From gene to protein in higher plant mitochondria

Philippe Giegé*a, Axel Brennickeb

a Department of Plant Sciences, Oxford University, South Parks Road, Oxford, OX1 3RB, UK
b Universität Ulm, Allgemeine Botanik, Albert-Einstein-Allee 11, 89069 Ulm, Germany

Abstract – Higher plant mitochondria contain a genetic system with a genome, transcription and translation processes, which have to be logistically integrated with the two other genomes in the nucleus and the plastid. In plant mitochondria, after transcripts have been synthesised, at least in some cases by a phage-type RNA polymerase, they have to go through a complex processing apparatus, which depends on protein factors imported from the cytosol. Processing involves cis- and trans-splicing, internal RNA editing and maturation at the transcript termini, these steps often occurring in parallel. Transcript life is terminated by RNA degradation mechanisms, one of which involves polyadenylation. RNA metabolism seems to be a key element of the regulation of gene expression in higher plant mitochondria. © 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

plant mitochondria / gene expression / maturation / degradation


mitochondrie de plante / expression génétique / maturation / dégradation

Version abrégée

La découverte de génomes dans les mitochondries a fortement contribué à réactualiser la théorie proposant une origine endosymbiotique aux cellules eucaryotes. Les mêmes fonctions fondamentales sont codées dans les génomes mitochondriaux. Néanmoins, ceux des plantes se caractérisent par leur grande taille, souvent 20 fois supérieure à celles des animaux.

Considérant la trentaine de protéines seulement codées dans les génomes mitochondriaux de plantes, la plupart des métabolismes présent dans ces organelles, incluant l’expression génétique, sont dépendant de protéines importées du cytosol. De même, les mitochondries de plantes ne disposent pas d’un lot d’ARNts

* Correspondence and reprints.
E-mail address: philippe.giege@plant-sciences.oxford.ac.uk (P. Giegé).
1. Structure of plant mitochondrial DNA

The finding of unique genomes in mitochondria strongly helped to reactualise the endosymbiotic theory for the evolution of the eukaryotic cell [1–3]. Sequence analysis has revealed that higher plant mitochondrial genomes, similar to their animal, fungal or protist counterparts, show a basal stability of gene content, but often differ in the presence or absence of individual genes. The same fundamental functions are found encoded in the mtDNAs of all eukaryotic cells. These are proteins involved in mitochondrial metabolism, i.e. genes encoding components of the complexes forming the respiratory chain, and in plants and protists also genes encoding proteins involved in the c-type cytochrome biogenesis, and components of the mitochondrial translational apparatus, i.e. a set of about 20 tRNAs, ribosomal RNAs and, mostly in plants and protists, genes for ribosomal proteins [4].

A major common characteristic of plant mitochondrial genomes is their large size, with an estimated 180–2 400 kb for the land plant genomes [5], with precise data available for the three completely sequenced genomes covering 187 kb in the liverwort Marchantia, 367 kb in Arabidopsis [6] and 369 kb in the sugar beet [7]. Gene density is comparatively low in plant mitochondria. The 16.5-kb human mitochondrial genome [8], though 22 times smaller than Arabidopsis mt genome, encodes a quarter of its genes. The actual coding regions represent only about 10 % of plant mt genomes, with large parts of these genomes (7 % in Arabidopsis) being taken up by repeated sequences, which are responsible for the physical organisation of the mitochondrial genome. In Arabidopsis, two large repeats of 6.5 and 4.2 kb have been shown to be active in recombination, resulting in the production of sub-genomic circular molecules [9]. Further large portions of the plant mtDNAs are covered by probably functionally insignificant open reading frames, by introns and by sequences imported from plastid and nuclear genomes. All in all, just over half of plant mitochondrial DNA is of as yet unaccounted function [6, 7].

