Gene expression in plant mitochondria: transcriptional and post-transcriptional control

Stefan Binder and Axel Brennicke*

Molekulare Botanik, Universität Ulm, 89069 Ulm, Germany

The informational content of the mitochondrial genome in plants is, although small, essential for each cell. Gene expression in these organelles involves a number of distinct transcriptional and post-transcriptional steps. The complex post-transcriptional processes of plant mitochondria such as 5' and 3' RNA processing, intron splicing, RNA editing and controlled RNA stability extensively modify individual steady-state RNA levels and influence the mRNA quantities available for translation. In this overview of the processes in mitochondrial gene expression, we focus on confirmed and potential sites of regulatory interference and discuss the evolutionary origins of the transcriptional and post-transcriptional processes.

Keywords: plant mitochondria; transcription; RNA processing

1. INTRODUCTION

The three genomes in the plant cell have to be coordinately expressed, particularly to yield functional photosynthesis and respiration, the basal sources of energy for the plant. The plastid and the mitochondrial genomes derive from formerly free-living prokaryotes and were taken up by the eukaryotic cell, thereby providing efficient machines of energy metabolism (Gray 1999; Lang *et al.* 1999; Race *et al.* 1999). In the progressive integration of the endosymbiont to the present-day mitochondrial genome, numerous genes have moved from the organelle to the nucleus. These include the genes for 99% of the 2000–3000 mitochondrial proteins. The transferred genes are now expressed under direct central control of the nucleus and their protein products are subsequently imported into the organelle.

The residual genes in the mitochondrion, 57 in the flowering plant *Arabidopsis thaliana* (Unseld *et al.* 1997) and a similar number in the liverwort *Marchantia polymorpha* (Oda *et al.* 1992), code for subunits of multiprotein complexes of the respiratory chain, of haem and cytochrome assembly and the ribosome (Unseld *et al.* 1997). The mitochondrial genes specify only some of the subunits in each of these complexes; the other polypeptides are contributed by nuclear genes. The respective genes in both compartments thus have to be tightly coordinated in their activity. Changes in the availability of these complexes demanded, for instance, by changes in the environment must consequently be met in both genetic systems, and respective responses have to be elicited both in the nucleus and in the mitochondrion.

In the analysis of mitochondrial gene expression, we focus on elucidating the levels of control in the cascade between DNA and the protein product (figure 1). The

investigation of the contribution of promoter activities and the various features of RNA processing and RNA stability will eventually assemble an overall picture of the control mechanisms within the organelle and the sites of interference from the nucleus or cytoplasm.

2. PROMOTERS IN PLANT MITOCHONDRIA

The physical arrangement of the 57 genes on the mitochondrial genome in *A. thaliana* and in other higher plants already suggests that several promoters will be required to transcribe these genes (Unseld *et al.* 1997; Dombrowski *et al.* 1998). The spacers of 2–60 kb between each gene, and the generally much shorter mRNAs detected for most genes in Northern blot analyses, are most easily rationalized by individual transcription units. This assumption is vindicated by the observation of more than 20 promoters in a global survey of transcription initiation sites. A number of these promoter regions have previously been identified and structural comparisons of these sites allowed the extraction of some of their characteristics.

In flowering plants one promoter type is characterized by a consensus sequence motif, which was identified in monocots as a tetranucleotide and in dicots as a nonanucleotide. At least in dicot plants the other promoters are defined by just the lack of such common structures. This negative definition does not preclude hidden similarities between several transcription initiation sites, but these have yet to be identified. Nothing more is known about this latter, divergent class of promoters, whereas the conserved promoters have been intensively investigated in monocot as well as in dicot plants, mostly through *in vitro* systems. Both classes of promoters can be found at transcriptional start sites of all types of RNAs: mRNA, tRNA and rRNA. In the *A. thaliana* mitochondrial genome 29 potential promoters, which match most of the criteria of the conserved type, can be found in the complete sequence. Of these 29 sites, 16 are located within 3 kb upstream of a recognized gene, including one near the 18S

* Author for correspondence (axel.brennicke@biologie.uni-ulm.de).

One contribution of 21 to a Discussion Meeting Issue 'Chloroplasts and mitochondria: functional genomics and evolution'.

Figure 1. In plant mitochondria a complex array of processing steps matures the primary transcript to the translatable mRNA. These maturation processes include RNA editing, intron splicing, and trimming of 5' and 3' termini. Control of gene expression may be exercised at any of these steps. In addition, gene activity may be modulated at the promoters, and during and after translation of the mRNA. Gene expression is terminated by degradation of the mRNA, which can be enhanced by polyadenylation.

rRNA gene, and are thus good candidates for genuine promoters (Dombrowski *et al.* 1998). In monocot plants the consensus motif of the conserved promoter type is centred around four conserved nucleotides just upstream from the first transcribed nucleotide (Rapp & Stern 1992). A second part of the element resides about 10 nucleotides further upstream and is characterized as an AT-rich region of six residues. In dicot plants, the promoter structures of this type are much better conserved, since a stretch of nine nucleotides shows considerable sequence identity between different genes as well as between different species. Through mutational analyses, the required promoter regions are found to extend beyond the obviously conserved regions (Hoffmann & Binder 2002). For example, in the AT-rich region not only is the AT content *per se* important, but also certain nucleotide identities are required for efficient

in vitro transcription initiation. An exchange from A to T between the different DNA strands in a given promoter will almost completely abolish recognition by the initiation complex. Intriguingly, the conserved nonanucleotide region extends into the transcribed region with two nucleotides. Further downstream into the transcribed region, the identity of the next two nucleotides was also found to be important for efficient promoter activation (Hoffmann & Binder 2002). Thus, many of the transcripts, at least those from this conserved type of promoter in dicot plants, will begin with the same nucleotides, characteristically the GA dinucleotide. This feature of the promoter overlapping the transcription initiation site has been observed in phage-type polymerases, and indeed RNA polymerases of this class have been identified for plant mitochondria as well as plastids.

