we again compared the pools of 5' termini cloned from +TAP and from -TAP samples (fourth and fifth column in Table 1). Mostly, the -TAP pool contained slightly shorter transcripts than the +TAP pool (for examples, compare sizes of +TAP and -TAP 5'-RACE signals obtained for *Patp6-1-916/913* and *Pcox2-683* in Figures 3 and 1, respectively), and particularly was deprived of longer transcript species that at their 5' extremities carried A or G nucleotides (see Table 1). These

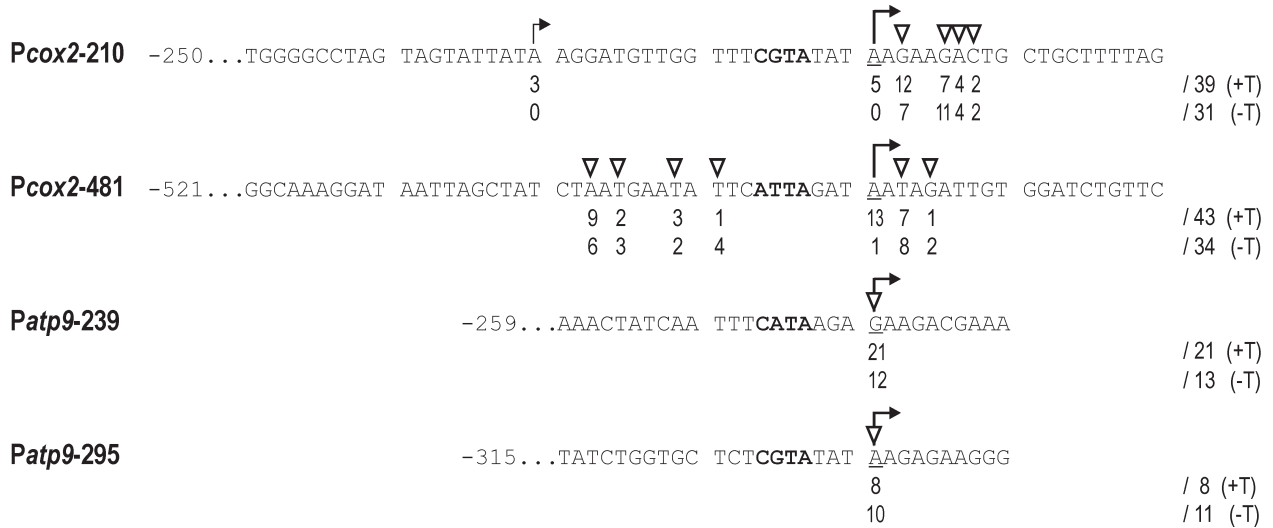


Figure 2. Transcript 5' termini detected through cloning and sequencing of 5'-RACE products designated Pcox2-210, Pcox2-481, Patp9-239 and Patp9-295 in Figure 1. Parts of the *cox2* and *atp9* upstream sequences that surround the four transcription initiation sites are shown; numbers preceding the sequences are the positions of the first displayed nucleotide with respect to the translational start. Numbers written below nucleotide positions indicate frequencies of clones that were found 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). Only 5' ends detected more than twice are marked. Numbers behind slashes indicate the numbers of clones that were sequenced in total for each promoter. Nucleotides corresponding to 5' ends that most likely result from processing events are indicated by open triangles. Transcription initiation sites, which gave rise to TAP-specific 5'-RACE products, are indicated by bent arrows. Processing sites and initiation sites were appointed as detailed in the text. Upstream of Pcox2-210, a small bent arrow marks position -231, which might be a transcription initiation site but did not yield a distinct band in 5'-RACE experiments.

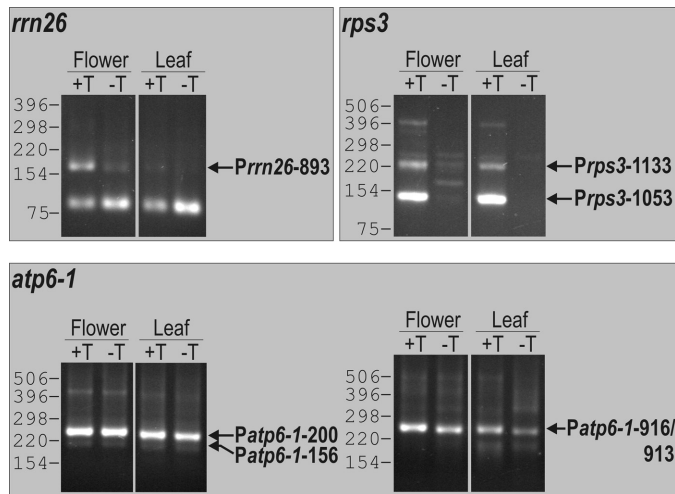


Figure 3. 5'-RACE analysis of the mitochondrial *rrn18*, *rps3* and *atp6-1* transcripts. Amplified products were separated on agarose gels alongside molecular weight markers; sizes are given in nucleotides (marker lane not displayed). TAP-specific products (lane +T) that correspond to primary transcript 5' ends are indicated by arrows and labelled with the name of the respective promoter as listed in Table 1. Initiation at Patp6-1-156 and Patp6-1-200 required confirmation through ribonuclease protection analysis of cap-labelled transcripts (compare Figure 4). Control experiments were performed by 5'-RACE from RNA mock-treated in TAP buffer without TAP (lanes -T).