2. Proteins and tRNAs have to be imported into the mitochondria from the cytosol

Given the small number of proteins (approximately 30) encoded in plant mitochondrial genomes, most metabo-
lisms taking place in this organelle, including genome expression, depend on imported proteins. About 4000 proteins [10] encoded in the nucleus are estimated to be imported from the cytosol to the mitochondrial compartment. Nuclear encoded mt proteins contain a signal peptide at their N-terminus, which is recognised by translocation cases of the outer (TOM) and inner mitochondrial membranes (TIM) at contact points of the two bilayers. Once translocated, the pre-proteins are released into the matrix by processing peptidases (MPP) that cleave off the signal peptides. This whole process is associated with chaperones involved in pre-protein unfolding, translocation (HSP70, MGE), refolding and assembly (MDJ, HSP60 and HSP10) (for review, see e.g. [11]).

It is very likely that an additional protein import apparatus exists in plant mitochondria, since the long time elusive orfx gene conserved in higher plant mitochondria [12] was identified to encode Mtt2, a component of a non-secretory protein transport pathway, analogous to one described in Escherichia coli [13].

Similarly to proteins, the set of rRNAs encoded in plant mitochondria is not sufficient to perform mitochondrial translation and tRNAs have to be imported from the cytosol [14, 15]. In Arabidopsis, 13 different tRNAs are missing to decode all of the codons used [4]. Several studies suggest that tRNAs may be imported in association with precursors (rRNA-synthetases, e.g. [16, 17]). This interaction, though necessary, alone, however, is sufficient to trigger tRNA import in mitochondria [18].

3. Mitochondrial RNA synthesis, transcription

3.1. Promoter sequences and cis-acting factors

Contrary to vertebrate mitochondria, where transcription initiation occurs at only one promoter per DNA strand, plant mitochondrial genome transcription is initiated at numerous sites. In the higher plant Oenothera, in vitro capping experiments showed that at least 15 promoters are active in the genome [19]. The individual promoters in plant mitochondria are of different strength as documented by the very heterogeneous run-on transcription levels observed for different genes in Arabidopsis [20]. Several promoters can be active for a single gene as described for atp9 and cox2 in maize mitochondria [21, 22]. Comparison between primer extension, nuclease S1 analysis and in vitro capping experiments revealed that higher plant mitochondrial transcripts do not always undergo processing in their 5’-regions, steady-state mRNA termini thus being in several instances true transcription initiation start points.

In vitro transcription assays and sequence analysis around 5’-transcript termini revealed that higher plant mitochondrial promoters are characterised by a conserved CRTA sequence (R being purines). For monocots, a little conserved 15 nucleotides long sequence containing the CRTA motif was shown to be sufficient for in vitro transcription initiation. In dicot mitochondria, a non-anucleotide motif was found conserved at transcription initiation sites [23]. This sequence is, however, not sufficient to give full in vitro transcription initiation. Comparison of various promoter structures from pea, Oenothera, soybean and potato revealed that an 18-nucleotide-long sequence is necessary and sufficient to initiate transcription at least for the genes investigated [24]. This sequence between positions −14 and +4 around the initiation point, the first transcribed nucleotide, contains an AT box in the 5’- and a purine-rich stretch in the 3’-region. This type of promoter sequence may not be responsible for all of the transcription initiation events, since these conserved motifs are not present in all promoters of plant mitochondria. This was observed first for the atp1 promoter in maize mitochondria and subsequently also seen for cox3 and atp6 promoters in the same plant [25, 26]. Similarly in dicots, e.g. in the Arabidopsis mitochondrial genome, numerous genes lack these or any other detectable consensus structures at their promoter sites. One or more other types of promoters must thus be active in plant mitochondrial transcription initiation.

3.2. RNA polymerase and other predicted trans-acting factors

In chloroplasts transcription is performed by at least three different RNA polymerases, a bacterial-type RNA polymerase (PEP) encoded in the plastid genome (i.e. a chloroplastic enzyme having a sequence homologous to RNA polymerases found in bacteria) and two nuclear-encoded bacteriophage-type RNA polymerases (NEP) (i.e. enzymes having their sequence homologous to bacteriophage RNA polymerases). In contrast, transcription in higher plant mitochondria seems to be catalysed by at least two nuclear-encoded phage-type enzymes; for a review see [27]. Detailed investigations in Arabidopsis have yielded in vitro as well as in vivo data demonstrating that three such single-subunit enzymes are coded by the nuclear genome, one of each being targeted to plastids and mitochondria and the third being imported into both organelles [28–30]. In the monocots maize and wheat similar mitochondrial RNA polymerases have been characterised, but their total number is not yet clear [31, 32].