3. RNA POLYMERASE AND ITS COFACTORS

The single-subunit phage-type RNA polymerases of plant organelles are similar to the enzymes in mammalian and fungal mitochondria, although their primary structures have diverged considerably. In the flowering plant *A. thaliana* three such genes are active in the nuclear genome, of which one protein product is targeted solely to the plastid, one exclusively to the mitochondrion and the third to both organelles (Hedtke *et al.* 2000). Phylogenetically, the three phage-type RNA polymerases in plant nuclear genomes arose by two independent duplication events in the plant lineage, and subsequently became integrated into the respective organellar transcription system.

It is probable that one, or both, of the mitochondrial RNA polymerases are active at the conserved promoters, but a definite functional assignment is still missing. Which RNA polymerase activates the variable promoter sites is at present completely unknown. While chloroplasts use an additional multi-subunit bacteria-like polymerase and yet a third (or fourth) so far largely uncharacterized activity, no such candidate is found for mitochondria in plants. No trace of the bacterial-type RNA polymerase found in the mitochondrial genome of the early branching protist *Reclinomonas americana* is left in the organellar or nuclear genomes of plants (Lang *et al.* 1997). Thus, a currently unknown RNA polymerase may have been activated to support mitochondrial transcription in plants.

An alternative scenario to that of both mitochondrial phage-like RNA polymerases addressing the conserved nonanucleotide motif containing promoters and invoking an additional as yet unknown enzyme, is that one of the two phage-type enzymes, either the single or the dual targeted one or both, may be guided to the non-conserved promoter sites by as yet unknown cofactors.

While the genuine phage-encoded enzymes do not require any cofactors for either promoter recognition or processivity, the mitochondrial phage-type RNA polymerases usually do need such factors. These have been identified in yeast and mammalian mitochondria as mtTFA and mtTFB (McCulloch *et al.* 2002). In plant mitochondria, *in vitro* promoter recognition is very fragile and rapidly loses activity in most manipulations. In particular, attempts at fractionation abolish all activity, suggesting that several loosely connected proteins may be involved in promoter recognition.

Investigations of proteins binding to the conserved promoter motifs in monocot and dicot plants indeed identified proteins, other than the polymerase itself, preferentially associated with this DNA sequence *in vitro*. In wheat (*Triticum aestivum*) a 67 kDa protein was purified that contains a DNA-binding domain (Ikeda & Gray 1999). This polypeptide stimulates *in vitro* transcriptional activity from a wheat *cox2* mitochondrial promoter and is thus a very promising candidate for a *bona fide* promoter-binding protein. In pea (*Pisum sativum*) a 43 kDa protein was purified from mitochondria by its promoter-binding properties, which shows high similarity to isovaleryl CoA dehydrogenase enzymes from other organisms (Däschner *et al.* 2001). *In vitro* assays with the protein synthesized from the cloned gene in *Escherichia coli* confirm this enzymatic activity. It remains unclear whether this protein also enhances promoter recognition *in vitro*. If it does, this

could be an interesting example of a dual-function protein, which could connect metabolic activities and requirements with gene expression in mitochondria. Such bifunctional molecular connections have been identified in yeast mitochondria, where several proteins are involved in gene expression, as well as in metabolic and energy-transforming processes. In mammalian mitochondria, the mtTFB transcription factor also has a second function as an RNA adenine methyltransferase (McCulloch *et al.* 2002). It is highly probable that many more such dual functions for proteins will be identified in the near future, particularly in the organelles.

4. RNA 5' AND 3' MATURATION

Maturation of primary transcripts by processing at 5' and 3' nucleolytic incisions is most obvious for tRNAs and rRNAs. These are generally synthesized as larger precursors from upstream promoters. The order of processing has been established for tRNAs, for which the primary event is the 5' processing by RNase P or its homologue (Marchfelder *et al.* 1990). The second step is an endonucleolytic cut by the RNase Z, which removes the 3' trailer (Kunzmann *et al.* 1998; Schiffer *et al.* 2002).

The rRNAs are likewise excised from longer precursors, but it is unclear whether the acting enzymes are site-specific endonucleases or a combination of endo- and exonucleases. The involvement of endonucleases is suggested, for example, by the finding that sequences surrounding the 5' termini of the 26S rRNAs are repeated at 5' processing sites in mRNAs (Schuster & Brennicke 1989; Maloney *et al.* 1989). These motifs may recruit the same enzymes also for processing of the mRNA molecules. The repeats cover two blocks of approximately 20 nucleotides each, certainly large enough to exclude chance similarities to other sequence contexts. At the same time, these sequence motifs would ensure specific targeting of the involved endonuclease(s) to these sites.