transcripts specific to the +TAP pool are most likely resulting from transcription initiation. Within three promoter regions located upstream of the *tRNA-fMet*, *atp6-1* and *atp8* genes, transcription was found to initiate at two different nucleotide positions. From 5'-RACE results it is likely that multiple initiations also occur around Prps3-1133 (data not shown).

Multiple promoters were detected for all investigated genes except *rrn26*, *cox1* and *orf291*. Owing to partly identical upstream and coding sequences of *cox2* and *orf291*, the transcriptional start site preceding *orf291* was fortuitously found using primers annealing to the *cox2* upstream region.

To analyse possible differences in promoter utilization between Arabidopsis leaves and flowers, TAP-specific 5'-RACE signals (lanes +T in Figures 1 and 3) that had been obtained from leaf and from flower RNA for a distinct gene were compared. No primary transcript 5' end was detected that was exclusively present in leaves or in flowers, indicating that transcription is initiated at identical sites in both tissues. An occasional enhancement of 5'-RACE signals from flower RNA can be attributed to the level of mitochondrial activity being generally higher in flowers than in green tissues (32,33).

Identification of transcription initiation sites by *in vitro* capping

As already observed for *atp9*, 5'-RACE analyses of those *atp1*, *atp6-1*, *atp6-2* and *atp8* transcript 5' termini mapping to the motifs CATAAGAGA and CGTATATAA did not support transcription initiation at these sequences (Table 1 and Figure 3). The perfect nonanucleotide motifs found at these sites prompted us to examine the corresponding 5' ends by an independent technique. As a method specifically detecting primary 5' ends, we employed ribonuclease protection of *in vitro*-capped transcripts to analyse the respective 5' termini of the *atp9* and *atp6-1* mRNAs. This method takes advantage of organellar transcripts being, unlike nuclear mRNAs, not capped at their 5' ends *in vivo*. Mitochondrial primary transcripts, which carry 5' triphosphates, are thus representing guanylyltransferase (capping enzyme) substrates

and can be 5' cap-labelled with the GMP moiety of [α - 32 P]GTP *in vitro*. Total Arabidopsis RNA was capped and then subjected to ribonuclease protection using RNA probes complementary to the genomic regions containing putative promoters. We included the *rrn26* primary transcript as a positive control in our capping study, since its 5' end had been established by 5'-RACE to map to a promoter that is identical to the sequence surrounding the predicted transcriptional start *Patp1*-1947, and moreover is highly similar to the hypothetical promoters *Patp6-1*-156, *Patp6-2*-148, *Patp8*-157 and *Patp9*-239. Additionally, we tested the *rrn18* transcript 5' ends coinciding with positions -69 and -156 for their ability to be capped *in vitro*.

Figure 4A shows the protected cap-labelled RNAs corresponding to the transcription initiation sites *Prrn26*-893, *Prrn18*-69 and *Prrn18*-156, and to the tandem promoters *Patp6-1*-156 and *Patp6-1*-200, and *Patp9*-239 and *Patp9*-295. The sizes of protected RNA fragments are in accordance with the expected lengths of transcript 5' segments annealing to RNA probes, as illustrated in Figure 4B. Most notably, *Prrn18*-69 as well as *Patp6-1*-156, *Patp6-1*-200, *Patp9*-239 and *Patp9*-295, which through 5'-RACE could not be confirmed as transcription initiation sites, were found to coincide with *in vitro*-cappable and thus primary RNA 5' termini. Capping of the *atp9* mRNA mapping to position -295 only yielded a very faint signal, which may be either because of rare utilization of *Patp9*-295 as a promoter, or due to rapid *in vivo* processing of this primary message. Through *in vitro* capping, we moreover verified transcription initiation at *Patp1*-1898, *Patp1*-1947 and *Patp6-2*-148 (data not shown). From our analyses of *in vitro*-cappable *atp1*, *atp6-1*, *atp6-2* and *atp9* mRNAs, we inferred that the transcript 5' terminus coinciding with *Patp8*-157 is derived from transcription initiation.

Mitochondrial promoter architecture in Arabidopsis

Table 1 aligns Arabidopsis mitochondrial promoter sequences with respect to experimentally defined transcription start sites and places promoters with similar core sequences in adjacent rows. At positions -7 to -4 with respect to the transcriptional start, the majority of promoters display the previously described core element CRTA (12), which here is almost always seen as part of the nonanucleotide motifs CGTATATAA or CATAAGAGA, or the sequences ATTA, AGTA, GGTA or AATA. Only in a few promoters is the distance between core element and start site altered by 1 bp.