As discussed above, dicot and particularly monocot promoters do not all conform to a single strict consensus sequence. Therefore, promoter recognition will require the different RNA polymerases to be specific for one type of promoter, and/or very likely also the additional presence of various transcription factors to modulate promoter specificity. One of these could be a DNA-binding protein recently characterised to attach to transcription initiation sites in wheat mitochondria [33]. In the dicot pea a different, apparently unrelated promoter-binding protein was identified in mitochondria, but its functional involvement in promoter recognition still needs to be determined [34]. Proteins similar to CDF2, which interacts with the phage-type RNA polymerase in the chloroplast, are possibly also...
involved in promoter recognition and activation in plant mitochondria [35].

4. Maturation of mitochondrial RNA

4.1. Splicing

4.1.1. Cis- and trans-splicing introns

Two classes of introns, group I and group II, are characteristic of organelles, first distinguished by their characteristic RNA secondary structures in fungal mitochondria by Michel et al. [36]. In higher plant mitochondria group II introns are generally found in several genes, while only one instance of a recently acquired, unique group I intron has been found in the cox1 genes of Peperomia and some other higher plants [37, 38].

Group II (and group I) introns sometimes encode proteins in their sequences, which function in the dispersion of the intron into new sites in the DNA or in its excision from the mRNA (maturases). Of the approximately 100 group II introns known up to now, approximately one-fourth potentially encode proteins, most of them maturases, which are usually necessary for splicing of at least their respective intron.

The genes encoded in the Arabidopsis mitochondrial genome are interrupted by altogether 23 group II introns with sizes varying from 485 to approximately 4 000 nucleotides [6]. Some genes are interrupted by more than one intron, e.g. nad7 has four introns, but only the intron between exons d and e of nad1 encodes a maturase-like open reading frame in most of the plants.

Trans-splicing is found in plant mitochondria in several instances, where the physically separated exons are flanked by partial group II intron sequences [39–41]. Trans-splicing designates splicing events where consecutive exons are not linearly separated by a continuous intron, but are present at distant positions on the genome, sometimes on the other strand. These trans-splicing events occur via the group II intron fragments, which connect the separated exons in two- or three-molecule interactions. In Arabidopsis mitochondria, five trans-splicing events are detected in nad1 [40], nad2 [42] and in nad5 [43], in the latter an only 22-nucleotide-long exon being integrated by two trans-splicing events.

4.1.2. Splicing mechanisms and factors involved

Similarly to nuclear introns, group I and II intron splicing occurs by transesterification reactions. In group I introns, free GTP is needed as cofactor with its 3’-OH group attacking the 5’-end of the intron. In group II introns, the 5’-end of the intron is attacked by an OH group of a conserved internal A residue. The highly conserved structure of the group II introns has been shown to be essential for splicing activity [44, 45]. Several introns of both organellar groups are characterised by their ability to perform autocatalytic splicing in vitro under essentially non-physiological conditions. In vivo, RNA structure can not account for splicing by itself, but including the maturases sometimes internally encoded, introns require protein factors to catalyse splicing. Recent studies have shown in tobacco that mitochondrial splicing is dependent on the expression of at least one nuclear-encoded factor, a probable member of the predicted splicing machinery [46].

4.2. RNA editing

4.2.1. Global effect

The discovery of RNA editing in plant mitochondria, manifested as mainly C to U conversions in mRNAs [47–49], provided an answer to the long-term puzzle concerning codon conservation in plant mitochondrial genes, including the use of a potentially specific genetic code. Besides C to U alterations in RNAs, U to C ‘reverse editing’ events have also been observed in both plant organelles, although rarely in flowering plants, e.g. [50–52]. RNA editing affects preferentially protein coding transcripts, but is occasionally also found in structural RNAs. In mRNAs, editing modifications result in increasing the conserved similarity of the deduced protein sequences with homologous proteins in other species. Editing thus ‘corrects’ affected codons to those conserved in ‘non-editing’ genetic systems. It has also been observed that the 441 editing modifications found in Arabidopsis mRNAs lead to an overall increase of the hydrophobicity of the mitochondrially encoded proteins [53]. In structural RNAs, usually tRNAs, RNA editing is believed to improve the spatial folding of these molecules, and thus to improve their functionality.