In addition to these conserved elements, numerous cryptic processing sites have been documented, which have to be recognized by endonuclease(s) and/or cofactors. Several of these enigmatic sites have been found associated with mitochondrial malfunctions. The most widely known phenotype of mutations in plant mitochondrial genomes manifests as cytoplasmically inherited male sterility. While the vegetative parts (such as leaves, stems and roots) of plants carrying such mitochondrial mutations appear normal by all accounts, pollen development is severely impaired and the plant is thus rendered male sterile. In sunflower (*Helianthus annuus*), *Brassica napus* or maize (*Zea mays*), for example, when expressed aberrant genes created by chance genomic recombination events disturb mitochondrial functionality, this results in malformation of pollen (Singh & Brown 1991; Monéger *et al.* 1994; Cui *et al.* 1996; Liu *et al.* 2001). Nuclear genes can rescue full mitochondrial function by cleaving these transcripts, thus preventing synthesis of the encoded mutant proteins. In citrus, a similar mechanism of masking a deleterious mitochondrial gene has been found (Ohtani *et al.* 2002). Here, citrus plants resistant to a certain fungal toxin have a divergent mode of processing a specific mitochondrial mRNA, which results in an altered protein product from this mitochondrial gene.

5. RNA STABILITY

RNA-internal, endonucleolytic processing will not necessarily lead to complete degradation of the RNA, as evidenced by the tRNAs and rRNAs, but also occurs in many mRNAs. RNA degradation as such is usually initiated at the 3' end of the molecule, and the elements holding exonucleolytic RNases in check and thus controlling stability are most often found in these regions of the RNA.

In plant mitochondria, a number of transcripts terminate in inverted repeat sequences, which can be folded, at least in theory, to form hairpin structures with varying loop sizes (Schuster *et al.* 1986). Usually two unrelated elements are found in tandem, suggesting that this double feature may be important for recognition by the respective *trans*-factor(s). In mitochondrial lysates extended RNA molecules are indeed degraded from their 3' termini up to the double stem-loop structures, resulting in steady-state transcripts ending just downstream of these structures (Dombrowski *et al.* 1997; Kuhn *et al.* 2001). The observed kinetics suggest the involvement of exonuclease(s) rather than endonucleolytic cuts. In chloroplasts, analogous inverted repeats have been found at some RNA termini that also mediate enhanced stability of the RNA. While in the latter organelle candidate proteins recognizing these structures have been functionally described, the *A. thaliana* genome sequence has so far yielded only suggestions for corresponding *trans*-factors in mitochondria, such as RNase II and also PNPase (see also below).

Mitochondria and chloroplasts are also similar in the influence of a homopolymeric nucleotide tract at the 3' end of a given RNA upon RNA stability. Usually, polyA and/or oligoA have been investigated for their effect, and were found to destabilize the upstream RNA sequence (Gagliardi & Leaver 1999; Kuhn *et al.* 2001). This enhanced degradation of the RNA is probably effected mostly by recruiting the respective activity, since a mitochondrial lysate will, *in vitro* at least, rapidly chew through an oligoA tract, but will stop at the hairpin structures. Interestingly, an oligoG stretch will be even more rapidly processed by the mitochondrial exonuclease than the oligoA tail, a puzzling result, especially since most exonucleases identified to date are inhibited by a stretch of G-residues and their velocity will decline having once entered a G-tract (J. Kuhn and S. Binder, unpublished results).

6. RNA SPLICING

The exclusively group II introns in plant mitochondria probably require cofactors for splicing, since all attempts to coax autocatalytic excision have failed, even with the smallest of these intervening sequences. In addition, the *trans*-splicing required for several introns in complex I genes is likely to involve additional factors to hold the upstream and downstream mRNA precursors in position (Bonen & Vogel 2001). Extrapolating from the complex situation of group II intron excision in *Chlamydomonas* chloroplasts, one would expect a large number of specialized and recruited proteins to be involved. To date, none of these admittedly hypothetical cofactors from mitochondria in plants has been clearly identified. Recently, organelle electroporation experiments have revealed some

of the intron internal sequences required for excision; these may be useful in future experiments to bait interacting cofactors (Farré & Araya 2002).

7. RNA EDITING

The full scale of RNA editing in plant mitochondria has been elucidated in *A. thaliana*, where 441 editing sites have been identified in the coding regions (Giegé & Brennicke 1999). These editing events change C to U in the coding regions of mRNAs, several events also having been documented in introns and non-coding leader or trailer sequences. In both *A. thaliana* mitochondria and chloroplasts no U to C changes were found (Bonnard *et al.* 1992). These also appear to be rare in other higher plants, but are relatively common in ferns and mosses (Malek *et al.* 1996). In some species, particularly in hornworts, more than a third of all editing events are U to C changes, which can eliminate genomic in-frame stop codons such as TAA or TGA.

Experiments with the recently developed mitochondrial electroporation system confirmed the inferences about the location of the *cis*-acting specificity elements previously drawn from genomic duplications of gene fragments (Farré *et al.* 2001). The *cis* information resides—at least in the instances investigated so far—upstream of the edited C and in a region of about 20 nucleotides. At one site, the identity of the two residues immediately downstream of the edited nucleotide also influences the editing efficiency.

For chloroplasts an *in vitro* system has been developed, which revealed similar *cis* requirements (Hirose & Sugiura 2001). Data from this as well as from *in vivo* analysis of transgenic plastids, however, showed that recognition of individual sites can be mediated by very variable parameters. For example, one site has been identified that needs more than 200 upstream nucleotides to be edited, while most others require no more than 20 upstream nucleotides. These results confirm the inference drawn from the comparison of RNA-editing sites in *A. thaliana* and other plant species, which were interpreted to indicate that each site is characterized by a very different, unrelated sequence context, without any common denominator (Giegé & Brennicke 1999). This lack of a consensus motif or common feature precludes a single *trans*-acting factor to home in on all editing sites, but suggests a larger number of participating molecules.