All primary transcripts characterized in this study originate from transcription initiation at either an A or a G nucleotide (21 and 12 out of 33 start sites, respectively). When comparing nucleotide frequencies within promoters, it appears that A and G start sites favour distinct nucleotides at adjoining positions. For example, while a G as initiating nucleotide is nearly always preceded by an A, initiation at an A essentially requires a T at position -1. Owing to these constraints on nucleotide frequencies particularly at positions around transcription initiation sites, we realigned promoter sequences in two subsets. The two alignments of promoters driving transcription from an A or a G are illustrated in Figure 5 as sequence logos. Among promoters directing initiation at an A, the sequence element TATATAA seems to be fairly frequent.

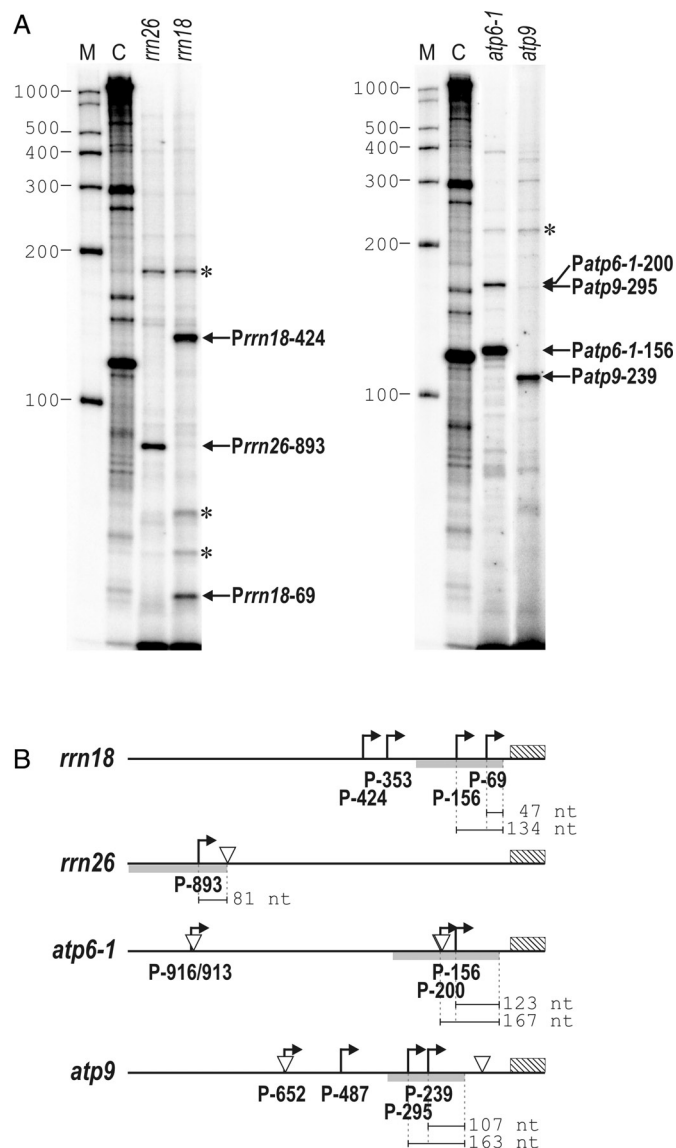


Figure 4. (A) Detection of selected *rrn18*, *rrn26*, *atp6-1* and *atp9* primary transcript 5' ends by ribonuclease protection of cap-labelled RNA. Protected RNA fragments were separated in polyacrylamide gels alongside a molecular weight marker (lane M); sizes are given in nucleotides. Lane C shows total capped RNA prior to ribonuclease protection. Lanes designated *rrn18*, *rrn26*, *atp6-1* and *atp9* show ribonuclease protection results obtained with riboprobes complementary to the *rrn18*, *rrn26*, *atp6-1* and *atp9* upstream regions as detailed in (B). Specific protected fragments, which correspond to primary transcripts, are indicated by arrows and labelled with the respective promoter name. Asterisks mark fairly strong signals that were considered non-specific, as they were seen with different riboprobes. (B) Diagram of the *rrn18*, *rrn26*, *atp6-1* and *atp9* 5'-UTR. Promoters identified in 5'-RACE and capping analyses are indicated by bent arrows; open triangles mark processing sites identified by 5'-RACE. The beginning of the mature rRNA or protein-coding sequence is indicated by a hatched bar. Grey bars are drawn below sequences complementary to the riboprobes that were annealed to cap-labelled primary transcripts in ribonuclease protection assays. The sizes of expected protected RNA fragments are given in nucleotides.

Promoters having a G nucleotide at position +1 mostly conform to the consensus CRTAAGAGA (R = A or G) that has been suggested previously for dicot mitochondrial promoters (20).