4.2.2. The editase

In order to identify the enzyme(s) catalysing the RNA editing activity in plant organelles, the first question addressed was the biochemical nature of the C to U and U to C conversions. Different in vitro and in organello experimental assays such as internal labelling with radioactive CTP of RNAs in purified mitochondria [54] and incubation of synthetic RNAs containing radiolabelled Cs with in vitro active extracts [55] have shown that the RNA backbone is not disrupted by the editing activity, the phosphate and ribose groups remaining unchanged. Other experiments showed that the amino group of the cytidine is removed by editing while the base labelled on its α-phosphate remains in the RNA [54, 56]. Zinc ions, typical co-factors of deaminases, are found to be necessary for the in vitro editing activity [55]. The present working hypothesis is that RNA editing occurs via a deamination or a transamination reaction. If deaminases are the editing enzymes, the rare ‘reverse’ editing should have to be catalysed by a separate CTP synthase activity. On the other hand, a transaminase could catalyse both C to U and U to C editing activities possibly by using different cofactors. The cytidine deaminases and transaminases found in plant genomes have characteristic conserved sequences. It is believed that the editing enzyme(s) must be encoded in the plant cell nucleus since no such characteristic
sequences has been found up to now in the complete mitochondrial genome sequences available [6, 7].

4.2.3. Editing site recognition

When comparing the editing site sequences at the genome level [53] no conserved sequence or RNA folding pattern can be observed. A bias towards an elevated frequency of pyrimidines directly in 5' of the sites is observed, but insufficient to label the respective Cs as ‘editable’. Individual RNA editing sites are thus most likely targeted one by one (or as small families) by separate specificity factors, probably acting in trans. Chaudhuri et al. [57] showed that for chloroplasts the introduction of a transgene inducing additional RNA copies had the consequence of decreasing the editing frequency. This suggests that a trans-acting factor may be present in limited amounts. This factor could a priori be a protein or a nucleic acid, but at least in plant mitochondrial transcripts recognition solely by site-specific proteins seems disfavoured by the large number of different editing sites. Indeed, an in vitro extract described by Araya et al. [58] appeared to be vulnerable to both protease and micrococal nuclease activities. This suggests that nucleic acids are involved in the RNA editing mechanism, which may be small RNA molecules in antisense to editing site sequences. The search for such molecules guiding the editing activity (gRNAs) has identified in Oenothera fragments of the 26S ribosomal RNA with limited antisense pairing to editing sites [59]. These similarities could, however, be due to statistically random nucleotide combinations and thus will have to be confirmed with other editing sites and in other plants.

4.2.4. RNA editing and the control of gene expression

The analysis of partially edited transcripts revealed that in plant mitochondria editing is not a sequential process scanning along the RNA thread, but that the editing machinery probably randomly homes in on the site to be edited. Analyses of large numbers of individual cDNA clones for a given gene reveal site-specific variations in editing frequencies, which result in a heterogeneously edited steady-state RNA population. The frequency of partial editing differs in mature RNAs, precursor RNAs and polysome-enriched fractions. Mature RNAs are found to be more frequently edited and polysome-enriched fractions are nearly all completely edited [60], suggesting that editing is a post-transcriptional process. The detection of editing in non-spliced pre mRNAs indicates that editing probably happens independently of splicing. In Arabidopsis mitochondria no correlation could be observed between the population of edited codons and the codon usage, suggesting that the RNA editing machinery acts independently of the translation apparatus [53]. RNA editing thus appears to take place anywhere between transcription and translation and could there play a role in regulation.