Nuclease insensitivity of the *in vitro* editing system in chloroplasts suggests that no RNAs are involved in the active lysate, implying that proteins mediate the specificity as well as contributing the enzymatic activity. The number of such specific proteins required in plastids for only 20–30 editing sites appears feasible and could be encoded by a small family of genes. In mitochondria, however, to specify 441 documented sites in *A. thaliana*, a much larger number of proteins would be required, probably counting in the range of hundred(s). To date only one large family of proteins, the so-called PPR proteins, has been identified in the sequence of the nuclear genome, which encompasses enough genes and in which many members display some of the required features (Small & Peeters 2000). This subgroup of 100–150 PPR proteins is predicted to have organellar import sequences, RNA-binding pockets and domains for protein–protein interaction. The latter

are needed when assuming that the actual enzyme is common for all sites and thus distinct from these *trans*-specificity determinants.

The catalysing enzyme and the biochemical reaction performed in RNA editing are currently unresolved. Several *in vitro* tests of C to U editing with different plants showed that the sugar phosphate backbone of the RNA remains intact, excluding a nucleotide excision–insertion repair mechanism (Rajasekhar & Mulligan 1993; Blanc *et al.* 1995; Yu & Schuster 1995). Since in these experiments radiolabelled UMP is generated from an RNA internal cytosine (labelled in the alpha-position with ^{32}P) after degradation of the RNA chain, a deamination reaction appears to be the best candidate. If a deaminase is involved, this enzyme may have evolved enhanced properties to also catalyse the reverse reaction, unless the C to U and U to C changes are mediated by different enzymes. Alternatively, it is still possible that an altogether different enzyme is involved in both reactions.

Regulation of gene expression may have recruited and taken hold of crucial editing sites such as those creating translation AUG initiation codons from ACG or those eliminating stop codons. To date, the only clear cases in proof of such regulatory interference of RNA editing in plant mitochondria have been documented for an intron and tRNAs, which can only be processed after editing. Excision from the precursor molecule of tRNA^{Phe} by RNases P and Z requires RNA editing, which may thus be a regulatory step limiting the availability of this tRNA species (Marchfelder *et al.* 1996; Maréchal-Drouard *et al.* 1996).

8. CANDIDATE GENES FOR PROCESSING ENZYMES IN THE NUCLEAR GENOME OF *ARABIDOPSIS THALIANA*

The genome sequence of *A. thaliana* (The Arabidopsis Genome Initiative 2000) gives access to genes with similarity to RNA interacting proteins in other organisms, which now can (and have to be) tested for their functional relevance. For example, cytidine deaminase genes have been pulled out of the database and some of them have been subsequently investigated biochemically and molecularly for their potential involvement in mitochondrial or chloroplast RNA editing, but no clear candidate has yet emerged (Faivre-Nitschke *et al.* 1999). Another gene identified by its similarity to a yeast protein codes for a mitochondrial RNA helicase (Gagliardi *et al.* 1999). Other enzymes identified *in silico* and tested experimentally for their mitochondrial localization include predicted proteins with a similarity to RNase II and PNPase (S. Binder, unpublished data). Numerous other similarities can be identified, which must now be investigated for their actual involvement in splicing and other processes.

9. TRANSLATION

Although *in organello* run-on translation in isolated plant mitochondria was developed in pioneering work more than 20 years ago (Forde *et al.* 1978), nothing much further is known about translation and regulation occurring at this step of gene expression. Translation initiation codons are generally not preceded by recognizable Shine–

Dalgarno sequences, but conserved sequence blocks have been described preceding several translational starts (Pring *et al.* 1992). These are, however, not found at all AUGs and it remains unclear how the ribosome is guided to and anchored at the start of translation. The recent identification of translation factors such as EF-Tu may allow easier access to the *cis* elements in the mRNA molecules.

10. PROTEIN TURNOVER

Extensive post-translational culling will be required to eliminate unwanted and superfluous protein products arising from non-stringent translation and insufficient assembly of final products. For example, numerous variants of proteins are apparently synthesized when incompletely edited mRNAs are translated, which is suggested by the observation that these RNAs are found associated with the ribosome (Phreaner *et al.* 1996). However, since only proteins corresponding to completely edited messengers appear to be inserted into the mature respiratory chain complexes and the ribosome (Grohmann *et al.* 1994; Williams *et al.* 1998) and protein products corresponding to incompletely edited mRNAs cannot be detected, these have to be rapidly turned over.

Processing of proteins has been observed in plant mitochondria for cytochrome oxidase subunit 2 (COX2), which undergoes the same presequence maturation steps as its yeast homologue. Other mitochondrial proteins are frequently processed at their amino terminus, where the first residue, the methionine, is removed, presumably by a specific, unidentified enzyme.

A LON-like serine-type protease has been characterized from *A. thaliana*, a homologue of which has been implicated in the rapid turnover of a mitochondrial protein involved in male sterility in the bean *Phaseolus vulgaris* (Sarria *et al.* 1998). The bacterial LON protease degrades defective and short-lived proteins so that 'LONG' forms of these proteins accumulate in the absence of the protease. When added to a mitochondrial extract, an antibody against the plant mitochondrial protease will similarly stabilize the aberrant male-sterility-connected protein and prolong its half-life. In addition to this LON-like enzyme, other proteases are expected to be active in plant mitochondria to remove excessive, aberrant and burnt-out polypeptides.