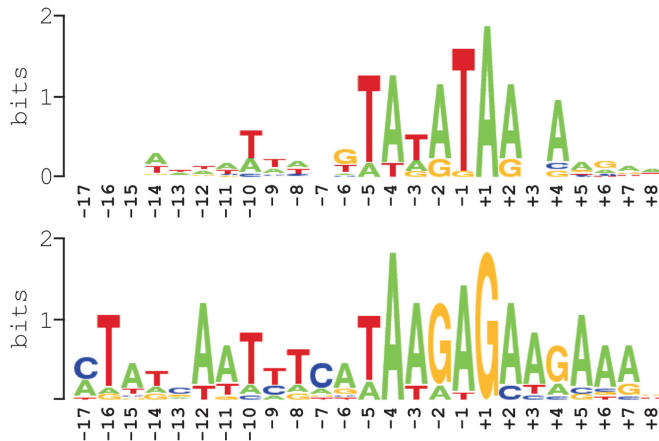


Figure 5. Summary of nucleotide sequences around experimentally defined transcription initiation sites in *Arabidopsis* mitochondria, as displayed in Table 1. Two sequence logos are shown that were generated using WebLogo [http://weblogo.berkeley.edu/logo.cgi, (45,46)] from an alignment of 20 promoter sequences activating transcription initiation at an adenine nucleotide (upper sequence logo) and from an alignment of 11 sequences supporting initiation at a guanine nucleotide (lower sequence logo). Position +1 corresponds to the transcriptional start.

DISCUSSION

Identification of transcription initiation sites by 5'-RACE and *in vitro* capping

Our current knowledge of *Arabidopsis* mitochondrial promoters is largely based on computational predictions performed on the mitochondrial genome sequence of this plant. While a set of 29 hypothetical mitochondrial promoters has emerged from searching the *Arabidopsis* mtDNA for motifs known to drive transcription initiation in other dicotyledonous plants (24,25), experimental evidence is limited to a single promoter in *Arabidopsis* mitochondria (26). Here, we have analysed the architecture and distribution of promoters in the upstream regions of 12 mitochondrial genes in *Arabidopsis* through mapping transcription start sites by a 5'-RACE technique that has previously been employed successfully for the analysis of bacterial transcript 5' termini (27–29).

The method proved a valid tool for mitochondrial 5' end detection and for 19 out of 30 mapped promoters clearly discriminated between primary and processed transcript ends. However, for selected 5' termini that were later identified through analysis of *in vitro*-capped transcripts to result from transcription initiation, 5'-RACE results did not support transcription initiation at the respective promoters. One example is *Prrn18-69*, for which the corresponding 5'-RACE product was significantly more abundant after amplification from transcripts not treated with TAP. Most likely, the misleading PCR result is due to a highly abundant primary transcript derived from the upstream promoter *Prrn18-156*, which is favoured in the 5'-RACE from TAP-treated transcripts, thereby outcompeting the 5'-RACE product generated from the much less abundant downstream 5' terminus (compare *rrn18* 5'-RACE products in Figure 1 and *in vitro*-capped transcripts in Figure 4).

For a group of primary transcripts of the *atp1*, *atp6-1*, *atp6-2*, *atp8* and *atp9* genes mapping to CGTATATAA or

CATAAGAGA motifs, 5'-RACE failed to distinguish primary from 5'-processed transcripts, owing to both types of transcripts mapping to the same initiating nucleotide. The processed transcripts could be derived from a modification of primary transcripts by a phosphatase or pyrophosphatase, or from endonucleolytic cleavage of transcripts initiated at upstream promoters.

Promoter architecture

Nearly 50% of the transcriptional starts identified in this study are located within sequences that conform to the nonanucleotide consensus CRTAAGAGA previously suggested for dicot mitochondrial promoters (20) or the motif CGTATATAA defined in this work (Table 1). About as many promoters deviate to a varying extent in the sequence surrounding the transcriptional start and in place of a CRTA core display ATTA or RGTA tetranucleotides that have emerged from this study as frequent promoter elements. A total of 20 out of 30 promoters support initiation at an adenine nucleotide, which complements previous reports on transcription initiation mostly at guanine nucleotides in dicot mitochondria (12). While purines appear obligatory at positions +1 and +2 with respect to the transcriptional start, they are moderately frequent at positions +3 to +8. Preceding the promoter core is usually a sequence rich in A and T nucleotides, which has been described for several mitochondrial genes in dicots as well as in maize to be important for unimpeded promoter function *in vitro* (13,17). Moreover, the all-A/T sequence TATATA is seen as an element not only of the CGTATATAA motif but also of deviating promoters, such as *Prrn18-353* and *Patp6-1-916/913* (Table 1), emphasizing the predisposition of A/T-rich nucleotide sequences to function as promoters. In line with this idea is the observation that of the two promoters *Patp8-710* and *PtrnM-574/573* showing neither a recognizable core motif nor the TATATA element, the latter is highly rich in A-T base pairs. The ability of mitochondrial sequences composed entirely of A and T nucleotides to support transcription initiation has been pointed out by Lupold *et al.* (34), based on their own observation of an all-A/T promoter of the maize *cox2* gene and an earlier report on yeast petite mutants presumably initiating mitochondrial RNA-synthesis at all-A/T sequences (35).