4.3. Transcript 3'-processing

In plant mitochondria, the 3'-untranslated regions often contain inverted repeat sequences [61]. In animal mitochondria, such structures are found to be regulated terminators of transcription [62] similar to bacteria, where inverted repeats have been found to act as rho-independent terminators [63]. In chloroplasts, stem-loop structures in the 3'-region of mRNA genes act as both processing signals and stability elements [64]. Proteins binding to such structures have been identified in these organelles to include both exo- and endonucleases [65]. Differential binding of these factors plays a role in the stability and in the turnover of the respective transcripts. In mitochondria of pea, the function of such an inverted repeat structure located in 3' of atp9 has been investigated [66]. In vitro transcription is not inhibited at this structure, but an in vitro processing system recognises the inverted repeat, which in this assay checks and stops the progressive exonucleolytic removal of nucleotides from the 3'-terminus. Thus, in plant mitochondria, inverted repeats are believed to be processing signals rather than transcription terminators. The DexH box RNA helicase recently characterised in Arabidopsis mitochondria [67] may help to solve such RNA structures at transcript 3'-ends, and may be part of the predicted machinery responsible for mRNA 3'-maturations. It is thus possible that this protein is involved in the control of transcript degradation in plant mitochondria.

5. Mitochondrial RNA degradation

In the nucleus and the cytoplasm polyadenylation enhances the stability of steady-state transcripts, whereas in plant mitochondria and in plastids polyadenylation rather destabilises affected RNAs. The large 3'-untranslated regions present in some of the mitochondrial mRNAs probably play individually distinct roles in modulating stability and maturation of transcripts. Transcript polyadenylation has recently been detected in plant mitochondrial mRNAs [68, 69]. For example, in a CMS (cytoplasmic male sterile) sunflower line, an atpA-orf522 transcript is polyadenylated in vivo. An increase in this polyadenylation activity can be correlated with an increase in transcript instability rather than with an prolonged lifespan. Nicely complementing these results is the ribonuclease activity detected in sunflower mitochondria which preferentially degrades polyadenylated transcripts [68]. Polyadenylation of transcripts has also been described in maize cox2 mRNAs [69]. These observations suggest that polyadenylation, as previously observed for chloroplasts [70] and bacteria [71], is also involved in the degradation of transcripts in plant mitochondria. The trans-acting signals predicted to trigger selective and regulatory polyadenylation remain as yet elusive. Such polyadenylation signals do not appear to be affected by RNA editing, since edited and unedited transcripts are equally found polyadenylated [69]. It is unclear whether any such specificity for poly-
Figure 1. This cartoon summarises gene expression mechanisms presently characterised or envisaged to play a role in the genetic system of plant mitochondria. Predicted but not yet identified factors are indicated by question marks. Black arrows show actions of proteins, black dashed arrows suggest potential protein actions and grey dashed arrows show the movement and localisation of proteins and tRNAs. The bold grey arrows separate the major stages of gene expression, i.e. transcription, transcript maturation, translation and transcript degradation processes.
adenylation is required at all, since only a few transcripts have been investigated for poly-A addition and it is still feasible to propose a non-selective addition of tails to all RNAs in plant mitochondria.

6. Translation and regulation at the protein level

6.1. Translation initiation sites

Sequence analysis has shown that translation is usually, but not always, initiated with an AUG codon in higher plant mitochondria. In Arabidopsis GGG, AAU and GUG are possible additional translation initiation triplets [6]. Similarly, in Oenothera GUG and in radish ACG were found to be potential translation initiation sites [72, 73]. Unlike in the chloroplast, where Shine-Dalgarno-like sequences have been found for some genes, the ribosome-binding sites presently remain elusive in plant mitochondria, e.g. [74].