11. RELATIVE CONTRIBUTIONS OF TRANSCRIPTIONAL AND POST-TRANSCRIPTIONAL REGULATION

The rather flexible genome and gene structures in plant mitochondria involve an apparent mobility of promoters and genes and their combinations. Noticeable particularly between species, but also seen in different ecotypes, genomic rearrangements can connect a different promoter to a given gene without any apparent negative effect. In many instances, several promoters have been found to transcribe one gene, some of these most probably involving such chance connections resulting from recombination events. In the instance of the *cox2* gene in maize mitochondria, sequence duplications created seven partially overlapping tandem promoters in addition to other further upstream

transcription initiation sites (Lupold *et al.* 1999). This tolerance of different promoters and variable numbers of promoters for one gene suggests that further downstream regulatory mechanisms exist which can buffer the resulting variations in transcript synthesis.

A comparative survey of transcriptional activities and steady-state levels of the various RNAs in *A. thaliana* mitochondria indeed revealed little correlation between relative promoter activities and transcript abundances, suggesting extensive post-transcriptional regulation (Giegé *et al.* 2000). Regulation of gene expression at the promoter level may thus be reduced to a general enhancement or dampening of mitochondrial activity rather than controlling individual gene activities.

Post-transcriptional regulation may occur at all levels of RNA maturation, but so far no clear indications for regulatory interference at introns, RNA editing or 5' processing have been obtained. Processing at the 3' end of mRNAs appears to be a major level of post-transcriptional regulation in plant mitochondria, where an interplay of destabilizing polyadenylation and stabilizing secondary structures determines the rate of RNA degradation (Kuhn *et al.* 2001). Unresolved mechanisms must modulate access of the 3' end to the respective *trans*-acting proteins, to yield the observed very divergent half-lives of the transcript populations from individual genes.

12. COORDINATION OF MITOCHONDRIAL AND NUCLEAR GENE EXPRESSION

Nuclear genes for mitochondrial proteins are regulated in concert with the mitochondrial gene activity. A number of studies have shown that, specifically during pollen maturation in growing anthers, expression of both sets of genes is enhanced and failure to do so inadvertently results in failure to develop functional pollen.

In nuclear genes for mitochondrial proteins, sequence elements have been identified which are necessary and sufficient to mediate the anther-specific boost in transcription. Two proteins, (apf1 and arr2, respectively) have been described, which bind to these elements at least in yeast one-hybrid systems (Brennicke *et al.* 1999; Lohrmann *et al.* 2001). However, these proteins may not be the immediate receivers of the developmental signal, since their expression is in turn regulated at their respective promoters, with transcriptional activity paralleling that of the nuclear genes for structural mitochondrial proteins. This opens two possible routes along which information on the next level of control may flow: either a straightforward regulation of their promoters shifts the next level to the respective promoter regulator, or the anther-sequence binding properties of the two proteins apf1 and arr2 are tuned by other interacting factors. Both these steps may act separately or in concert to integrate the developmental signal. Interacting gene expression between nucleus and mitochondria has also been seen with ATPase subunits during anther development in male-sterile and fertility-restored sunflower florets (Monéger *et al.* 1994). At which level mitochondrial and nuclear gene expression are connected and integrated should be accessible by unravelling these next molecular interactors with both nuclear-encoded and mitochondrially located genes and their transcription factors or processing enzymes.

13. SPECIFIC STRATEGIES OF REGULATION HAVE EVOLVED DURING INTEGRATION OF THE ENDOSYMBIOTIC GENOME

The above-reviewed facets of plant-mitochondrial gene expression show in the comparison with bacterial, nuclear, plastid and other mitochondrial systems that unique modes and processes of gene expression, especially in terms of post-transcriptional maturation steps, have evolved in plants. Some of these such as RNA editing, *trans*-splicing of introns, the 3' instability conferred by polyA tracts, the positive stability elements in the form of RNA foldback, etc., are unique to mitochondria in plants and are not found in other mitochondria and bacterial and cytoplasmic systems. Intriguingly, many of these specific strategies are also found in the plastid, suggesting that they evolved in concert for both plant organelles, and/or started in one of the compartments and were subsequently adopted by the other.

Plants have thus developed unique pathways of integrating the mitochondrial endosymbiotic genome, adapting to their unique cellular situation with the additional third genome, the plastid DNA. The molecular identification of the individual RNA maturation steps and the *trans*-acting proteins involved will show to what extent mechanisms and factors were originally developed in the plastid or in the mitochondrion and which organelle later recruited these inventions. Estimations of these origins will derive from primary structural analyses of the factors involved and their distribution in different plant and non-plant species. Analysis of the RNA polymerase has shown analogous molecules having been recruited for mitochondria in animals and fungi, suggesting that this enzyme is a hallmark of an early step of the mitochondrial integration (Gray & Lang 1998). The duplicated gene copy(ies) in plants suggests a secondary adaptation of this imported gene for the plastid compartment. In particular, the dual targeting of one of the RNA polymerase products to both plastid and mitochondrion supports the other evidence for a dominant post-transcriptional control of gene expression in both organelles, which have to respond to and accommodate very different biochemical functions and situations (Hedtkke *et al.* 2000).

Likewise, to determine the origin of the plant-specific mechanism of RNA editing will require more data on the proteins and other cofactors involved. The distribution of this process in the plant kingdom is paralleled, without exception, in plastids and mitochondria thus precluding an easy assignment of a primary location.

14. OUTLOOK

The first steps towards unravelling the various levels of controls have been initiated through the identification of *trans*-acting factors in the mitochondrion as well as in the nucleus. The first prerequisite, the identification of the sites of regulation, still needs to be extended and research is required to identify the precise interferences to give access to further *trans*-acting factors. Particularly interesting will be the identification of connections between biochemical and electrochemical requirements and gene expression. Many more metabolic enzymes, as well as sensors of redox levels, may have side functions connecting

into the regulatory grid of the important post-transcriptional interferences with gene expression.