The detection of transcription initiation sites in the *Arabidopsis* mtDNA was strongly biased towards the identification of promoters displaying a CRTAAGAGA or CGTATATAA motif, as 9 out of 12 genes were selected for the analysis of transcriptional starts owing to the prior observation of at least one of these motifs in their 5' regions. Any of these motifs was confirmed to be part of an active promoter. Yet, since we found 50% of the identified promoters to deviate from these motifs, considerably more than half of all mitochondrial promoters in *Arabidopsis* may be expected to show divergent sequences. This underlines the necessity to experimentally define mitochondrial transcription initiation sites and the limited possibility of predicting these sites, based on conserved promoter motifs.

Defined promoter elements, such as particular core sequences or the TATATA motif, appear to be distributed randomly between different promoters. This precludes *Arabidopsis* mitochondrial promoters from being classified into distinct groups that could be related to different

mitochondrial RNA polymerases, based on merely a comparison of promoter sequences. In *Arabidopsis*, the nucleus-encoded phage-type RNA polymerase RpoTnp is targeted not only to mitochondria but also to plastids (6). This enzyme may thus be expected to recognize promoters of similar structure in both organelles. Many of the mitochondrial promoters characterized in the present study contain a CRTA core, which resembles the YRTA motif displayed by a subset of plastid promoters that are most likely used by phage-type RNA polymerases (19,36). Yet, the relatively small number of plastid promoters of this type studied thus far and the variability of mitochondrial promoter structures make it difficult to assign a distinct subset of mitochondrial promoters to RpoTnp. In addition, promoter recognition by this RNA polymerase may be mediated by different, yet unidentified cofactors in plastids and in mitochondria, and may therefore depend on different promoter sequences.

Frequent duplications and rearrangements of plant mitochondrial genomes during evolution (37) appear to have been important mechanisms of establishing promoter sequences throughout the mtDNA that are recognized by the nucleus-encoded transcriptional apparatus. Extended sequences showing high similarity to the *rrn18* 5' region and comprising both *Prrn18*-69 and *Prrn18*-156 are seen upstream of the *orf275-orf149-nad5c-nad4L-orf25* and *orf153b-orf118-orf114-nad3-rps12-orf117-ccb203* gene clusters. Although it was not tested whether transcription is initiated within the duplicated *rrn18*-like regions, it is possible that they direct transcription of the two clusters. A duplication of promoter regions is most apparent for the *atp1*, *atp6*, *atp8*, *atp9* and *rrn26* genes, which possess homologous 5' sequence stretches that comprise one or two promoters (compare promoter sequences in Table 1 and promoter distribution in Figure 6).

Promoter distribution

For 9 out of 12 genes, we detected multiple transcription initiation sites (Figure 6). Six genes were found to be transcribed from three or even four promoters. Interestingly, in a series of initiation sites of one gene, promoters containing the conserved nonanucleotide motifs CGTATATAA or CRTAAGAGA were usually positioned downstream of promoters with deviating or apparently no motifs. It may be speculated that transcription is initiated at additional sites even further upstream of the defined promoters, but that it has been impossible to detect these as the experimental strategy applied in this study does probably not cover regions exceeding 2 kb. Identifying additional upstream sites would require an extensive primer walking strategy.

While in dicots, multiple promoters have so far only been reported for the pea *cox2* gene (38), transcription initiation at multiple sites has been described for several mitochondrial genes in maize, rice and sorghum (34,39–42). It is unlikely that frequent promoter multiplicity, as revealed here for the first time for a dicotyledonous plant, would be restricted to *Arabidopsis*. 5'-RACE analyses, which easily trace even small amounts of primary transcripts, might uncover multiple promoters also in the mitochondria of other dicots.

Within the regions investigated for transcription start sites, additional promoters not displayed in Table 1 may be active from which minor primary transcripts derive. Transcripts mapping to the possible transcription initiation sites *Pcox2*-231