6.2. Regulation of gene expression at the protein level

Protein sequence comparisons have shown that only proteins encoded by fully edited transcripts should yield evolutionary conserved and functional proteins. Translation of fully, partially and non-edited transcripts would produce a family of protein isoforms deviating at non-silent editing sites. Protein analysis in maize mitochondria using specific antibodies produced against proteins coded by non-edited, partially edited and fully edited rps12 fragments, revealed that ‘partially edited proteins’ are indeed synthesised [75]. On the other hand, direct sequencing of the proteins ATP9 from wheat [76], NAD9 from potato [77] and ATP6 from petunia [78] revealed that only ‘fully edited proteins’ are present in the respective multi-subunit polyprotein complexes. These observations suggest that while all the mRNAs are translated indiscriminately of their editing status, selection occurs post-translationally, and non-functional proteins are not incorporated into the complexes and are presumably rapidly degraded.

Other recent studies in Arabidopsis have shown that the protein complex subunit stoichiometries correlate better with the steady-state transcription levels than with the respective run-on transcription levels [20]. This suggests that regulation of gene expression occurs at the post-transcriptional and post-translational levels.

7. Origin and evolution of unique features

The evolutionary histories of some of the specific features in the genetic system of plant mitochondria reflect the fluidity and flexibility of genetic systems, particularly of such comparatively small genomes (figure 1). The two presently identified RNA polymerases appear to be derived from phage-type ancestral enzymes, which in the emerging eukaryotic cells presumably replaced the original bacteria-like polymerase. Evidence for this scenario has been found in the protozoon Reclinomonas americana, which harbours the only known mitochondrial genome still coding for the bacterial-type enzyme [79].

The trans-splicing introns have evolved from originally cis-arranged continuous exon–intron structures, which were disrupted by genome recombination and reshuffling events during the evolution of the higher plants. In depth investigations of intron structures in ferns, fern allies and a hornwort revealed cis-homologues for all of the higher plant trans-splicing introns and showed that some of these introns invaded the plant line very early during evolution [80].

RNA editing is also a feature that evolved during the evolution of plants, most likely in the early land plants. The presence of this type of editing in all land plants but the Marchantiidae may reflect the evolutionary history of this process, its origin thus being placed within the line of the liverworts [81, 52].

Acknowledgements. We apologise to all colleagues whose work could not be cited because of space limitations. We thank Professor C.J. Leaver for his continuous encouragement. P.G. was funded by an EMBO-fellowship. Support was provided by the Fonds der chemischen Industrie.

References

Evidence for a novel mitochondrial promoter preceding the
557
plant mitochondria, J. Biol. Chem. 274 (1999) 10094
(1996) 429
perennial
Teosintes
sation of a gene encoding a single-subunit bacteriophage-type RNA
[22] Newton K.J., Winberg B., Yamato K., Lupold S., Stern D.B.,
ö
[23] Giese A., Thalheim C., Brennicke A., Binder S., Correlation of
Ö
[25] Gray M.W., Lang B.F., Transcription in chloroplasts and mitochon-
[26] Trac R.L., Stern D.B., Mitochondrial transcription initiation: pro-
[27] Gualberto J.M., Lamattina L., Bonnard G., Weil J.H., Grienen-
[28] Ikeda T.M., Gray M.W., Identi-
[31] Young D.A., Allen R.L., Harvey A.J., Lonsdale D.M., Characteri-
sation of a gene encoding a single-subunit bacteriophage-type RNA
[32] Ikeda T.M., Gray M.W., Characterization of a DNA-binding pro-
[34] Hatzak F., Dombrowski S., Brennicke A., Binder S., Characteriza-
[39] Chapdelaine Y., Bonen L., The wheat mitochondrial genome for
[40] Wissinger B., Schuster W., Brennicke A., Trans-splicing in Oenothera mitocondria: nad1 mRNAs are edited in exon and trans-
intron is associated with lack of several
[45] Steinhauser S., Beckert S., Capesius I., Malek O., Knoop V., Plant
RNA editing of a conserved
nad5 mRNA is edited in higher plant mitochondria, EMBO J. 10 (1991) 3483–3493.
berger J.M., RNA editing in wheat mitochondria results in the conserva-
[51] Hiesel R., Wissinger B., Schuster W., Brennicke A., RNA editing in
nad2 gene: evidence for twelve C to U and one U to C
conversions and for sequence similarities around editing sites, Nucleic