Research in the authors' laboratories is supported by grants from the DFG, the Rudolph and Clothilde Eberhardt Stiftung and the Fonds der chemischen Industrie.

REFERENCES

- Blanc, V., Litvak, S. & Araya, A. 1995 RNA editing in wheat mitochondria proceeds by a deamination mechanism. *FEBS Lett.* **373**, 56–60.
- Bonen, L. & Vogel, J. 2001 The ins and outs of group II introns. *Trends Genet.* **17**, 322–331.
- Bonnard, G., Gualberto, J. M., Lamattina, L. & Grienemberger, J. M. 1992 RNA editing in plant mitochondria. *Crit. Rev. Plant Sci.* **10**, 503–521.
- Brennicke, A., Zabaleta, E., Dombrowski, S., Hoffmann, M. & Binder, S. 1999 Transcription signals of mitochondrial and nuclear genes for mitochondrial proteins in dicot plants. *J. Hered.* **90**, 345–350.
- Cui, X., Wise, R. P. & Schnable, P. S. 1996 The rf2 nuclear restorer gene of male-sterile T-cytoplasm maize. *Science* **272**, 1334–1336.
- Däschner, K., Couée, I. & Binder, S. 2001 The mitochondrial isovaleryl-coenzyme A dehydrogenase of *Arabidopsis* oxidizes intermediates of leucine and valine catabolism. *Plant Physiol.* **126**, 601–612.
- Dombrowski, S., Brennicke, A. & Binder, S. 1997 3'-Inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. *EMBO J.* **16**, 5069–5076.
- Dombrowski, S., Hoffmann, M., Kuhn, J., Brennicke, A. & Binder, S. 1998 On mitochondrial promoters in *Arabidopsis thaliana* and other flowering plants. In *Plant mitochondria: from gene to function* (ed. I. M. Möller, P. Gardeström, K. Glimelius & E. Glaser), pp. 165–170. Leiden, The Netherlands: Backhuys Publishers.
- Faivre-Nitschke, S. E., Grienemberger, J. M. & Gualberto, J. M. 1999 A prokaryotic-type cytidine deaminase from *Arabidopsis thaliana* gene expression and functional characterization. *Eur. J. Biochem.* **263**, 896–903.
- Farré, J. C. & Araya, A. 2002 RNA splicing in higher plant mitochondria: determination of functional elements in group II intron from a chimeric cox II gene in electroporated wheat mitochondria. *Plant J.* **29**, 203–213.
- Farré, J. C., Leon, G., Jordana, X. & Araya, A. 2001 *cis* Recognition elements in plant mitochondrion RNA editing. *Mol. Cell Biol.* **21**, 6731–6737.
- Forde, B. G., Oliver, R. J. C. & Leaver, C. J. 1978 Variation in mitochondrial translation products associated with male-sterile cytoplasms in maize. *Proc. Natl Acad. Sci. USA* **75**, 3841–3845.
- Gagliardi, D. & Leaver, C. 1999 Polyadenylation accelerates the degradation of the mitochondrial mRNA associated with cytoplasmic male sterility in sunflower. *EMBO J.* **18**, 3757–3766.
- Gagliardi, D., Kuhn, J., Spadinger, U., Brennicke, A., Leaver, C. J. & Binder, S. 1999 An RNA helicase (AtSUV3) is present in *Arabidopsis thaliana* mitochondria. *FEBS Lett.* **458**, 337–342.
- Giegé, P. & Brennicke, A. 1999 RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in open reading frames. *Proc. Natl Acad. Sci. USA* **96**, 15 324–15 329.
- Giegé, P., Hoffmann, M., Binder, S. & Brennicke, A. 2000 RNA degradation buffers asymmetries of transcription in *Arabidopsis* mitochondria. *EMBO Rep.* **1**, 164–170.
- Gray, M. W. 1999 Evolution of organellar genomes. *Curr. Opin. Genet. Dev.* **9**, 678–687.
- Gray, M. W. & Lang, B. F. 1998 Transcription in chloroplasts and mitochondria: a tale of two polymerases. *Trends Microbiol.* **6**, 1–3.
- Grohmann, L., Thieck, O., Herz, U., Schröder, W. & Brennicke, A. 1994 Translation of *nad9* mRNA in mitochondria from *Solanum tuberosum* is restricted to completely edited transcripts. *Nucleic Acids Res.* **22**, 3304–3311.
- Hedtkke, B., Börner, T. & Weihe, A. 2000 One RNA polymerase serving two genomes. *EMBO Rep.* **1**, 435–440.
- Hirose, T. & Sugiura, M. 2001 Involvement of a site-specific trans-acting factor and a common RNA binding protein in the editing of chloroplast mRNAs: development of a chloroplast *in vitro* RNA editing system. *EMBO J.* **20**, 1144–1152.
- Hoffmann, M. & Binder, S. 2002 Functional importance of nucleotide identities within the pea *atp9* mitochondrial promoter sequence. *J. Mol. Biol.* **320**, 943–950.
- Ikeda, T. M. & Gray, M. W. 1999 Characterization of a DNA-binding protein implicated in transcription in wheat mitochondria. *Mol. Cell Biol.* **19**, 8113–8122.
- Kuhn, J., Tengler, U. & Binder, S. 2001 Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria. *Mol. Cell Biol.* **21**, 731–742.
- Kunzmann, A., Brennicke, A. & Marchfelder, A. 1998 5' end maturation and RNA editing have to precede tRNA 3' processing in plant mitochondria. *Proc. Natl Acad. Sci. USA* **95**, 108–113.
- Lang, B. F., Burger, G., O'Kelly, C. J., Cedergren, R., Golding, G. B., Lemieux, C., Sankoff, D., Turmel, M. & Gray, M. W. 1997 An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* **387**, 493–497.
- Lang, B. F., Gray, M. W. & Burger, G. 1999 Mitochondrial genome evolution and the origin of eukaryotes. *A. Rev. Genet.* **33**, 351–397.
- Liu, F., Cui, X., Horner, H. T., Weiner, H. & Schnable, P. S. 2001 Mitochondrial aldehyde dehydrogenase activity is required for male fertility in maize. *Plant Cell* **13**, 1063–1078.
- Lohrmann, J., Sweere, U., Zabaleta, E., Bäurle, I., Keitel, C., Kozma-Bognar, L., Brennicke, A., Schaefer, E., Kudla, J. & Harter, K. 2001 The response regulator ARR2: a pollen-specific transcription factor involved in the expression of nuclear genes for components of mitochondrial Complex I in *Arabidopsis*. *Mol. Gen. Genom.* **265**, 2–13.
- Lupold, D. S., Caoile, A. G. & 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.
- McCulloch, V., Seidel-Rogol, L. & Shadel, G. S. 2002 A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine. *Mol. Cell Biol.* **22**, 1116–1125.
- Malek, O., Lättig, K., Brennicke, A. & Knoop, V. 1996 RNA editing in bryophytes and a molecular phylogeny of land plants. *EMBO J.* **15**, 1403–1411.
- Maloney, A. P., Traynor, P. L., Levings III, C. S. & Walbot, V. 1989 Identification in maize mitochondrial 26S rRNA of a short 5'-end sequence possibly involved in transcription initiation and processing. *Curr. Genet.* **15**, 207–212.
- Marchfelder, A., Schuster, W. & Brennicke, A. 1990 *In vitro* processing of mitochondrial and plastid derived tRNA precursors in a plant mitochondrial extract. *Nucleic Acids Res.* **18**, 1401–1406.
- Marchfelder, A., Brennicke, A. & Binder, S. 1996 RNA editing is required for efficient excision of tRNA^{Phe} from precursors in plant mitochondria. *J. Biol. Chem.* **271**, 1898–1903.
- Maréchal-Drouard, L., Kumar, R., Remacle, C. & Small, I. 1996 RNA editing of larch mitochondrial tRNA^{Phe} precursors is a prerequisite for processing. *Nucleic Acids Res.* **24**, 3229–3234.