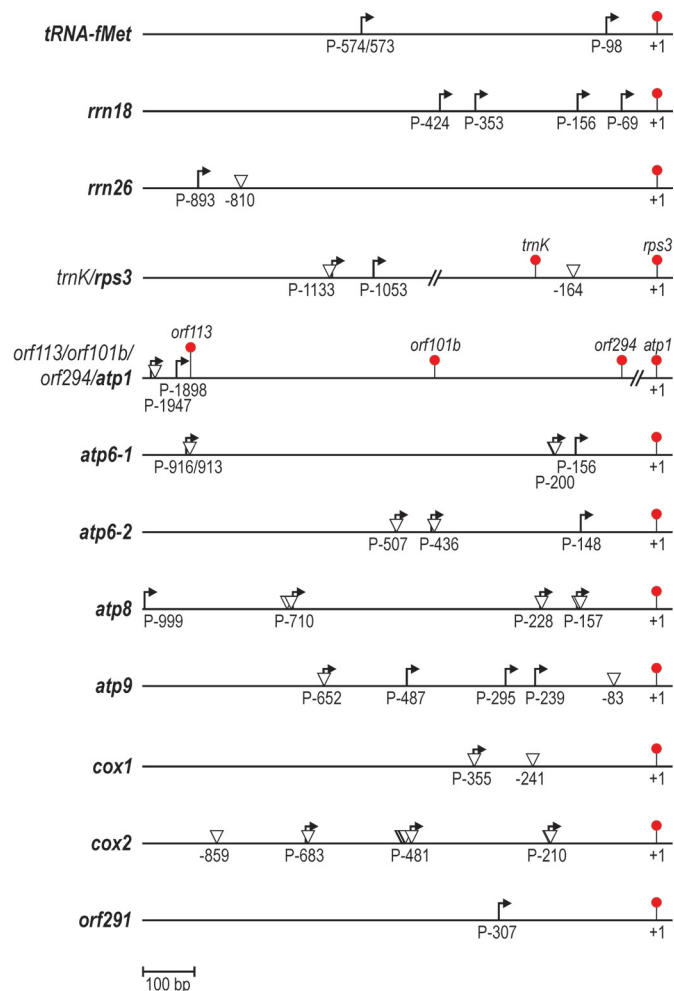


Figure 6. Linear diagram of upstream regions of the mitochondrial genes *trnM*, *rrn18*, *rrn26*, *rps3*, *cox1*, *cox2*, *atp1*, *atp6-1*, *atp6-2*, *atp8*, *atp9* and *orf291* for which transcript 5' ends were mapped in this study. Positions of transcriptional starts and processing sites are given as negative distances (in base pairs) from translational starts or mature rRNA or tRNA 5' ends equalling position +1, which is highlighted by a filled circle. The letter P precedes transcription initiation site positions; other symbols as in Figure 4.

(AGTATTATAA, see Figure 2) and *Prps3*-1149 (AGTAGATAG) did not give rise to distinct bands in the 5'-RACE, but were revealed as by-products of cloning and extensive sequencing of more abundant 5'-RACE products resulting from nearby transcriptional starts. These possible promoters are not listed in Table 1 because not enough data has been accumulated to substantiate transcription initiation at these sites. The cloning of transcript 5' ends not producing signals above the limit of detection in *in vitro* capping (*Pcox2*-231) or 5'-RACE experiments implies that a multitude of minor transcription start sites within mtDNA sequences may exist.

The role of mitochondrial promoter multiplicity in plant mitochondria is hitherto unknown. While variation in mitochondrial promoter usage has been described between *Zea mays* plants possessing different alleles of a nuclear gene (21), there has been no evidence to date that in a single plant species with a distinct nuclear background, mitochondrial genes would be differentially transcribed. The present study confirms mitochondrial transcription to initiate at

identical sites in leaves and in flowers of Arabidopsis plants, indicating that tissue-specific regulation of mitochondrial genes on the level of promoter selection is of only minor or no importance in Arabidopsis. This is in line with previous conclusions that in mitochondrial gene expression, regulatory mechanisms are aimed predominantly at posttranscriptional steps (43). Different promoters of a distinct gene may, however, vary in their activity between different tissues and possibly also between different developmental stages. Quantitative analyses of primary transcripts through primer extension experiments or quantitative real-time PCR would contribute to elucidating the activity and function of individual promoters controlling one gene.

It has been suggested that multiple promoters might enable regulatory mechanisms, such as differential promoter usage, producing different 5'-untranslated regions (5'-UTR) that possibly influence translational yield (34). Such mechanisms would be contrasted by analyses that detected a similar distribution of *cox3* transcripts with different 5'-untranslated sequences in both polysomal fractions and total RNA prepared from maize mitochondria, implying that mitochondrial ribosomes non-preferentially associate with differently initiated transcripts (44). From our data on primary transcript 5' termini in Arabidopsis mitochondria, we would favour a point of view that considers multiple promoters the result of a relaxed promoter specificity of the mitochondrial transcription machinery (34). Further experiments employing an *in vitro* transcription system will be required to clarify whether specific promoter types are recognized by different RNA polymerase complexes containing particular transcriptional cofactors or different core enzymes.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank Wolfgang Hess for valuable discussions of experimental data and for critical reading of the manuscript, Stefan Binder for sharing his knowledge on mitochondrial transcription in Arabidopsis and Thomas Hübschmann for helpful comments on capping experiments. The assistance of Martin Meixner in the sequencing work and of Conny Stock, Carola Emanuel and Monika Swiatecka in RNA preparation is gratefully acknowledged. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG; SFB429, TP A5). Funding to pay the Open Access publication charges for this article was provided by the DFG.