- Monéger, F., Smart, C. J. & Leaver, C. J. 1994 Nuclear restoration of cytoplasmic male sterility in sunflower is associated with the tissue-specific regulation of a novel mitochondrial gene. *EMBO J.* **13**, 8–17.
- Oda, K. (and 10 others) 1992 Gene organisation deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. *J. Mol. Biol.* **223**, 1–7.
- Ohtani, K., Yamamoto, H. & Akimitsu, K. 2002 Sensitivity to *Alternaria alternata* toxin in citrus because of altered mitochondrial RNA processing. *Proc. Natl Acad. Sci. USA* **99**, 2439–2444.
- Phreaner, C. G., Williams, M. A. & Mulligan, M. 1996 Incomplete editing of rps12 transcripts results in the synthesis of polymorphic polypeptides in plant mitochondria. *Plant Cell* **8**, 107–117.
- Pring, D. R., Mullen, J. A. & Kempken, F. 1992 Conserved sequence blocks 5' to start codons of plant mitochondrial genes. *Plant Mol. Biol.* **9**, 313–317.
- Race, H. L., Herrmann, R. G. & Martin, W. 1999 Why have organelles retained genomes? *Trends Genet.* **15**, 364–370.
- Rapp, W. D. & 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.
- Rajasekhar, V. K. & Mulligan, R. M. 1993 RNA editing in plant mitochondria: alpha-phosphate is retained during C-to-U conversion in mRNAs. *Plant Cell* **5**, 1843–1852.
- Sarria, R., Lyznik, A., Valejos, C. E. & Mackenzie, S. A. 1998 A cytoplasmic male sterility-associated mitochondrial peptide in common bean is post-translationally regulated. *Plant Cell* **10**, 1217–1228.
- Schiffer, S., Rösch, S. & Marchfelder, A. 2002 Assigning a function to a conserved group of proteins: the tRNA 3'-processing enzymes. *EMBO J.* **21**, 2769–2777.
- Schuster, W. & Brennicke, A. 1989 Conserved sequence elements at putative processing sites. *Curr. Genet.* **15**, 187–192.
- Schuster, W., Hiesel, R., Isaac, P. G., Leaver, C. J. & Brennicke, A. 1986 Transcript termini of messenger RNAs in higher plant mitochondria. *Nucleic Acids Res.* **14**, 5943–5954.
- Singh, M. & Brown, G. G. 1991 Suppression of cytoplasmic male sterility by nuclear genes alters expression of a novel mitochondrial gene region. *Plant Cell* **3**, 1349–1362.
- Small, I. D. & Peeters, N. 2000 The PPR motif—a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem. Sci.* **25**, 46–47.
- The Arabidopsis Genome Initiative 2000 Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.
- Unsel, M., Marienfeld, J. R., Brandt, P. & Brennicke, A. 1997 The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genet.* **15**, 57–61.
- Williams, M. A., Tallakson, W. A., Phreaner, C. G. & Mulligan, R. M. 1998 Editing and translation of ribosomal protein S13 transcripts: unedited translation products are not detectable in maize mitochondria. *Curr. Genet.* **34**, 221–226.
- Yu, W. & Schuster, W. 1995 Evidence for a site-specific cytidine deamination reaction involved in C to U RNA editing of plant mitochondria. *J. Biol. Chem.* **270**, 18 227–18 233.