REFERENCES

- Kubo, T., Nishizawa, S., Sugawara, A., Itchoda, N., Estiati, A. and Mikami, T. (2000) The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA(Cys)(GCA). *Nucleic Acids Res.*, **28**, 2571–2576.
- Unsel, M., Marienfeld, J.R., Brandt, P. and Brennicke, A. (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366 924 nucleotides. *Nature Genet.*, **15**, 57–61.
- Handa, H. (2003) The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res.*, **31**, 5907–5916.
- Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A. and Kadowaki, K. (2002) The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Mol. Genet. Genomics*, **268**, 434–445.
- Hedtke, B., Börner, T. and Weihe, A. (1997) Mitochondrial and chloroplast phage-type RNA polymerases in Arabidopsis. *Science*, **277**, 809–811.
- Hedtke, B., Börner, T. and Weihe, A. (2000) One RNA polymerase serving two genomes. *EMBO Rep.*, **1**, 435–440.
- Hedtke, B., Legen, J., Weihe, A., Herrmann, R.G. and Börner, T. (2002) Six active phage-type RNA polymerase genes in *Nicotiana tabacum*. *Plant J.*, **30**, 625–637.
- Chang, C.C., Sheen, J., Bligny, M., Niwa, Y., Lerbs-Mache, S. and Stern, D.B. (1999) Functional analysis of two maize cDNAs encoding T7-like RNA polymerases. *Plant Cell*, **11**, 911–926.
- Ikeda, T.M. and Gray, M.W. (1999) Identification and characterization of T3/T7 bacteriophage-like RNA polymerase sequences in wheat. *Plant Mol. Biol.*, **40**, 567–578.
- Richter, U., Kiessling, J., Hedtke, B., Decker, E., Reski, R., Börner, T. and Weihe, A. (2002) Two RpoT genes of *Physcomitrella patens* encode phage-type RNA polymerases with dual targeting to mitochondria and plastids. *Gene*, **290**, 95–105.
- Tracy, R.L. and Stern, D.B. (1995) Mitochondrial transcription initiation: promoter structures and RNA polymerases. *Curr. Genet.*, **28**, 205–216.
- Fey, J. and Marechal-Drouard, L. (1999) Compilation and analysis of plant mitochondrial promoter sequences: an illustration of a divergent evolution between monocot and dicot mitochondria. *Biochem. Biophys. Res. Commun.*, **256**, 409–414.
- Rapp, W.D., Lupold, D.S., Mack, S. and Stern, D.B. (1993) Architecture of the maize mitochondrial *atp1* promoter as determined by linker-scanning and point mutagenesis. *Mol. Cell. Biol.*, **13**, 7232–7238.
- Rapp, W.D. and Stern, D.B. (1992) A conserved 11 nucleotide sequence contains an essential promoter element of the maize mitochondrial *atp1* gene. *EMBO J.*, **11**, 1065–1073.
- Caoile, A.G. and Stern, D.B. (1997) A conserved core element is functionally important for maize mitochondrial promoter activity *in vitro*. *Nucleic Acids Res.*, **25**, 4055–4060.
- Binder, S., Hatzack, F. and Brennicke, A. (1995) A novel pea mitochondrial *in vitro* transcription system recognizes homologous and heterologous mRNA and tRNA promoters. *J. Biol. Chem.*, **270**, 22182–22189.
- Dombrowski, S., Hoffmann, M., Guha, C. and Binder, S. (1999) Continuous primary sequence requirements in the 18-nucleotide promoter of dicot plant mitochondria. *J. Biol. Chem.*, **274**, 10094–10099.
- Hoffmann, M. and Binder, S. (2002) Functional importance of nucleotide identities within the pea *atp9* mitochondrial promoter sequence. *J. Mol. Biol.*, **320**, 943–950.
- Hess, W.R. and Börner, T. (1999) Organellar RNA polymerases of higher plants. *Int. Rev. Cytol.*, **190**, 1–59.
- Binder, S., Marchfelder, A. and Brennicke, A. (1996) Regulation of gene expression in plant mitochondria. *Plant Mol. Biol.*, **32**, 303–314.
- Newton, K.J., Winberg, B., Yamato, K., Lupold, S. and Stern, D.B. (1995) Evidence for a novel mitochondrial promoter preceding the *cox2* gene of perennial teosintes. *EMBO J.*, **14**, 585–593.
- Edqvist, J. and Bergman, P. (2002) Nuclear identity specifies transcriptional initiation in plant mitochondria. *Plant Mol. Biol.*, **49**, 59–68.
- Emanuel, C., Weihe, A., Graner, A., Hess, W.R. and Börner, T. (2004) Chloroplast development affects expression of phage-type RNA polymerases in barley leaves. *Plant J.*, **38**, 460–472.
- Dombrowski, S., Hoffmann, M., Kuhn, J., Brennicke, A. and Binder, S. (1998) On mitochondrial promoters in *Arabidopsis thaliana* and other flowering plants. In Möller, I.M., Glaser, E. and Glimelius, K. (eds.), *Plant Mitochondria*. Backhuys Publishers, The Netherlands, pp. 165–170.
- Brennicke, A., Zabaleta, E., Dombrowski, S., Hoffmann, M. and Binder, S. (1999) Transcription signals of mitochondrial and nuclear genes for mitochondrial proteins in dicot plants. *J. Hered.*, **90**, 345–350.
- Giese, A., Thalheim, C., Brennicke, A. and Binder, S. (1996) Correlation of nonanucleotide motifs with transcript initiation of 18S rRNA genes in mitochondria of pea, potato and Arabidopsis. *Mol. Gen. Genet.*, **252**, 429–436.
- Bensing, B.A., Meyer, B.J. and Dunny, G.M. (1996) Sensitive detection of bacterial transcription initiation sites and differentiation from RNA

- processing sites in the pheromone-induced plasmid transfer system of *Enterococcus faecalis*. *Proc. Natl Acad. Sci. USA*, **93**, 7794–7799.
28. Argaman, L., Hersberg, R., Vogel, J., Bejerano, G., Wagner, E.G., Margalit, H. and Altuvia, S. (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.*, **11**, 941–950.
 29. Vogel, J., Axmann, I.M., Herzel, H. and Hess, W.R. (2003) Experimental and computational analysis of transcriptional start sites in the cyanobacterium *Prochlorococcus* MED4. *Nucleic Acids Res.*, **31**, 2890–2899.
 30. Brown, G.G., Auchincloss, A.H., Covello, P.S., Gray, M.W., Menassa, R. and Singh, M. (1991) Characterization of transcription initiation sites on the soybean mitochondrial genome allows identification of a transcription-associated sequence motif. *Mol. Gen. Genet.*, **228**, 345–355.
 31. Lizama, L., Holuigue, L. and Jordana, X. (1994) Transcription initiation sites for the potato mitochondrial gene coding for subunit 9 of ATP synthase (*atp9*). *FEBS Lett.*, **349**, 243–248.
 32. Huang, J., Struck, F., Matzinger, D.F. and Levings, C.S., III (1994) Flower-enhanced expression of a nuclear-encoded mitochondrial respiratory protein is associated with changes in mitochondrion number. *Plant Cell*, **6**, 439–448.
 33. Smart, C.J., Moneger, F. and Leaver, C.J. (1994) Cell-specific regulation of gene expression in mitochondria during anther development in sunflower. *Plant Cell*, **6**, 811–825.
 34. Lupold, D.S., Caoile, A.G. and Stern, D.B. (1999) The maize mitochondrial *cox2* gene has five promoters in two genomic regions, including a complex promoter consisting of seven overlapping units. *J. Biol. Chem.*, **274**, 3897–3903.
 35. Fangman, W.L. and Dujon, B. (1984) Yeast mitochondrial genomes consisting of only A.T base pairs replicate and exhibit suppressiveness. *Proc. Natl Acad. Sci. USA*, **81**, 7156–7160.
 36. Liere, K., Kaden, D., Maliga, P. and Börner, T. (2004) Overexpression of phage-type RNA polymerase RpoTp in tobacco demonstrates its role in chloroplast transcription by recognizing a distinct promoter type. *Nucleic Acids Res.*, **32**, 1159–1165.
 37. Palmer, J.D. (1990) Contrasting modes and tempos of genome evolution in land plant organelles. *Trends Genet.*, **6**, 115–120.
 38. Kuhn, J. and Binder, S. (2002) RT-PCR analysis of 5' to 3'-end-ligated mRNAs identifies the extremities of *cox2* transcripts in pea mitochondria. *Nucleic Acids Res.*, **30**, 439–446.
 39. Mulligan, R.M., Lau, G.T. and Walbot, V. (1988) Numerous transcription initiation sites exist for the maize mitochondrial genes for subunit 9 of the ATP synthase and subunit 3 of cytochrome oxidase. *Proc. Natl Acad. Sci. USA*, **85**, 7998–8002.
 40. Nakazono, M., Ishikawa, M., Yoshida, K.T., Tsutsumi, N. and Hirai, A. (1996) Multiple initiation sites for transcription of a gene for subunit 1 of F1-ATPase (*atp1*) in rice mitochondria. *Curr. Genet.*, **29**, 417–422.
 41. Nakazono, M., Nishiwaki, S., Tsutsumi, N. and Hirai, A. (1996) A chloroplast-derived sequence is utilized as a source of promoter sequences for the gene for subunit 9 of NADH dehydrogenase (*nad9*) in rice mitochondria. *Mol. Gen. Genet.*, **252**, 371–378.
 42. Yan, B. and Pring, D.R. (1997) Transcriptional initiation sites in sorghum mitochondrial DNA indicate conserved and variable features. *Curr. Genet.*, **32**, 287–295.
 43. Giege, P., Hoffmann, M., Binder, S. and Brennicke, A. (2000) RNA degradation buffers asymmetries of transcription in Arabidopsis mitochondria. *EMBO Rep.*, **1**, 164–170.
 44. Yang, A.J. and Mulligan, R.M. (1993) Distribution of maize mitochondrial transcripts in polysomal RNA: evidence for non-selectivity in recruitment of mRNAs. *Curr. Genet.*, **23**, 532–536.
 45. Schneider, T.D. and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.*, **18**, 6097–6100.
 46. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.