Discussion

M. Clokie (*Department of Biological Sciences, University of Warwick, Warwick, UK*). I was just wondering why it is that plant mitochondria have evolved so much more slowly than the chloroplast and the nuclear genome.

A. Brennicke. That is a good question. Maybe one rea-

son is that there are so many copies and there is so much recombination, so the different copies have a very good chance of maintaining the sequences. Similarly, you find that the sequences within the inverted repeat of the chloroplast evolved much more slowly than those in the unique regions. RNA editing yields no argument for the slow evolution, because RNA editing can be lost and gained rapidly.

J. C. Gray (*Department of Plant Sciences, University of Cambridge, Cambridge, UK*). What is known about the U to C editing? I mean, it seems to me that that is a much more difficult form of editing to do than to go the other way around.

A. Brennicke. Since we do not even know the biochemistry of the C to U editing, which could be done by deamination, involving a deaminase, as in the apolipoprotein B editing. However, these deaminases are not able to catalyse the reverse reaction, which is much more complicated, needs energy and about which we do not know anything at present. *In vitro*, we do need ATP—a fairly high level of ATP—for the deamination reaction and we have no clue about how the reverse reaction works. There could be a second enzyme involved if the C to U editing is indeed a deaminase. However, it is unlikely that two enzymes have evolved at the same time, so I like to think a transaminase or something like it is involved, but so far there is no evidence for or against that.

D. H. Williamson (*Parasitology Division, National Institute of Medical Research, Mill Hill, London, UK*). I am very interested in this RNA editing. The trypanosomal-kinetoplast system uses a lot of guide RNAs. Is there any evidence for guide RNAs here, and could the repeats you get in the plant mitochondrial genome, which you mentioned at the beginning of your talk, be a source of guide RNAs, or is the system totally different?

A. Brennicke. Basically the system is totally different, because we have U insertion and excision in the trypanosome, whereas here we have the alteration from C to U. So far, we do not have any evidence for or against guide RNAs. Actually there is evidence that the sugar phosphate backbone is not cut, so RNA editing in plants seems to be a deamination or amination mechanism. To take the second part of your question, there is indeed enough space on the genome, and there are enough repeats to potentially code for guide RNAs. There are numerous bits and pieces of genes repeated that could very well code for guide RNAs, but in the chloroplast *in vitro* system that Hirose and Sugiura developed, micrococcal nuclease digestion gave no evidence that an RNA is involved there. Of course guide RNA could always be hidden in proteins, but it seems probable that no guide RNA is involved. On the other hand, a protein-only identification system of *cis* elements may be feasible in chloroplasts where you only have 30 editing sites, but if you have 450 editing sites, as in plant mitochondria, you would need a lot of different proteins.

R. J. M. Wilson (*Division of Parasitology, National Institute of Medical Research, Mill Hill*). I thought I saw in one of your diagrams you had a pseudo-gene that had no editing. Is that correct?

A. Brennicke. There are indeed numerous pseudo-genes, copies of genes that are incomplete. These are probably not real genes, although some are transcribed; of

these pseudo-genes some show no editing, but others do have RNA editing. From these transcribed duplications, we can deduce some of the parameters of RNA editing, one being that the sequence 5' to an editing site is important, while the sequence downstream three or four nucleotides is not. For example, two introns contain the same sequence and the same editing site. They vary three nucleotides downstream, but have 40 identical nucleotides upstream. This is similar to what Sugiura's group found in chloroplasts. Thus 20, 30 or 40 nucleotides upstream, but only very few nucleotides downstream are important. This is very different from the trypanosome system, where you have the guide RNA home in on the downstream side, and also from the apolipoprotein B RNA editing, where the editing enzyme moors downstream of the editing site.

A. E. Douglas (*Department of Biology, University of York, York, UK*). Regarding the selection pressure for this RNA editing, is it simply a complex way of solving what is fundamentally bad design?

A. Brennicke. Probably not, because what could RNA editing solve? It would solve the problem if you do not want Ts in your genome or if you want to avoid specific dinucleotides. TT-dinucleotides are fragile to UV radiation, so we screened the genome to see if there is a higher frequency of these around these editing sites, but there is not. This is the only potentially useful thing we can think of; otherwise I think that RNA editing is just to make life

complicated. According to Mike Gray, initially there was a mutation in an enzyme like a deaminase, which enabled it to change Cs to Us in the mRNA; and once this was established, the genome was free to interchange T and C. Once such a mutation has occurred in the genome, you cannot get rid of RNA editing any more. Subsequently more genome sequences mutate and RNA editing is established. Another reason to keep RNA editing is to keep genes in the mitochondrial genome, so that they do not move to the nucleus. However, in plants you very often find that genes do move to the nucleus in spite of RNA editing, which may be easily done going through the RNA. If you take an edited mRNA, do a reverse transcription step, and then insert the cDNA sequence into the nucleus, you have a fully edited genomic sequence. This has apparently happened many times. For example, from one plant species to the next, a gene like *cox2* has moved from the mitochondrial gene into the nucleus, lost all editing, lost all mitochondrial introns, gained pre-sequence and proper promoters and now is a functional gene in the nucleus.

GLOSSARY

ATPase: adenosine triphosphatase
 PNPase: polynucleotidephosphorylase
 PPR: pentatricopeptide repeat
 RNase: ribonuclease